



FACULTADE DE MEDICINA  
E ODONTOLOXÍA

**TRABALLO DE  
FIN DE GRAO**

DIAGNÓSTICO GENÉTICO DEL TRASTORNO DEL ESPECTRO AUTISTA (TEA): ANÁLISIS MEDIANTE SECUENCIACIÓN DEL GENOMA COMPLETO.

DIAGNÓSTICO XENÉTICO DO TRASTORNO DO ESPECTRO AUTSTA (TEA): ANÁLISE MEDIANTE SECUENCIACIÓN DO XENOMA COMPLETO.

GENETIC DIAGNOSIS OF AUTISM SPECTRUM DISORDER (ASD): WHOLE GENOME SEQUENCING ANALYSIS.

**AUTOR: Ecénarro del Río, Luis**

**TITOR: Carracedo Álvarez, Ángel María**

**COTITOR: Rodríguez López, Julio**

**DEPARTAMENTO: Ciencias Forenses, Anatomía Patológica,  
Ginecología y Obstetricia, y Pediatría.**

Curso académico: 2023-2024. Convocatoria: Xuño

## ÍNDICE

<b>1. Introducción. Los Trastornos del Espectro Autista.</b> .....	<b>7</b>
<b>1.1 Definición</b> .....	<b>7</b>
<b>1.2. Criterios Diagnósticos de los TEA (DSM V)</b> .....	<b>8</b>
<b>1.3 Epidemiología</b> .....	<b>11</b>
<b>1.4 Clínica y detección precoz de los TEA</b> .....	<b>12</b>
<b>1.5 Etiología de los TEA</b> .....	<b>14</b>
<b>1.5.1 Factores ambientales.</b> .....	<b>14</b>
<b>1.5.2 Factores genéticos. Variantes genéticas.</b> .....	<b>16</b>
<b>1.5.3 Diagnóstico genético en el TEA. El análisis de genoma completo.</b> .....	<b>18</b>
<b>2. Objetivos y plan de trabajo</b> .....	<b>20</b>
<b>3. Material y métodos</b> .....	<b>21</b>
<b>3.1 Parámetros de calidad</b> .....	<b>21</b>
<b>3.2 Filtrado interno</b> .....	<b>22</b>
<b>3.3 Efecto molecular sobre la transcripción del gen.</b> .....	<b>22</b>
<b>3.4 Frecuencias poblacionales</b> .....	<b>24</b>
<b>3.5 Por fenotipo. Bases de datos de enfermedad y fenotipos clínicos.</b> .....	<b>24</b>
<b>3.6 Modo de herencia de las variantes y zigosis. Trío genético</b> .....	<b>25</b>
<b>3.7 Clasificación de variantes según los criterios ACGM y ACGS.</b> .....	<b>26</b>
<b>3.8 Método de trabajo</b> .....	<b>28</b>
<b>4. Resultados</b> .....	<b>30</b>
<b>4.1 Probando 1</b> .....	<b>30</b>
<b>4.2 Probando 2</b> .....	<b>33</b>
<b>4.3 Probando 3</b> .....	<b>34</b>
<b>4.4 Probando 4</b> .....	<b>37</b>
<b>4.5 Probando 5</b> .....	<b>40</b>
<b>5. Discusión</b> .....	<b>41</b>
<b>6. Conclusiones</b> .....	<b>45</b>
<b>7. Referencias bibliográficas</b> .....	<b>47</b>

## Índice de abreviaturas y siglas

<b>TEA:</b> Trastorno del Espectro Autista.	<b>CMA:</b> <i>Cromosomal Micro-Array</i>
<b>ASD:</b> <i>Autism Spectrum Disorder</i>	<b>G:</b> Guanina
<b>WES:</b> <i>Whole Exome Sequencing</i>	<b>C:</b> Citosina
<b>WGS:</b> <i>Whole Genome Sequencing</i>	<b>A:</b> Adenosina
<b>DSM:</b> <i>Diagnostic and Statistical Manual of Mental Disorders.</i>	<b>U:</b> Uracilo
<b>TGD-NOS:</b> trastorno generalizado del desarrollo no especificado	<b>T:</b> Timina
<b>CIE:</b> Clasificación internacional de enfermedades.	<b>SV:</b> Structural Variants
<b>CDC:</b> <i>Centers for Disease Control and Prevention</i>	<b>UPD:</b> disomías uniparentales
<b>MMRW:</b> <i>Morbidity and Mortality Weekly report</i>	<b>STR/TRE:</b> <i>short tandem repeats/ Expansiones y Repeticiones en tandem</i>
<b>MFOR:</b> <i>Male Female Odds Ratio</i>	<b>del:</b> <i>delección</i>
<b>ADDM:</b> Red de Monitoreo de Autismo y Discapacidades del Desarrollo	<b>dup:</b> <i>duplicación</i>
<b>T.O.C:</b> Trastorno Obsesivo Compulsivo	<b>indel:</b> <i>inserción o delección</i>
<b>TDAH:</b> Trastornos por déficit de atención e hiperactividad	<b>Emedgene:</b> <i>Software Emedgene Analyze (Illumina, inc.)</i>
<b>AEPed:</b> asociación Española de Pediatría	<b>Missense:</b> Variante de sentido errado
<b>M-CHAT:</b> Cuestionario Modificado de Detección Temprana de Autismo	<b>Nonsense:</b> Variante sin sentido
<b>CAST:</b> Test Infantil del Síndrome de Asperger	<b>Frameshift:</b> Variante de cambio de marco de lectura
<b>VPA:</b> <i>Valproic Acid/ Ácido valproico</i>	<b>5'3' UTR:</b> <i>5'3' untranslated regions</i>
<b>SNP/ SNV:</b> <i>Single Nucleotide Polimorfism/ Single Nucleotide Variant</i>	<b>pLI:</b> <i>probability of loss-of-function intolerance</i>
<b>CNV:</b> <i>Copy Number Variant</i>	<b>ExAC:</b> <i>Exome Aggregation Consortium</i>
<b>MAF:</b> <i>Minor allele frequency</i>	<b>P:</b> Patogénica ( <i>pathogenic</i> )
<b>ADN:</b> Ácido desoxirribonucleico	<b>LP:</b> Probablemente patogénica ( <i>likely pathogenic</i> )
<b>ADNmt:</b> Ácido desoxirribonucleico mitocondrial	<b>VUS:</b> Variante de significado incierto ( <i>variant of uncertain significance</i> )
<b>ARN:</b> Ácido ribonucleico	<b>LB:</b> Probablemente benigna ( <i>likely benign</i> )
<b>ARNm:</b> Ácido ribonucleico mensajero	<b>B:</b> Benigna ( <i>benign</i> )
<b>microARN:</b> micro Ácido ribonucleico	<b>PVS:</b> <i>Pathogenic very Strong</i>
<b>OMIM:</b> Online Mendelian Inheritance in Man	<b>PS:</b> <i>Pathogenic Strong</i>
<b>HPO:</b> Human Phenotype Ontology	<b>PM:</b> <i>Pathogenic Moderate</i>
<b>PanelApp:</b> Genomics England PanelApp	<b>PP:</b> <i>Pathogenic Supporting</i>
<b>De novo:</b> En latín "de nuevo"	<b>BA:</b> <i>Benign Stand Alone</i>
<b>In silico:</b> "Hecho por computadora o vía simulación computacional".	<b>BS:</b> <i>Benign Strong</i>
<b>ACMG:</b> American College of Medical Genetics and Genomics	<b>BP:</b> <i>Benign Supporting.</i>
<b>AMP:</b> Association for Molecular Pathology	<b>AI/IA:</b> <i>Artificial Intelligence/ Inteligencia artificial</i>
<b>ACGS:</b> The Association for Clinical Genomic Science	<b>AD:</b> Herencia autonómica dominante
<b>Array-CGH:</b> <i>Array-Comparative Genomic Hybridization</i>	<b>AR:</b> Herencia autonómica recesiva
	<b>NMD:</b> <i>Nonsense Mediated Decay</i>
	<b>Leu:</b> Leucina
	<b>Val:</b> Valina
	<b>Phe:</b> Fenilalanina
	<b>Thr:</b> Treonina
	<b>Gly:</b> Glicina
	<b>Glu:</b> Glutamato/ Ácido glutámico
	<b>IRDIRC:</b> Consorcio Internacional de enfermedades raras

## Resumen

**ANTECEDENTES:** Los trastornos del espectro autista (TEA) son un grupo de afecciones diversas, relacionadas con el neurodesarrollo y cuyo diagnóstico se establece en base a unos criterios clínicos. Constituyen, actualmente, una de las enfermedades con mayor importancia en la investigación debido al notable aumento en su incidencia en los últimos años y al impacto que suponen no sólo en la vida de los pacientes sino también en sus familias y entorno. Aunque se trata de una enfermedad de etiología compleja y múltiple, sabemos que la genética juega un papel crucial en estos trastornos, de hecho, se considera que es el trastorno del neurodesarrollo con mayor heredabilidad (83%).

**OBJETIVO:** El principal objetivo de este trabajo es realizar un análisis de genoma completo de cinco tríos de TEA (probando afectado y progenitores sanos) con el objetivo encontrar una variante causal que nos proporcione el diagnóstico genético. Junto a ello, se quiere comprobar si el análisis del genoma completo en tríos nos podría ofrecer una mayor eficacia, fiabilidad y rapidez diagnóstica como prueba de primer nivel en el diagnóstico de TEA, analizando, para ello, el caso de un paciente que ha sido estudiado mediante WES (*whole exome sequencing*) previamente con resultado negativo.

**MÉTODOS:** En este trabajo, partimos de cinco tríos genéticos, es decir, cinco pacientes con un diagnóstico clínico de TEA (probandos afectados) y sus respectivos padre y madre (progenitores sanos o afectados). De cada probando afecto, tenemos el conjunto de las variantes de su genoma, las cuales priorizaremos, filtraremos y anotaremos funcionalmente gracias a la herramienta software *Emedgene Analyze* (Illumina, inc.), con el objetivo de reducir el número y quedarnos con variantes candidatas para analizarlas en detalle y encontrar la variante o variantes causales si es que existen. Una vez obtenido un número manejable de variantes candidatas, llevaremos a cabo el proceso de estudio y selección de estas basándonos en criterios como la clasificación de patogenicidad, herencia, compatibilidad del fenotipo con el descrito en la literatura y bases de datos etc., con el objetivo de llegar al diagnóstico genético final.

**RESULTADOS:** Se ha conseguido llegar al diagnóstico genético en cuatro de los cinco probandos afectados, incluyendo aquel con un resultado negativo en un estudio previo de WES (*whole exome sequencing*); encontrado y caracterizando la variante causal del fenotipo clínico en cada uno de estos casos. En el otro probando, no se ha conseguido encontrar la variante causal, sin que ello implique que no exista. La aparición, en el futuro, tanto de nuevos datos del paciente, como de avances tecnológicos y de investigación podrían aportar más datos al estudio de este caso.

## Resumo

**ANTECEDENTES:** Os trastornos do espectro autista (TEA) son un grupo de afeccións diversas, relacionadas co neurodesarrollo e cuxo diagnóstico se establece en base a uns criterios clínicos. Constitúen, actualmente, unha das enfermidades con maior importancia na investigación debido ao notable aumento na súa incidencia nos últimos anos e ao impacto que supoñen non só na vida dos pacientes senón tamén nas súas familias e contorna. Aínda que se trata dunha enfermidade de etioloxía complexa e múltiple, sabemos que a xenética xoga un papel crucial nestes trastornos, de feito, considérase que é o trastorno do neurodesarrollo con maior heredabilidade (83%).

**OBXECTIVO:** O principal obxectivo deste traballo é realizar unha análise de xenoma completo de cinco tríos de TEA (probando afectado e proxenitores sans) co obxectivo atopar unha variante causal que nos proporcione o diagnóstico xenético. Xunto a iso, quérese comprobar se a análise do xenoma completo en tríos poderíanos ofrecer unha maior eficacia, fiabilidade e rapidez diagnóstica como proba de primeiro nivel no diagnóstico de TEA, analizando, para iso, o caso dun paciente que foi estudado mediante WES (*whole exome secuencing*) previamente con resultado negativo.

**MÉTODOS:** Neste traballo, partimos de cinco tríos xenéticos, é dicir, cinco pacientes cun diagnóstico clínico de TEA (probandos afectados) e o seu respectivos pai e nai (proxenitores sans ou afectados). De cada probando afecto, temos o conxunto das variantes do seu xenoma, as cales priorizaremos, filtraremos e anotaremos funcionalmente grazas á ferramenta software Emedgene Analyze (Illumina, inc.), co obxectivo de reducir o número e quedarnos con variantes candidatas para analizalas en detalle e atopar a variante ou variantes causais se é que existen. Unha vez obtido un número manexable de variantes candidatas, levaremos a cabo o proceso de estudo e selección destas baseándonos en criterios como a clasificación de patoxenicidade, herdanza, compatibilidade do fenotipo co descrito na literatura e bases de datos etc., co obxectivo de chegar ao diagnóstico xenético final.

**RESULTADOS:** Conseguiuse chegar ao diagnóstico xenético en catro dos cinco probandos afectados, incluíndo aquel con un resultado negativo nun estudo previo de WES (*whole exome sequencing*); atopado e caracterizando a variante causal do fenotipo clínico en cada un destes casos. No outro probando, non se conseguiu atopar a variante causal, sen que iso implique que non exista. A aparición, no futuro, tanto de novos datos do paciente, como de avances tecnolóxicos e de investigación poderían achegar máis datos ao estudo deste caso.

## Abstract

**BACKGROUND:** Autism Spectrum Disorders (ASD) are a group of diverse conditions related to neurodevelopment and whose diagnosis is established based on clinical criteria. They are currently one of the most important diseases in research due to the notable increase in their incidence in recent years and the impact they have not only on the patients' lives but also on their families and environment. Although it is a disease of complex and multiple etiology, we know that genetics plays a crucial role in these disorders, in fact, it is considered to be the neurodevelopmental disorder with the highest heritability (83%).

**OBJECTIVE:** The main objective of this work is to perform a whole genome analysis of five ASD trios (testing affected and healthy parents) with the aim of finding a causal variant that will provide us with the genetic diagnosis. Together with this, we want to check if the analysis of the complete genome in trios could offer us greater efficiency, reliability, and diagnostic speed as a first level test in the diagnosis of ASD, analyzing, for this purpose, the case of a patient who has been studied by WES (whole exome sequencing) previously with negative results.

**METHODS:** In this project, we started from five genetic trios, that is, five patients with a clinical diagnosis of ASD (affected probands) and their respective fathers and mothers (healthy or affected parents). For each affected proband, we have the set of variants of their genome, which we will prioritize, filter and functionally annotate thanks to the Emedgene Analyze software tool (Illumina, inc.), with the aim of reducing the number and keeping candidate variants to analyze them in detail and find the causal variant(s) if they exist. Once we have obtained a manageable number of candidate variants, we will carry out the process of study and selection of these based on criteria such as classification of pathogenicity, inheritance, compatibility of the phenotype with that described in the literature and databases, etc., with the aim of arriving at the final genetic diagnosis.

**RESULTS:** Genetic diagnosis has been achieved in four of the five affected probands, including the one with a negative result in a previous WES (whole exome sequencing) study, finding and characterizing the causal variant of the clinical phenotype in each of these cases. In the other proband, the causal variant has not been found, although this does not imply that it does not exist. The appearance, eventually, of new patient data, as well as technological and research advances, could contribute more data to the study of this case.

# 1. Introducción. Los Trastornos del Espectro Autista.

## 1.1 Definición

Los trastornos del espectro autista, TEA (ASD en inglés por “Autism Spectrum Disorders”) son un grupo de afecciones diversas, relacionadas con el neurodesarrollo. Se caracterizan por algún grado de dificultad en la interacción social y la comunicación, así como por patrones atípicos de actividad y comportamiento en los pacientes que los padecen; por ejemplo, dificultad para pasar de una actividad a otra, gran atención a los detalles y reacciones poco habituales a las sensaciones. (1)

Aunque los individuos con TEA son muy diferentes entre sí, el trastorno se caracteriza por rasgos básicos en dos áreas: comunicación social y conductas sensitivo-motoras restringidas y repetitivas, independientemente de la cultura, la raza, la etnia o el grupo socioeconómico.

Actualmente está ampliamente aceptado que este trastorno es de etiología compleja y múltiple, además de poseer diversos subtipos y trayectorias evolutivas. (2)

Factores ambientales, genéticos y epigenéticos tienen su importancia en la patogénesis de este grupo de trastornos, y algunos de ellos permanecen todavía desconocidos. Aun así, en la última década se han hecho grandes progresos en la investigación de la naturaleza de esta enfermedad. A pesar de que estos avances aún no han tenido un impacto transformador en la práctica clínica, hay motivos para el optimismo: las listas fiables de genes de riesgo son amplias y crecen rápidamente; las proteínas codificadas identificadas ya han empezado a apuntar a un número relativamente pequeño de áreas de la biología, en las que los avances paralelos en neurociencia y genómica funcional están aportando conocimientos profundos. (3)

Los trastornos del espectro autista constituyen en la actualidad una de las enfermedades con mayor importancia en la investigación debido al notable aumento en su incidencia en los últimos años y al impacto que suponen no sólo en la vida de los pacientes sino también en sus familias y entorno.

Se trata de una entidad clínicamente heterogénea ya que en ella incluimos tanto a individuos con problemas menores durante su neurodesarrollo, los cuales pueden llegar a realizar una vida normal en su vida adulta, como aquellos con síntomas graves e inhabilitantes, con diversas comorbilidades físicas y mentales. De ahí que utilicemos la palabra espectro en su definición, ya que se trata de un amplio rango que abarca desde cuadros clínicos leves hasta pacientes con sintomatología muy grave, todos bajo unos criterios clínicos establecidos desde hace años por el DSM (*“Diagnostic and Statistical Manual of Mental Disorders”*).

El paso al DSM-V (*“Diagnostic and Statistical Manual of Mental Disorders”, Fifth edition*), publicado en 2013, estuvo marcado por la ampliación de la definición y la reducción de la especificidad de los síntomas relacionados con el autismo, lo que supuso cambios sustanciales en los criterios diagnósticos. Los diagnósticos de trastorno autista, síndrome de Asperger y trastorno generalizado del desarrollo no especificado (TGD-NOS) se eliminaron como clasificaciones diagnósticas y se agruparon en dos diagnósticos: trastorno del espectro autista y trastorno de la comunicación social. (4)

## 1.2. Criterios Diagnósticos de los TEA (DSM V)

El inicio de los síntomas del TEA suele producirse a los 3 años, aunque los síntomas pueden no manifestarse plenamente hasta la edad escolar o más tarde, y algunas investigaciones sugieren que los síntomas pueden surgir entre los 6 y los 18 meses (5). Los niños más gravemente afectados o con comorbilidad asociada tienen más probabilidades de ser identificados y diagnosticados de forma fiable a edades más tempranas que los casos más leves. (6)

<b>Criterios Diagnósticos de los TEA (DSM V)</b>
<b>A. Deficiencias persistentes en la comunicación social y en la interacción social en diversos contextos, manifestado por lo siguiente, actualmente o por los antecedentes (los ejemplos son ilustrativos, pero no exhaustivos):</b>
1. Las deficiencias en la reciprocidad socioemocional, varían, por ejemplo, desde un acercamiento social anormal y fracaso de la conversación normal en ambos sentidos pasando por la disminución en intereses, emociones o afectos compartidos hasta el fracaso en iniciar o responder a interacciones sociales.
2. Las deficiencias en las conductas comunicativas no verbales utilizadas en la interacción social, varían, por ejemplo, desde una comunicación verbal y no verbal poco integrada pasando por anomalías del contacto visual y del lenguaje corporal o deficiencias de la comprensión y el uso de gestos, hasta una falta total de expresión facial y de comunicación no verbal.
3. Las deficiencias en el desarrollo, mantenimiento y comprensión de las relaciones, varían, por ejemplo, desde dificultades para ajustar el comportamiento en diversos contextos sociales pasando por dificultades para compartir juegos imaginativos o para hacer amigos, hasta la ausencia de interés por otras personas.
<b>B. Patrones restrictivos y repetitivos de comportamiento, intereses o actividades, que se manifiestan en dos o más de los siguientes puntos, actualmente o por los antecedentes (los ejemplos son ilustrativos, pero no exhaustivos):</b>
1. Movimientos, utilización de objetos o habla estereotipados o repetitivos (p. ej., estereotipias motoras simples, alineación de los juguetes o cambio de lugar de los objetos, ecolalia, frases idiosincrásicas).

<p>2. Insistencia en la monotonía, excesiva inflexibilidad de rutinas o patrones ritualizados de comportamiento verbal o no verbal (p. ej., gran angustia frente a cambios pequeños, dificultades con las transiciones, patrones de pensamiento rígidos, rituales de saludo, necesidad de tomar el mismo camino o de comer los mismos alimentos cada día).</p>
<p>3. Intereses muy restringidos y fijos que son anormales en cuanto a su intensidad o foco de interés (p. ej., fuerte apego o preocupación por objetos inusuales, intereses excesivamente circunscritos o perseverantes).</p>
<p>4. Hiper o hiporreactividad a los estímulos sensoriales o interés inhabitual por aspectos sensoriales del entorno (p. ej., indiferencia aparente al dolor/temperatura, respuesta adversa a sonidos o texturas específicos, olfateo o palpación excesiva de objetos, fascinación visual por las luces o el movimiento).</p>
<p><b>C. Los síntomas han de estar presentes en las primeras fases del período de desarrollo (pero pueden no manifestarse totalmente hasta que la demanda social supera las capacidades limitadas, o pueden estar enmascarados por estrategias aprendidas en fases posteriores de la vida).</b></p>
<p><b>D. Los síntomas causan un deterioro clínicamente significativo en lo social, laboral u otras áreas importantes del funcionamiento habitual.</b></p>
<p><b>E. Estas alteraciones no se explican mejor por la discapacidad intelectual (trastorno del desarrollo intelectual) o por el retraso global del desarrollo. La discapacidad intelectual y el trastorno del espectro del autismo con frecuencia coinciden; para hacer diagnósticos de comorbilidades de un trastorno del espectro del autismo y discapacidad intelectual, la comunicación social ha de estar por debajo de lo previsto para el nivel general de desarrollo.</b></p>

Tabla 1: Criterios diagnósticos del Trastorno del Espectro Autista según la DSM V.

Junto con estos criterios diagnósticos, se ha de especificar si va acompañado de los siguientes:

1. Con o sin déficit intelectual acompañante
2. Con o sin deterioro del lenguaje acompañante.
3. Asociado a una afección médica o genética, o a un factor ambiental conocidos
4. Asociado a otro trastorno del desarrollo neurológico, mental o del comportamiento.
5. Con catatonía. (7)

Estos anteriores, también llamados modificadores de la enfermedad son ciertamente importantes en lo que se refiere al pronóstico. En concreto, hay evidencia de que aquellos pacientes con diagnóstico de TEA acompañado de déficit intelectual y deterioro del lenguaje (casos menos prevalentes) van a presentar una peor evolución en cuanto a la gravedad del cuadro (tabla 2).

Para especificar la gravedad del cuadro, la DSM establece tres grados basándose en el nivel de deterioro de la comunicación social y en los patrones de comportamiento restringidos y repetitivos que presente el paciente

<b>Nivel de gravedad</b>	<b>Comunicación social</b>	<b>Comportamientos restringidos y repetitivos</b>
<b>Dentro de la normalidad</b>	Peculiar o aislado pero carente de interferencias	No interferencia
<b>Síntomas subclínicos</b>	Ciertos síntomas en esta o ambas dimensiones, pero sin alteraciones significativas	Presencia de atípico o excesivo interés sin que llegue a interferir
<b>Grado 1 “necesita ayuda”</b>	Alteraciones significativas, pero sin apoyo in situ	Interferencia significativa en uno o más contextos
<b>Grado 2 “necesita ayuda notable”</b>	Marcado déficit: respuestas breves o atípicas, limitada iniciación	Inflexibilidad y dificultades del cambio del foco que interfieren con frecuencia
<b>Grado 3 “necesita ayuda muy notable”</b>	Mínima comunicación social	Notable alteración en la vida diaria por inflexibilidad y dificultades de cambio y fijación de la atención

Tabla 2: Niveles de gravedad del Trastorno del espectro autista según la DSM V. (7)

Atendiendo al manual imperante en Europa, la CIE-10 (*“The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines”*) se codificaría de forma semejante, teniendo en cuenta también dichos especificadores. En este manual se codificaría con código 6A02, con diferentes subtipos atendiendo a especificadores. (8)

Una revisión reciente en Cochrane destaca una buena estabilidad diagnóstica, puesto que 9 de cada 10 niños diagnosticados con un TEA, antes de los 6 años, seguían cumpliendo los criterios diagnósticos pasado un año. (6)

### **1.3 Epidemiología**

La Organización Mundial de la Salud estimó que el 0,76% de los niños del mundo tenían TEA en 2010, aunque esta estimación se basó en estudios realizados en países que representan sólo el 16% de la población infantil mundial. Otras revisiones sistemáticas de estudios de prevalencia a nivel internacional han producido estimaciones resumidas similares, de aproximadamente el 0,7%. En Estados Unidos, el CDC (Centres for Disease Control and Prevention) estimaba en 2012 que aproximadamente el 1,5% de los niños de 8 años tenían un TEA; sin embargo, en su último informe de MMRW (Morbidity and Mortality Weekly report), publicado en marzo de 2023 nos habla de una prevalencia total calculada del 2,76% de los niños de 8 años (uno de cada 36 niños de esta edad padecería este trastorno). A su vez, este informe documenta un cambio de tendencia respecto al grupo étnico al que pertenecen los niños diagnosticados; aunque desde el año 2000, la prevalencia de TEA ha aumentado constantemente entre todos los grupos, durante 2018-2020, los aumentos fueron mayores para los niños negros e hispanos que para los niños blancos o caucásicos. (9)

Sin embargo, las estimaciones resumidas ocultan una variabilidad considerable en función de la geografía, el enfoque metodológico y el tiempo. La estimación internacional de prevalencia más alta reciente fue del 2,64%, para niños de 7 a 12 años en Corea del Sur en 2005-2009. (10)

Respecto a su distribución según el sexo, hay evidencia de que es un trastorno más frecuente en varones. Un metaanálisis de estudios de prevalencia, publicado en 2017 donde se analizan 54 estudios, con 13784284 participantes, de los cuales 53712 tenían TEA (43972 niños y 9740 niñas), demostró que la Odds Ratio hombres-mujeres (MFOR) global agrupado fue de 4,20 (IC 95%: 3,84-4,60). (11)

Se cree que una razón para que haya una diferencia tan marcada en la prevalencia de TEA entre varones y mujeres es que existe en estas últimas un fenómeno de camuflaje, que consiste en comportamientos complejos de copia y/o enmascaramiento de algunos rasgos de la personalidad que favorece la adaptación a las exigencias específicas del entorno. Esto se produce más frecuentemente en las niñas que presentan rasgos de comportamiento autista, y por ello las diferencias se acortan cuando comparamos sexos en edades más avanzadas. Esta hipótesis ha sido comprobada en una revisión sistemática con resultados positivos por Tubío-Fungueiriño M. y col. (12)

Además de los datos de prevalencia en los últimos años, es importante recalcar el incremento de incidencia y prevalencia que están teniendo los TEA, el cual ya podíamos apreciar en el último informe del CDC que advertía un incremento de más de un punto en la prevalencia en Estados Unidos. Pero esta tendencia se ve reflejada igualmente en numerosos estudios, incluidos a nivel nacional.

En España, se han realizado varios estudios epidemiológicos de autismo y se ha encontrado una prevalencia del 0,61% en preescolares, en un estudio realizado en las Islas Canarias, y otro estudio realizado en Tarragona encontró una prevalencia de 1,55% en preescolares y del 1% en niños en educación primaria. En Cataluña, un estudio de incidencia y prevalencia del diagnóstico de autismo en niños entre 2-17 años, utilizando datos administrativos del sistema público de salud encontró una prevalencia del: 1,23% en 2017, 1,95% para chicos y 0,46% para chicas. La incidencia del diagnóstico de autismo se había incrementado del 0,07% en 2009, al 0,23% en el 2017. (13)

Una revisión de los datos de 2000 a 2014 sobre niños de ocho años en comunidades seleccionadas de EE. UU. realizada por la Red de Monitoreo de Autismo y Discapacidades del Desarrollo (ADDM) de los CDC mostró que las estimaciones de prevalencia de TEA aumentaron más del 150% durante este período, de 6,7 por 1.000 en 2000 a 16,8 por 1.000 en 2014. (14)

La causa de este marcado incremento está todavía por determinar. No obstante, es muy probable que una parte guarde relación con los cambios en los criterios diagnósticos (DSM V en 2013), una mayor concienciación social y un mayor conocimiento por parte de los profesionales sanitarios que deriva en mejores herramientas de cribado y diagnóstico.

## **1.4 Clínica y detección precoz de los TEA**

Como se ha visto con anterioridad, los signos y síntomas clásicos de los TEA se agrupan en dos grandes esferas del neurodesarrollo. A continuación, vemos un resumen de algunos de los más importantes que se engloban en cada una de ellas:

-Alteraciones de la comunicación e interacción social:

- Disminución del juego simbólico, empatía, reciprocidad socio emocional y las relaciones con los iguales.
- Atención selectiva a ciertos aspectos particulares de los estímulos. Por ejemplo, al hacer un puzle se fijan generalmente en la forma y no en la imagen general del puzle.
- La percepción de los detalles está aumentada y la percepción de los movimientos globales está disminuida.
- En general usan las expresiones faciales, los gestos y el contacto visual en menor medida que los iguales.
- Dificultades en la comunicación con fallo del lenguaje hablado no compensado por otras formas de comunicación, ecolalia.
- Lenguaje estereotipado, restrictivo e idiosincrásico.
- Escasa variedad de imitación espontánea y de juego funcional y simbólico

-Patrones de conducta restringidos, repetitivos y estereotipados:

- Preocupaciones estereotipadas e intereses restrictivos.
- Adherencia inflexible a rutinas; las transiciones pueden representar un grave problema y pequeños cambios en su día a día pueden suponer para ellos graves problemas.
- Manierismos motores, estereotipias motoras. No solo son exclusivas de los TEA, también se pueden dar en T.O.C, síndrome de Tourette, Rett.
- Preocupaciones persistentes con partes de objetos (rueditas de los coches, ojos de las muñecas, etc.). (15)

Las manifestaciones clínicas de los TEA pueden cambiar con la edad y no ser evidentes hasta que no aumentan las demandas del entorno, especialmente en los niños con una buena capacidad intelectual y un lenguaje conservado.

En los niños pequeños (menores de dos años), los motivos de derivación pueden ser retraso del desarrollo psicomotor, retraso del lenguaje, irritabilidad, alteración de conducta o, directamente, por sospecha de TEA de la familia o del pediatra. También se puede realizar una detección precoz durante el seguimiento de recién nacidos de riesgo neurológico o de aquellos que hayan sufrido una agresión grave del sistema nervioso central.

En los niños mayores de 3 años, la sintomatología suele ser más clara, y por tanto la derivación a especialistas es realizada con más frecuencia por sospecha de TEA (dificultades en la interacción social junto con patrones repetitivos del comportamiento). Sin embargo, en ocasiones esto no es así, por tanto, la derivación puede ser por motivos diferentes como problemas de conducta, de socialización o de expresión y control emocional, problemas de aprendizaje o del lenguaje, movimientos anormales... (5)

Habitualmente, las manifestaciones clínicas suelen comenzar alrededor de los dos o tres años y suelen ser los padres los primeros en detectarlos. A pesar de ello, la edad media de diagnóstico en Europa ronda los 4 años y medio. Hay una serie de signos de alarma que nos deben alertar sobre un posible TEA:

No responde a su nombre a los 12 meses de edad

No muestra objetos interesantes (punto en un avión en vuelo) a los 14 meses de edad

No presenta juego de simulación o ficción (fingir que “alimenta” de verdad a un muñeco) a los 18 meses de edad

Evita el contacto visual y quiere estar solo

Tiene dificultad para entender los sentimientos de otras personas o para hablar de sus propios sentimientos

Tiene un retraso en la adquisición del habla y las competencias lingüísticas

Repite palabras o frases una y otra vez (ecolalia)

Da respuestas no relacionadas con las preguntas

Se altera ante cambios menores

Tiene intereses obsesivos

Aletea las manos, mueve el cuerpo o da vueltas en círculos

Tiene reacciones inusuales a la manera en que las cosas suenan, huelen, saben, se ven o se sienten

Tabla 3: Signos de alarma precoces de TEA según la Asociación Española de Pediatría (AEPed). (5)

La detección temprana es importante porque proporciona la oportunidad de intervenir con el diagnóstico precoz. Esto es lo más recomendable para identificar y tratar posibles comorbilidades y comenzar con el tratamiento o actividades terapéuticas específicas del trastorno (hablamos de terapias ocupacionales, remodelación conductual y psicoterapia principalmente), mejorando así, la calidad de vida y el pronóstico.

Para llevarla a cabo existen numerosas herramientas de cribado que se pueden realizar a distintos niveles (padres, profesores, pediatras de atención primaria etc.). Las más importantes son: para niños entre 16-30 meses, M-CHAT. Para niños más mayores, CAST (4-6 años), o la escala Autónoma para la detección de S. Asperger o Autismo de Alto Funcionamiento (por encima de 6 años). (16)

El diagnóstico definitivo de TEA sigue siendo clínico como se ha expuesto en este trabajo previamente (criterios diagnósticos de la DSM V) y, aunque se van desarrollando biomarcadores moleculares y funcionales que pueden ayudar al diagnóstico precoz, están pendientes de validación. Uno de ellos ha sido evaluado por Jones et al. en un proyecto de investigación donde evalúan si la medición y seguimiento del compromiso visual social (*eye tracking*), es decir, la forma en que los niños observan y aprenden del entorno social, ofrece validez neurobiológica y clínica como posible medida diagnóstico temprano del TEA. Este reciente estudio ofrece resultados prometedores ya que ha demostrado una sensibilidad del 71,0% y una especificidad del 80,7%, en un estudio de cohortes de casi 500 pacientes validado con diagnóstico clínico experto. (17)

Una vez identificado un paciente con síntomas compatibles o bien con una prueba de cribado positivo, sería fundamental llevar a cabo un estudio genético precoz. Esto podría permitir, no sólo confirmar el caso y advertir (en caso de genes asociados) de posibles comorbilidades, sino también ofrecer consejo genético a la familia en caso de nuevo deseo gestacional.

## **1.5 Etiología de los TEA**

Los TEA, como veremos más adelante, son trastornos del neurodesarrollo predominantemente genéticos, que aparecen durante la primera infancia, pero se van a manifestar durante toda la vida del paciente. A pesar de ello, también se han identificado diversos factores de riesgo ambientales (principalmente prenatales) para el desarrollo de TEA.

### **1.5.1 Factores ambientales.**

Se han planteado muchos factores de riesgo de TEA. Varias revisiones sistemáticas y metaanálisis han descrito factores prenatales y perinatales, así como factores dietéticos y de estilo de vida maternos. A pesar de que es probable que hoy en día los casos provocados únicamente por factores de riesgo ambientales son la minoría, si conviene tener claro cuáles se han identificado con evidencia firme en la literatura, para ofrecer esta información a las familias.

La edad materna avanzada ( $\geq 40$  años) y la edad paterna ( $\geq 50$  años) se han asociado de forma independiente con el riesgo de TEA en varios estudios, al igual que los intervalos entre sucesivos embarazos cortos ( $< 24$  meses). Los factores no óptimos no

específicos durante el embarazo, incluidas las condiciones metabólicas maternas, el aumento de peso y la hipertensión, así como factores más específicos (como el ingreso materno en el hospital debido a infecciones bacterianas o víricas, preeclampsia o los antecedentes familiares de enfermedades autoinmunes) también se han asociado a un riesgo ligeramente mayor de TEA y retraso del desarrollo combinados.

El nacimiento prematuro (<32 semanas), el bajo peso al nacer (<1500 g), el ser pequeño para la edad gestacional y el ser grande para la edad gestacional (>95 percentil de peso al nacer) se han asociado de forma independiente con un mayor riesgo de TEA, aunque no está claro si estos factores son causales o marcadores de riesgo. No obstante, estos niños deben ser controlados para detectar TEA durante la infancia tardía y los primeros años de vida. No se han encontrado asociaciones consistentes entre el parto por cesárea o la concepción asistida y el riesgo de TEA.

Respecto al uso de medicación externa durante el embarazo, existe una amplia literatura, de la que podemos destacar lo siguiente:

Actualmente existen pruebas concluyentes que determinan que la exposición prenatal al ácido valproico (antiepiléptico y eutimizante empleado en el tratamiento de trastorno bipolar) se ha asociado con un mayor riesgo de TEA. Christianson et al. describieron por primera vez un posible vínculo entre la exposición materna al VPA y el TEA de la descendencia. Posteriormente, estudios más amplios confirmaron la asociación entre la exposición intrauterina al VPA y el autismo. Basándose en los criterios del DSM, las estadísticas hallaron que el 8,9% de los 56 niños con exposición prenatal al VPA en el estudio de monoterapia desarrollaron autismo o síndrome de Asperger. La exposición al valproato provoca un adelgazamiento de la corteza prefrontal temprana, la amígdala basolateral y el hipocampo.

En cuanto a los antidepresivos, incluidos los inhibidores selectivos de la recaptación de serotonina, estudios bien controlados no han sugerido un riesgo inequívoco, a pesar de las alertas sobre estos fármacos basadas en estudios previos (probablemente sometidos a algún sesgo de confusión).

Los suplementos preconceptuales de ácido fólico se han asociado con un menor riesgo de TEA y discapacidades generales del desarrollo, con una interacción significativa entre genes y ambiente. (18, 19)

Respecto a la vacunación frente a diversas enfermedades infecciosas y su asociación con los TEA (un tema ciertamente controvertido en el pasado), se puede afirmar con firmeza que no existe ningún tipo de asociación, ni mucho menos causalidad. (20)

## 1.5.2 Factores genéticos. Variantes genéticas.

La genética juega un papel fundamental en la patogénesis de los Trastornos del Espectro Autista. Una medida de ello es que los TEA son el trastorno del neurodesarrollo con mayor heredabilidad, la cual refleja la proporción de variabilidad fenotípica que es debida a la variabilidad genética (sobre el total de genes y ambiente). Esta ha sido calculada en diversos estudios, de los cuales destaca el llevado a cabo por Seven Sandin et al. en 2017 en Suecia que incluía una cohorte de 37.000 gemelos y otra de más de 2 millones y medio de hermanos y que estimó una heredabilidad del 83% para el TEA, siendo esta mayor que la del cáncer de mama (28%) o el cáncer colorrectal (35%) y similar al de la esquizofrenia, por ejemplo. Estos resultados, sugieren que los factores genéticos constituyen la mayor parte del riesgo individual de padecer TEA. (21)

A pesar de ello, el TEA no se hereda de forma mendeliana, sino que se trata de trastornos poligénicos ya que existe una gran cantidad de genes implicados (algunos todavía no se han descubierto) y con una baja penetrancia. Aun así, se están llevando a cabo cada vez más estudios para identificar qué genes están involucrados en el desarrollo de la enfermedad y se ha visto que hay tres principales grupos en que podemos dividirlos:

- Genes asociados al neurodesarrollo.
- Genes implicados en la formación y mantenimiento de sinapsis neuronales (este grupo está muy relacionado también con la aparición de epilepsias congénitas, a veces, comórbidas con el TEA).
- Genes implicados en el remodelado de la cromatina.

El genoma humano está compuesto por más de 3000 millones de pares de bases y sobre 22000 genes, sin embargo, la variación interindividual supone tan solo el 1% de todo ello (el 99% es igual entre nosotros). Ese 1% corresponde por tanto a lo que denominamos variantes genéticas o mutaciones, las cuales dividimos en:

- Variantes de un solo nucleótido o SNPs.
- Variantes del número de copias o CNVs.
- Traslocaciones.
- Inversiones.
- Variantes cromosómicas (haploidías, poliploidías y aneuploidías)

Estas, a su vez, dependiendo de la frecuencia con la que las encontremos en la población general se clasifican en “variación rara” (presentes en menos del 1% de la población) y “variación común” (presentes en más del 1%). Según su localización, su capacidad para producir enfermedad y su génesis, las mutaciones pueden afectar al ADN codificante o no codificante (este último conforma la vasta mayoría de la longitud total del genoma humano), pueden ser patogénicas o no patogénicas (recordemos que la mutación no deja de ser un mecanismo que permite la evolución de los organismos), y pueden ser *de novo* (cambio en la secuencia de ADN que aparece por primera vez en un sujeto, sin estar presente en sus progenitores o generaciones anteriores) o heredadas; y dentro de estas últimas recesivas, dominantes o ligadas a X.

La estructura genética de este grupo de trastornos es, como comentamos, compleja. Distinguimos, dentro de las variaciones implicadas en la patogénesis de la enfermedad, dos grupos:

-Variación común: Presentan una frecuencia del alelo menor (MAF) mayor del 1% (o 5% para variantes recesivas). Muchas veces comunes a la población general y con un amplio número de genes implicados. En este grupo encontramos algo más de la mitad de los TEA. a mayoría son varones.

-Variación rara: presentan una MAF menor del 1% (o 5% para variantes recesivas) Se trata de mutaciones *de novo* (lo más habitual) o heredadas un poco más frecuentemente por la madre (debido que muchas se heredan ligadas al cromosoma X), en genes concretos y se suele asociar a otras comorbilidades. Se estima que representan alrededor del 45 % de los casos, aunque normalmente se consigue el diagnóstico genético en el 25% (el resto se trata de variación perdida o *missing variation* que estamos tratando de buscar con trabajos como este). (22)

Un aspecto importante que hemos de aclarar es cuándo hablamos de TEA síndrómico y cuándo de TEA comórbido. Hasta el 70% de pacientes con TEA tienen comorbilidades, tanto psiquiátricas (trastornos de conducta, TDAH, depresión, T.O.C...) como no psiquiátricas (epilepsia, problemas gastrointestinales...) y dentro de ellas, la discapacidad intelectual es la más frecuente. Sin embargo, hablamos de TEA síndrómico cuándo el TEA se halla dentro de un conjunto de características clínicas de un síndrome genético conocido, y entre ellos destacan síndrome de Down, síndrome de X frágil y Prader-Willi/Angelman, como los más asociados al trastorno.

Aunque las variantes comunes contribuyen en gran medida al riesgo de TEA, son difíciles de reconocer, ya que se asocian a efectos sutiles, y muchas permanecen desconocidas. Por lo tanto, gran parte de nuestro conocimiento sobre los genes subyacentes al TEA procede de estudios que identificaron variantes con un riesgo de moderado a alto. Se ha estimado que las variantes en más de 400 genes y varias variaciones en el número de copias (CNVs) pueden representar variantes de riesgo alto a moderado para la enfermedad.

En cuanto al modo de transmisión de las variantes, antes se creía que un patrón de herencia poligénico o multifactorial explicaba la mayoría de los casos. No obstante, con el tiempo se ha encontrado que muchas personas con TEA presentan mutaciones raras que afectan el desarrollo neuronal y podrían ser suficientes para causar el trastorno por sí solas. Además, en algunas familias, una variante genética rara potencialmente perjudicial se encuentra tanto en individuos afectados como en no afectados, sugiriendo un patrón de herencia monogénico con penetrancia incompleta. Por lo tanto, se han revisado los patrones de herencia del TEA y se considera que la interacción entre variantes comunes y raras es más adecuada para explicar la arquitectura genética del trastorno. También hay casos donde una baja carga de variantes de bajo riesgo podría llevar al desarrollo del trastorno si se combinan con algunas variantes de riesgo moderado. En estas circunstancias, el riesgo de recurrencia de TEA en la familia es mayor que en la población general debido a la segregación de alelos de riesgo dentro de la familia. Finalmente, el TEA podría ser causado por una mutación rara y altamente perjudicial. Estas mutaciones, que suelen ser de alto riesgo, están asociadas a una alta penetrancia y suelen ser eventos *de novo*. En este caso, el riesgo de recurrencia en la familia sería solo un poco mayor que en la población general, debido a la posibilidad de mosaicismo en los padres. (23)

Actualmente, debido al continuo desarrollo de la genómica y descubrimiento de nuevas variantes, se trabaja con diversas bases de datos y recursos web creadas para el compendio y clasificación de estas (Franklin, Varsome, Intervar, ClinVar...) así como para su correlación con el fenotipo (Clingen, OMIM, HPO y otros).

En 2015, el American College of Medical Genetics and Genomics (ACMG) y la Association for Molecular Pathology (AMP) publicaron normas y directrices para la interpretación de variantes de secuencia (Richards et al Genetics in Medicine 2015). Estas directrices describen un marco para clasificar las variantes como "patogénicas", "probablemente patogénicas", "de significado incierto", "probablemente benignas" o "benignas" según una serie de criterios con niveles de evidencia definidos como muy fuerte, fuerte, moderado o de apoyo. Esta clasificación ha sido adoptada también posteriormente por el sistema nacional de salud inglés (National Health System), creando las guías "ACGS Best Practice Guidelines for Variant Classification in Rare Disease" y es en la que se basan los recursos que hemos nombrado anteriormente para la clasificación de variantes en apoyo del trabajo de diagnóstico genético. (24)

Pero para poder emplear estas fuentes de información hemos de realizar antes las pruebas diagnósticas de secuenciación que nos permitan llegar hasta las variantes. Actualmente, la batería de pruebas genéticas usadas en la práctica clínica incluye:

- Estudio del cariotipo; para identificar posibles variantes cromosómicas.
- Pruebas de Array-CGH (CMA de "chromosomal micro-Array"). Su objetivo es caracterizar diversas CNVs (deleciones o duplicaciones demasiado pequeñas como para ser visibles con la resolución del cariotipo) comunes en la población en diferentes puntos del genoma de cara a obtener un banco que nos sirva de comparación para identificar posibles variantes patogénicas.
- Secuenciación y análisis de exomas (WES de "whole exome sequencing").

Estas tres nos ofrecen una rentabilidad diagnóstica de alrededor del 25%, lo que quiere decir que en al rededor del 75% de los casos no se encuentran variantes patogénicas o bien se desconocen los genes implicados (la mayor parte de los TEA con variación común) (22).

### **1.5.3 Diagnóstico genético en el TEA. El análisis de genoma completo.**

A diferencia de enfoques específicos como la secuenciación del exoma o el CMA, que analizan una porción limitada del genoma, la secuenciación del genoma completo ofrece una visión íntegra de todo el genoma. Es una prueba tremendamente prometedora para aplicaciones de descubrimiento, como la identificación de variantes causales y el ensamblaje de genomas nuevos. (25)

Se sabe que muchas variantes patogénicas (no sólo de TEA sino en general) se encuentran en partes no codificantes del genoma, y WES no suele proporcionar información sobre estas, aunque es cierto que algunas regiones no codificantes específicas (por ejemplo, aquellas de importancia clínica conocida) se pueden secuenciar además de los exomas. Además, la secuenciación WES puede no capturar algunas partes exónicas del genoma, en particular regiones que son ricas en GC.

Por el contrario, en WGS, casi todo el genoma, incluidas las partes no codificantes, está secuenciado. La inclusión de partes no codificantes del genoma proporciona información crucial para comprender la arquitectura genética de los rasgos de la enfermedad, ya que se estima que cerca del 90 % de los intervalos genómicos relacionados con la enfermedad están en un espacio no codificante. WGS está viendo cada vez más uso en las evaluaciones clínicas directas, identificando contribuyentes genéticos tanto a la enfermedad aguda como al riesgo de enfermedad.

Con el análisis estandarizado de variantes en muestras de WGS podemos identificar polimorfismos comunes de un solo nucleótido (SNP), así como variantes raras y *de novo* de un solo nucleótido (SNV), inserciones/deleciones cortas (denominadas *indels*), variantes del ADN mitocondrial (ADNmt) y variantes estructurales (SV, incluyendo variantes de número de copias [CNV], inversiones, inserciones mayores, disomías uniparentales [UPD] y expansiones de repeticiones en tándem o *short tandem repeats* [STR o TRE]).

Lowther et al, para evaluar la utilidad clínica de la secuenciación de genoma completo, comparan su rendimiento diagnóstico frente a tres pruebas (WES, CMA y análisis del cariotipo) en 1612 familias de cuatro miembros que incluían un individuo con TEA y en 295 familias prenatales, y concluyen que WGS identificó una variante diagnóstica en el 7,8% de los casos con TEA, casi 2 veces más que CMA (4,3%) y 3 veces más que WES (2,7%). Sin embargo, cuando capturamos sistemáticamente las CNVs de los datos del exoma, el rendimiento diagnóstico de WES (7,4%) se acercó mucho más a WGS, aunque no lo superó. Esto último da una idea de que la mayor parte de la ventaja diagnóstica que ofrece esta prueba podría tener que ver con el análisis de dichas variantes estructurales (CNVs, traslocaciones, inversiones...), y en concreto las expansiones de repeticiones en tándem (STRs), indetectables, por ejemplo, con las pruebas CMA, las cuales detectan CNVs en regiones pequeñas del genoma. (26)

Estos resultados demuestran que la estrategia de la secuenciación y análisis de genoma completo (WGS) ofrece una potencial ventaja diagnóstica en cuanto a identificación de variantes en regiones no codificantes, así como en la captación de CNVs, indels y SNPs en una sola prueba. Por ello, aunque es verdad que la complejidad técnica del análisis de genoma completo es mayor (supone la secuenciación de alrededor e 3 billones de pares de bases frente a 30 millones en el caso de WES) podría ofrecernos una mayor capacidad diagnóstica para aquellos casos de TEA no identificados genéticamente, así como abrir un nuevo frente investigador en lo que respecta al papel de las secuencias no codificantes en la génesis de la enfermedad.

## 2. Objetivos y plan de trabajo

En este trabajo, partimos de cinco tríos genéticos, es decir, cinco pacientes con un diagnóstico clínico de TEA (probandos afectados) y sus respectivos padre y madre (progenitores sanos o afectados). El genoma de cada uno de los integrantes del trío fue secuenciado en el contexto de un proyecto de investigación de la Fundación Pública Galega de Medicina Xenómica.

De cada probando afecto, tenemos el conjunto de las variantes de su genoma, las cuales priorizaremos, filtraremos y anotaremos funcionalmente gracias a la herramienta software *Emedgene Analyze* (Illumina, inc.), con el objetivo de reducir el número y quedarnos con variantes candidatas para analizarlas en detalle y encontrar la variante o variantes causales si es que existen. Una vez obtenido un número manejable de variantes candidatas, llevaremos a cabo el proceso de estudio y selección de estas basándonos en criterios como la clasificación de patogenicidad, herencia, compatibilidad del fenotipo con el descrito en la literatura y bases de datos etc. Todo ello con la meta final de obtener un diagnóstico genético que poder ofrecer al paciente y su familia.

Por tanto, los objetivos principales de este trabajo son:

- i. Realizar un análisis de genoma completo de cinco tríos de TEA (probando afectado y progenitores sanos) con el objetivo encontrar una variante causal que nos proporcione el diagnóstico genético del paciente en cuestión.
- ii. Comprobar si el análisis del genoma completo en tríos nos podría ofrecer una mayor eficacia, fiabilidad y rapidez diagnóstica como prueba de primer nivel en el diagnóstico de TEA, analizando, para ello, el caso de un paciente que ha sido estudiado mediante WES (*whole exome secuencing*) previamente con resultado negativo.
- iii. Comprender, mediante los anteriores, cómo se lleva a cabo el proceso de estudio diagnóstico genético, en este caso, de pacientes con patología del neurodesarrollo.

### 3. Material y métodos

En el desarrollo de un estudio genético de diagnóstico hemos de pasar de moléculas (una muestra de ADN en sangre de nuestro paciente) a datos informáticos. En este trabajo vamos a usar los datos que nos ofrece la secuenciación del genoma completo (WGS) de cinco probandos distintos diagnosticados de TEA junto con la de sus respectivos progenitores (es a lo que llamamos trío genético).

Un genoma completo contiene alrededor de 4 000 000 variantes genéticas, las cuales tenemos que someter a un proceso de priorización de variantes, con el objetivo de quedarnos con un número lo más reducido posible de ellas, que serían las variantes candidatas, y luego analizarlas en detalle para intentar llegar a la variante causal en el caso de que exista.

Para ello, vamos a emplear una serie de parámetros que nos sirven para ir descartando variantes hasta quedarnos con un número manejable de las mismas.

En este trabajo usaremos la herramienta software *Emedgene Analyze* de Illumina inc., la cual cuenta con un programa de Inteligencia Artificial que, empleando una serie de criterios que vamos a explicar a continuación, prioriza y nos muestra de manera automática unas variantes candidatas en cada probando. Esto facilita la labor del genetista al poder trabajar con estas variantes en primera instancia, sin tener que aplicar todos los procedimientos de filtrado uno a uno. Aun así, esto no nos impide recurrir a otras variantes no priorizadas por la aplicación ya que, de igual manera, se puede acceder manualmente a los filtros.

Esto nos permite ahondar en cada una de las candidatas, buscando información compatible con nuestro probando en las distintas bases de datos y herramientas web existentes en busca de la causante de la patología.

Vamos a explicar cuáles son los criterios de cribado más importantes en que se basa nuestro filtrado de variantes, junto con algunos de los instrumentos de los que disponemos para evaluarlos:

#### 3.1 Parámetros de calidad

Valoraremos varios aspectos básicos a la hora de valorar una variante:

- Profundidad de lectura. Esto es número de veces que se lee un nucleótido durante la secuenciación. Una mayor profundidad de lectura aumenta la confianza en los resultados finales.
- Cobertura: Es el porcentaje de la secuencia del gen, exoma o genoma (este último en nuestro trabajo) que se ha leído con la prueba, por lo que cuánto más cercano esté al 100% mayor capacidad diagnóstica teórica tendrá nuestro estudio genético.

- **Distribución alélica.** Se refiere a la pérdida de la ratio 50:50 entre el alelo de referencia y el alelo alternativo cuando una mutación es heterocigota, ya que se espera que se lea el alelo de referencia alrededor del 50% de las veces, y el otro 50% el alelo alternativo. Sin embargo, si el número de lecturas del alelo alternativo es muy inferior al número de lecturas del alelo de referencia, es muy probable que la variante realmente sea un artefacto y, por tanto, podemos eliminarla. Sería una excepción cuando la mutación es homocigota, ya que solo veremos el alelo alternativo. También se rompería esta proporción en el caso de que un paciente presente una mutación somática; por ejemplo, que provenga de un tumor, entonces la distribución del alelo de referencia y el alelo alternativo ya no sería del 50:50 sino que variaría; pero estas variantes no son el objetivo de este estudio por no ser relevantes en la etiología del TEA.

Todos estos parámetros aparecen reflejados en la aplicación Emedgene, la cual dará una valoración de calidad a cada variante que seleccionemos. De esta forma, va a ofrecer una clasificación en tres niveles de calidad para cada variante: baja, moderada o alta.

### **3.2 Filtrado interno**

La herramienta, junto con los registros internos del laboratorio (errores de secuenciado previos) es capaz de identificar defectos en el proceso de lectura de ADN que dan lugar a falsos positivos. Estas han de ser eliminadas en primer lugar.

### **3.3 Efecto molecular sobre la transcripción del gen.**

Debemos tener en cuenta el efecto de la variante en la transcripción de la secuencia genética.

Dividimos las variantes puntuales o SNVs, según su efecto molecular en la secuencia genética en:

- *Missense* o variantes de sentido errado, en la cual se produce un cambio en un único nucleótido, provocando la aparición de un codón que codifica para un aminoácido diferente.
- *Nonsense* o variantes sin sentido, en la cual el cambio en el nucleótido provoca un codón de finalización de lectura. Son las que más probabilidades tienen de provocar pérdida de función y, por tanto, ser patogénicas.
- *Frameshift* o variantes de cambio de marco de lectura, donde una inserción o deleción provoca una agrupación en codones diferentes que cambia la secuencia de lectura. Pueden provocar la formación de una proteína anómala o también afectar a su longitud si dan lugar a un codón de parada nuevo. Presentan también una alta probabilidad de ser patogénicas.
- Variantes sinónimas, que son alteraciones en la secuencia de un gen que no resultan en un cambio del aminoácido en la proteína final. Esto ocurre porque el

código genético es redundante y varios tripletes de nucleótidos pueden codificar para el mismo aminoácido.

- Variantes de *Splicing* del sitio canónico: a veces las variantes se pueden situar en las secuencias reguladoras de *splicing* (que es el proceso por el cual los intrones son escindidos del transcripto de ARN mensajero primario y los exones se unen para generar ARN mensajero maduro). Esto puede provocar errores en el proceso de empalme y por tanto en la posterior traducción dando lugar a trastornos hereditarios.

Por otra parte, tenemos las variantes estructurales entre las que encontramos las CNVs, traslocaciones, inversiones y STRs, cuyo efecto sobre la secuencia de ADN depende fundamentalmente de su localización. En muchas ocasiones (por ejemplo, en casos de variantes intrónicas) dicho efecto es casi siempre desconocido.

Un tipo de variantes especiales son las que denominamos reguladoras. En nuestro trabajo cobra especial importancia las variantes localizadas en regiones no codificantes, denominadas 5'3' UTR (*5'3' untranslated regions*). El 5'UTR es la secuencia de ARN inmediatamente anterior al ARN codificante. Las variaciones genéticas en las UTR pueden modificar elementos reguladores que afectan a la interacción de las UTR con proteínas y microARN. Las consecuencias funcionales generales incluyen la modulación de la transcripción del ARNm, la estructura secundaria, la estabilidad, la localización, la traducción y el acceso a reguladores como los microARN y las proteínas de unión al ARN. Se sabe que las alteraciones de estos mecanismos reguladores modifican las vías moleculares y los procesos celulares, pudiendo dar lugar a procesos patológicos. (27)

Así mismo, es importante tener en cuenta otros parámetros como:

-La posición de la variante en el gen y en la proteína (si forma parte o no de un dominio funcional).

-La haploinsuficiencia, que es la situación que se presenta cuando una copia de un gen se inactiva o se elimina y la copia funcional que queda del gen no es suficiente para producir la cantidad de producto génico necesaria para conservar el funcionamiento normal. (28)

-La triplosensibilidad es la misma situación, pero para duplicaciones o triplicaciones. Es decir, es la sensibilidad del gen que se altera con este tipo de variantes para inducir un fenotipo patológico en el individuo.

-La tolerancia del gen a variantes de pérdida de función o pLI (probability of loss-of-function intolerance): Los genes intolerantes a variantes de pérdida de función apenas tienen variantes de este tipo, precisamente porque no las toleran y la existencia de este tipo de variantes suelen ser deletéreas para la función del gen. Se mide como la probabilidad de que el gen en el que se encuentra la variante pertenezca a un grupo de genes que contienen <10% de sus variantes esperadas del tipo *los-of-function*. Los valores oscilan entre 0 y 1. Los genes con valores mayores (más cercanos a uno) son más intolerantes a las mutaciones. Suele ser un valor que va parejo al de haploinsuficiencia, aunque no siempre. (29)

La información acerca de la posición de las variantes en el genoma, así como su efecto sobre la transcripción de las diferentes secuencias, la obtenemos del proceso de anotación funcional que lleva a cabo el software Emedgene. Esta herramienta cuenta con

una plataforma de análisis genómico automatizada y basada en la IA explicable que admite la entrada de datos de WGS para un análisis e interpretación de variantes eficiente. El software permite captar variantes de nucleótido único (SNV) inserciones/delecciones (indels), variantes en el número de copias (CNV) variantes de ADN mitocondrial (ADNmt), variantes estructurales y repeticiones cortas en tándem (STR).

### 3.4 Frecuencias poblacionales

Sabemos que aquellas variantes que se encuentren presentes de una manera recurrente en la población general tienen unas posibilidades muy bajas de corresponder a patología. Normalmente están implicadas en el desarrollo de rasgos fenotípicos frecuentes como por ejemplo el color de ojos, de pelo etc., o bien con tienen una función todavía no conocida pero cuya frecuencia excede ampliamente a la de la propia enfermedad, hallándose de manera reiterada en personas sanas. Por ello, el cribado por frecuencia poblacional es uno de los más importantes para ayudar en la diferenciación de variantes benignas.

La reciente disponibilidad de bases de datos de referencia muy grandes, como el conjunto de datos del Exome Aggregation Consortium (ExAC), que ha caracterizado las frecuencias alélicas poblacionales de diez millones de variantes genómicas mediante el análisis de datos de secuenciación del exoma (ES) de más de 60.000 humanos, brinda la oportunidad de obtener estimaciones de frecuencia robustas incluso para variantes raras. (30)

Otro recurso habitualmente empleado es (<https://gnomad.broadinstitute.org>.) Se trata de una base de agregación de genoma que proporciona datos de genética de poblaciones y asociación con enfermedades.

Se ha demostrado que los límites de frecuencia poblacional por debajo del 1% para trastornos de herencia dominante y del 5% para trastornos de herencia recesiva son los indicados para detectar variantes causales en enfermedades raras. Estos son los límites de frecuencia alélica que emplearemos en este trabajo.

### 3.5 Por fenotipo. Bases de datos de enfermedad y fenotipos clínicos.

Hemos de comprobar si existe evidencia en la literatura de que nuestra variante ha sido relacionada con una condición clínica compatible con la de nuestro paciente. Para ello existen una serie de bases de datos clínicas que correlacionan variantes descritas con fenotipos o directamente patologías en los pacientes.

Las más importantes y que emplearemos en este trabajo son:

- OMIM (<https://www.omim.org>): Online Mendelian Inheritance in Man es un recurso desarrollado en el Instituto McKusick-Nathans de Medicina Genética de la Facultad de Medicina de la Universidad Johns Hopkins que contiene un compendio exhaustivo y fidedigno de genes y fenotipos genéticos humanos. Está disponible gratuitamente y se actualiza mensualmente. Contiene información sobre todos los trastornos mendelianos conocidos y más de 16 000 genes. OMIM se centra en la relación entre fenotipo y genotipo.

- Genomics England Panel App (<https://panelapp.genomicsengland.co.uk>): es una base de conocimientos de acceso público que permite crear, almacenar y consultar paneles virtuales de genes relacionados con trastornos y enfermedades humanas. Incluye una herramienta de *crowdsourcing* que permite que los genes y las entidades genómicas (repeticiones cortas en tándem/STR y variantes del número de copias/CNV) sean añadidos o revisados por expertos de toda la comunidad científica mundial, proporcionando una oportunidad para la estandarización de los paneles de genes, y un consenso sobre qué genes tienen suficientes pruebas de asociación con la enfermedad. Esta herramienta también nos proporciona información sobre el modo de herencia de dichas alteraciones genéticas (información que usamos como explicamos en el apartado 4.4)
- HPO (<https://hpo.jax.org/app/>): La Ontología del Fenotipo Humano (HPO) proporciona un vocabulario estandarizado de las anomalías fenotípicas que se encuentran en las enfermedades humanas. Cada término de la HPO describe una anomalía fenotípica. La HPO contiene actualmente más de 13.000 términos y más de 156.000 anotaciones a enfermedades hereditarias. El proyecto HPO y otros han desarrollado software para el diagnóstico diferencial basado en el fenotipo, el diagnóstico genómico y la investigación traslacional. Permite buscar fenotipos asociados a genes alterados o genes asociados a fenotipos determinados.
- SFARI (<https://gene.sfari.org>): es una base de datos actualizada dedicada a la recopilación de genes y variantes asociadas a los trastornos del espectro autista.

### 3.6 Modo de herencia de las variantes y zigosis. Trío genético.

Tenemos que emplear la información que nos ofrece la condición clínica y la secuenciación genética del trío familiar ya que esto nos permite descartar variantes sabiendo la zigosis que se presenta en el probando comparada con el modo de herencia que tiene la condición clínica a la que se asocia el gen en el que se encuentran.

Por ejemplo, si una variante está en heterocigosis y el gen se asocia a una condición de herencia autosómica recesiva, es poco probable que se la variante causal, ya que necesitaríamos que estuviera en homocigosis o en heterocigosis compuesta (dos variantes diferentes en trans, es decir, en alelos distintos, en el mismo gen).

De todos modos, hay casos en que tenemos en el probando una sola variante patogénica en heterocigosis en un gen asociado a una patología de herencia autosómica recesiva, pero si el fenotipo es muy coincidente y específico, podemos reservarla e informarla si hay pruebas bioquímicas específicas que nos puedan confirmar la condición (en ocasiones si es muy específico el fenotipo, se puede informar aunque no haya pruebas bioquímicas), porque puede ser que el otro alelo no sea funcional por alguna causa genética o epigenética que no somos capaces de detectar con la tecnología actual.

El estudio de tríos nos permite saber si las variantes detectadas son *de novo*. Sabiendo que ambos padres están sanos, es poco probable que una variante heredada en un gen asociado a una condición autosómico dominante sea la causal (debería tener penetrancia incompleta), con lo que priorizar las variantes *de novo* es una estrategia de primera línea.

Las variantes patogénicas que están en genes asociados a condiciones clínicas de herencia autosómico recesivas pueden ser causales, aunque sean heredadas de los progenitores (de hecho, lo más frecuente es que sea heredada una variante de cada progenitor).

Las variantes causales en genes ligados al X pueden ser heredadas de la madre en varones si la herencia es ligada al X recesiva, es decir, que las mujeres son portadoras no afectas y los varones sí.

### 3.7 Clasificación de variantes según los criterios ACGM y ACGS.

Las variantes han de ser clasificadas siguiendo las recomendaciones de la ACGM y ACGS. Estas se basan en la creación de un sistema de clasificación de cinco niveles para las variantes relevantes para la enfermedad a estudio.

- Patogénica (*pathogenic*, P)
- Probablemente patogénica (*likely pathogenic*, LP)
- Variante de significado incierto (*variant of uncertain significance*, VUS)
- Probablemente benigna (*likely benign*, LB)
- Benigna (*benign*, B)

Para llegar a incluir una variante en alguno de estos grupos el equipo de la ACMG utiliza una serie de criterios de evidencia, basados en las características de las variantes que hemos visto hasta ahora. Estas son, por ejemplo, predicciones del efecto molecular, análisis de frecuencias poblacionales, relación genotipo-fenotipo, segregación, haploinsuficiencias... Estos criterios de evidencia van a ir en apoyo de una clasificación patogénica (P) o benigna (B).

Los diferentes tipos de pruebas (funcionales, genéticas, poblacionales, *in silico*, etc.) se estratifican según el nivel de evidencia que aporten (de apoyo, moderado, fuerte, muy fuerte) y se asigna a cada una un código alfanumérico que valora su contribución al nivel de patogenicidad. Existe 1 código para *Pathogenic very Strong* (PVS), 4 para *Pathogenic Strong* (PS), 6 para *Pathogenic Moderate* (PM), 4 para *Pathogenic Supporting* (PP), 1 para *Benign Stand Alone* (BA), 4 para *Benign Strong* (BS) y 7 para *Benign Supporting* (BP).

Estos códigos se emplearán, siguiendo un conjunto de criterios de combinación, para clasificar a la variante a estudio, en uno de los 5 grupos que hemos nombrado al principio. La siguiente tabla elaborada por Richards et al (2015), sienta las bases para la elaboración de los criterios de clasificación. (24)

Clasificación	Criterios de clasificación
Patogénica, P	<ul style="list-style-type: none"> <li>i. 1 <i>Pathogenic very Strong</i> (PVS) Y: <ul style="list-style-type: none"> <li>a) <math>\geq 1</math> Strong (PS) O</li> <li>b) <math>\geq 1</math> Moderate (PM) O</li> <li>c) <math>\geq 2</math> Supporting (PP) O</li> <li>d) 1 <i>Moderate</i> (PM) Y 1 Supporting (PS)</li> </ul> </li> <li>ii. <math>\geq 2</math> Pathogenic Strong (PS)</li> <li>iii. 2 Pathogenic Strong Y: <ul style="list-style-type: none"> <li>a) <math>\geq 3</math> Moderate (PM) O</li> <li>b) <math>\geq 2</math> Moderate (PM) Y <math>\geq 2</math> Supporting (PS) O</li> <li>c) <math>\geq 1</math> Moderate (PM) Y <math>\geq 4</math> Supporting (PS)</li> </ul> </li> </ul>
Probablemente patogénica, LP	<ul style="list-style-type: none"> <li>i. <math>\geq 1</math> Pathogenic very Strong (PVS) Y 1 Moderate (PM)</li> <li>ii. 1 Strong (PS) Y <ul style="list-style-type: none"> <li>a) 1-2 Moderate (PM)</li> <li>b) <math>\geq 2</math> Supporting (PS)</li> </ul> </li> <li>iii. <math>\geq 3</math> Moderate</li> <li>iv. 2 Moderate (PM) Y <math>\geq 2</math> Supporting (PS)</li> <li>v. 1 Moderate (PM) Y <math>\geq 4</math> Supporting (PS)</li> </ul>
Probablemente benigna, LB	<ul style="list-style-type: none"> <li>i. 1 Strong (BS) and 1 Supporting (BP) OR</li> <li>ii. <math>\geq 2</math> Supporting (BP)</li> </ul>
Benigna, B	<ul style="list-style-type: none"> <li>i. 1 Stand Alone (BA)</li> <li>ii. <math>\geq 2</math> Strong (BS)</li> </ul>
Significado incierto, VUS	Variantes que no cumplen los criterios o cuyos criterios para benigno y patogénico son contradictorios.

Tabla 4: Criterios para la clasificación de variantes, Richards et al. (2015)

Hay que destacar que la ACMG ha desarrollado esta clasificación para enfermedades de herencia mendeliana, por lo que, aunque cada vez se esté extendiendo su uso al resto patologías de origen genético (los TEA, como hemos explicado tienen en general, un modo de herencia complejo), hay que tomar estas recomendaciones con cautela.

Para nuestra clasificación, priorizaremos aquellas variantes *Patogenic o Likely Patogenic*, aunque sin descartar las *Variant of Uncertain Significance* en caso de encajar en el resto de los criterios.

La herramienta web Franklin (<https://franklin.genoox.com>), es una forma avanzada y actualizada de clasificación de variantes con criterios ACMG y es la que emplearemos para trabajar de forma individualizada en nuestras variantes candidatas. Podremos ver no solo la clasificación final, sino que obtendremos un desglose de cuáles son los criterios de evidencia que nuestra variante cumple para llegar a ella. Además, este recurso ofrece accesos asociados a muchas de las herramientas web y bases de datos que hemos nombrado anteriormente para obtener información más detallada a cerca de su efecto molecular, fenotipo etc.

Otro recurso interesante para encontrar información sobre la clasificación de variantes es Varsome (<https://varsome.com>).

### 3.8 Método de trabajo

Como hemos dicho, todos estos procesos de filtrado, junto con algunos más adicionales, se llevarán a cabo de forma integrada gracias a la plataforma online Emedgene Analyze de Illumina Inc., la cual emplea Inteligencia Artificial y una serie sets de filtros definidos por el usuario (como los que vamos a explicar a continuación) que permiten trabajar con un número mucho mayor de variantes secuenciadas al emplear WGS. De esta manera el genetista puede priorizar y filtrar las variantes de una manera rápida y secuencial.

Por tanto, en el proceso de priorización de variantes que aplicaremos en cada uno de nuestros probandos se aplicarán los siguientes sets predefinidos:

- Variantes priorizadas por Inteligencia Artificial de Emedgene (*Most likely by AI*).
- Variantes patogénicas conocidas (estudiadas, identificadas, publicadas...) (*Known pathogenic*).
- SNVs con un alto impacto asociado al fenotipo. En este grupo incluiremos aquellas variantes de novo en genes contenidos en el panel de PanelApp, SFARI y en genes relacionados con el autismo identificados por la herramienta Emedgene.
- CNVs en genes relacionados con el fenotipo (*CNV phenomatch*) y genes con sensibilidad de dosis (*CNV dosagegenes*). También nos fijaremos en los filtros de CNVs *de novo* de más de 100 kpb y 1Mb que corresponden a variantes de gran tamaño en el genoma.
- Variantes en genes ligados a X (*Rare X-linked*), en ADN mitocondrial (*MTDNA Rare*) y variantes tipo STR (TRE en su traducción al español) con una frecuencia poblacional inferior al 5% ( $STR < 5\%$ ).
- Variantes en genes candidato en homocigosis y heterocigosis compuesta.

- Variantes contenidas en la región intrónica 5' 3'UTR (*5' 3' untranslated region*) que asocien un fenotipo similar (*5' 3' UTR intronic de novo phenomatch*). Esta es una secuencia del gen que se sitúa justo antes del codón de iniciación (es decir, que no se transcribe a proteína) y que sabemos que tiene gran importancia en la regulación de la transcripción de muchos genes.
- Emplearemos también el filtro *Rare de novo*, para buscar variantes *de novo* (son muy frecuentes en los TEA) con una baja frecuencia poblacional.

Como hemos señalado previamente en este trabajo, tanto las CNVs como las 5' 3' UTR variants, STRs y aquellas encontradas en ADN mitocondrial son identificables gracias al análisis del genoma completo y serían más difíciles o incluso imposibles de analizar en un estudio de análisis de exoma.

Por último, una vez obtengamos alguna variante candidata con estos filtros emplearemos la herramienta Franklin y el recurso web OMIM para obtener información más detallada a cerca de su clasificación ACMG y consecuencias fenotípicas asociadas, respectivamente.

Emplearemos también, en caso de variantes tipo CNV, la herramienta DECIPHER (<https://www.deciphergenomics.org>) (31) para buscar información de sus características moleculares y fenotípicas, así como la herramienta de clasificación de CNVs de CLINGEN (<https://clinicalgenome.org>) y la página web <https://rarechromo.org>, que es una base de datos de condiciones genéticas raras provocadas por alteraciones cromosómicas.

Las aplicaciones bioinformáticas MetaDome y genomAD (<https://stuart.radboudumc.nl/metadome/> y <https://gnomad.broadinstitute.org>) nos sirven como complemento para evaluar la tolerancia de diferentes genes a variantes tipo *missense*.

Para valorar las variantes del sitio de *splicing* emplearemos dos herramientas. Por un lado los denominados predictores de *splicing in silico* (*splicing AI* y *SNV Ada*) que encontramos en la herramienta Franklin, y, además, el recurso web Alamut de sophia Genetics (<https://www.sophiagenetics.com/platform/alamut-visual-plus/>) el cual nos permite recrear virtualmente la variante que estamos estudiando para comprobar el efecto *in silico* que tendría sobre el *splicing* (pérdida de secuencias exónicas, aparición de secuencias anómalas por alteraciones en el empalme, unión de factores de transcripción etc.)

## 4. Resultados

### 4.1 Probando 1

Se trata de un paciente cuya condición clínica ha sido informada como rasgos autistas, retraso en la adquisición del lenguaje y el habla y retraso en el desarrollo motor. La madre del sujeto ha sido informada como afecta clínicamente de un cuadro similar (retraso del desarrollo del lenguaje y rasgos autistas) mientras que el padre está sano.

Se conoce que este paciente había sido estudiado previamente en diagnóstico genético con una prueba de WES con un resultado negativo.

En este caso, aplicando los filtros establecidos en este trabajo se han encontrado tres variantes candidatas que se muestran a continuación.

Gen	Nomenclatura	Efecto molecular	Heredada/ <i>De novo</i>	Zigosidad	Fenotipo asociado y herencia	Clasificación ACMG
NRXN1	NC_000002.11: g(50801906_50935037)del.	CNV. Delección	Heredada (madre)	Heterocigoto	Síndrome Pitt-Hopkins-like 2, que asocia retraso mental, retraso motor hiperventilación y TEA.	Patogénica
PPP1R3 F	PPP1R3F (NM_033215.5): c.151T>G p.(Leu51Val)	SNV missense	<i>De novo</i>	Hemicigoto (ligado a X)	Asociada a trastornos del neurodesarrollo y rasgos autistas.  Ligado a X	Significado incierto (VUS)
OPHN1	OPHN1 (NM_002547.): c.1270T>G p.(Phe424Val)	SNV missense	<i>De novo</i>	Hemicigoto (ligado a X)	Retraso en el neurodesarrollo con rasgos faciales dismórficos. Hipoplasia cerebelar.  Ligado a X	Significado incierto (VUS)

-La variante OPHN1(NM\_002547.3):c.1270T>G p.(Phe424Val) a pesar de concordar parcialmente con el fenotipo de nuestro paciente obtiene una baja puntuación en la clasificación ACMG en Franklin, siendo clasificada como VUS debido a que se ha encontrado en homocigosis en una proporción mayor de la esperada para ser causante de la enfermedad.

-Con la variante PPP1R3F(NM\_033215.5):c.151T>G p.(Leu51Val) nos encontramos dos problemas, a pesar de que sí existe bibliografía que lo asocia al TEA: por una parte Emedgene nos la clasifica con una calidad de secuenciación baja en el probando. Además, aparece clasificada en Franklin como VUS debido a que los predictores *in silico* computacionales predicen de forma unánime un efecto benigno en el gen. (32)

-La variante NC\_000002.11:g(50801906\_50935037)del. Consiste en una microdelección de 133 131 kbp desde la posición 50\_80 19 06 hasta la posición 50 93 50 37 del brazo corto del cromosoma 2. Esta microdelección solapa parcialmente con el gen de la neurexina 1 (NRXN1), abarcando los exones 6, 7 y 8 pero comenzando y acabando en el interior de dos regiones intrónicas. La hemos localizado con el filtro *CNV phenomatch* y sus parámetros de calidad agregada en Emedgene (profundidad y cobertura) aparecen como altos. La frecuencia poblacional es de 1/5919 según DECIPHER. La variante se presenta en heterocigosis en el probando y ha sido heredada de la madre (que presenta un fenotipo similar al probando, aunque subclínico y no diagnosticado). El fenotipo asociado lo hemos encontrado en OMIM, que nos habla del síndrome Pitt-Hopkins-like 2 (código OMIM: 614325) que asocia retraso en el desarrollo motor, TEA, retraso mental y a veces hiperventilación, lo cual concuerda con nuestro paciente. No se aclara en Emedgene si esta condición se hereda de forma recesiva o dominante, pero en OMIM, indican que el síndrome asociado es de herencia autosómica recesiva. La búsqueda en DECIPHER mediante las coordenadas de la delección da como resultado la siguiente variante, que coincide con la nuestra:

DECIPHER Patient	Sex	Location	Size	Inheritance / Genotype	Pathogenicity / Contribution	Phenotype(s)	Contact
271876	46XY	2 50654886 50786517 GRCh38 Deletion	131.63 kb	Unknown Heterozygous			
333396	46XX	2 50665709 50810025 GRCh38 Deletion	144.32 kb	Unknown Heterozygous		Intellectual disability	
359217	46XY	2 50665709 50810025 GRCh38 Deletion	144.32 kb	Unknown Heterozygous	Uncertain	Specific learning disability, Speech apraxia	
266041	46XY	2 50578337 50732545 GRCh38 Deletion	154.21 kb	Inherited from normal parent Heterozygous		Autism, Delayed speech and language development, Intellectual disability	
271770	46XY	2 50654886 50830715 GRCh38 Deletion	175.83 kb	Unknown Heterozygous			
308316	46XY	2 50882515 50862503 GRCh38 Deletion	179.99 kb	Paternally inherited Heterozygous	Uncertain	Delayed speech and language development, Posterior embryotoxon, Smooth philtrum	
356009	46XY	2 50623553 50810025 GRCh38 Deletion	186.47 kb	Unknown Heterozygous			
501658	46XY	2 50561837 50750646 GRCh38 Deletion	191.81 kb	Unknown Heterozygous	Likely pathogenic	Epicanthus, Immunodeficiency, Recurrent infections, Recurrent otitis media, Sensorineural hearing impairment	

Imagen 1: Búsqueda en DECIPHER por coordenadas de la variante NC\_000002.11:g(50801906\_50935037)del.

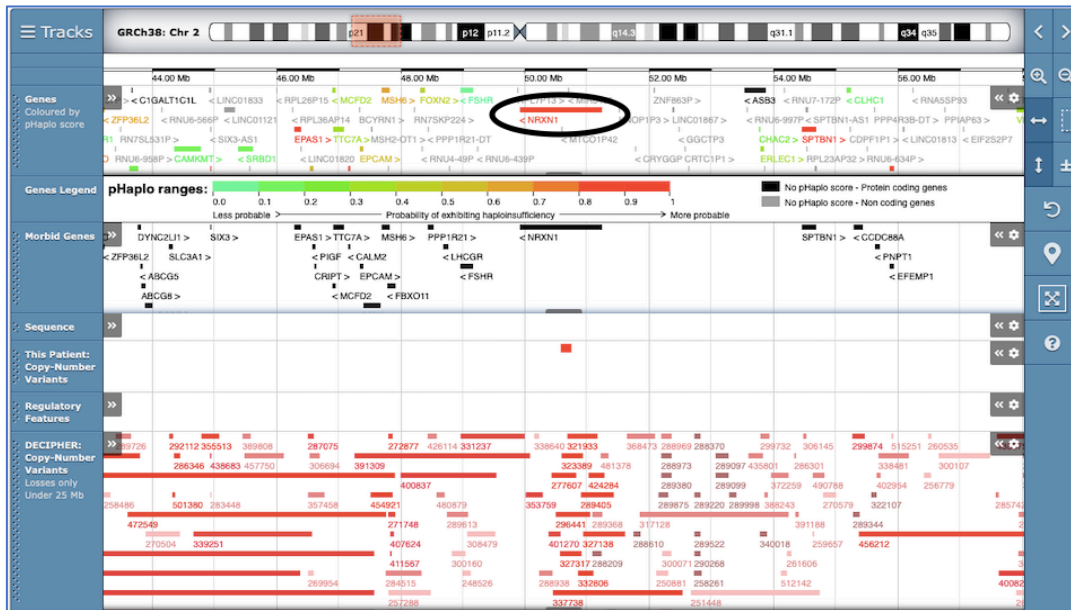


Imagen 2: Búsqueda en el browser de DECIPHER de la variante coincidente por coordenadas.

Como se puede ver en DECIPHER, la variante aparece asociada a un fenotipo muy similar al de nuestro caso e indicando que el probando presenta esta delección en heterocigosis y ha sido heredada por vía materna (lo que coincide con nuestro caso); además, vemos como el gen NRXR1 es severamente haploinsuficiente (pHaplo range cercano a 1). Esta situación implica que independientemente de la herencia de la condición asociada, el alelo no mutado no es capaz de compensar el efecto en la función proteica provocado por la delección. Dicho de otra manera, la función normal de la proteína es muy dependiente de que existen dos alelos funcionales del gen de NRXR1; las variantes de pérdida de función en uno de los alelos hacen que la proteína sintetizada no sea suficiente para que se ejerza una función normal y, como consecuencia, se produce un fenotipo clínico. Esto implica que una variante de pérdida de función en heterocigosis en este gen, como lo es la delección que se ha encontrado, puede ser una variante causal.

Con toda esta evidencia, podemos concluir que la variante NC\_000002.11:g(50801906\_50935037)del. es la variante causal del cuadro clínico del paciente, con lo que hemos llegado al diagnóstico genético de este caso y se podría elaborar el informe.

## 4.2 Probando 2

Paciente con rasgos autistas del comportamiento, junto con estreñimiento crónico, manchas café con leche y problemas de sueño como rasgos fenotípicos importantes. Ambos progenitores han sido informados como sanos y sin rasgos similares al probando. Tras realizar la búsqueda de variantes empleando los filtros establecidos en Emedgene para este trabajo, solamente hemos encontramos una posible variante:

Gen	Nomenclatura	Efecto molecular	Heredada/ <i>De novo</i>	Zigosidad	Fenotipo asociado y herencia	Clasificación ACMG
CDH22	CDH22 (NM_021248.3) :c.1916-24T>G	SNV variante intrónica	<i>De novo</i>	Heterocigoto	No conocido	Significado incierto (VUS)

-La variante CDH22 (NM\_021248.3):c.1916-24T>G se ha encontrado empleando el filtro *SFRAI de novo* y es una SNV localizada *de novo* en un intrón y se encuentra en heterocigosis en el probando. La calidad de la secuenciación es moderada en Emedgene. No tenemos información en Emedgene acerca del fenotipo que podría asociar. Cuando se realiza la búsqueda en Franklin, vemos que aparece clasificada como VUS, y los predictores *in silico* de la aplicación indican que probablemente sea benigna. Se han buscado publicaciones a cerca de la misma asociada a distintas características clínicas y no encontramos ninguna que la relacione con TEA ni rasgos del comportamiento autista. Se conoce que el gen CDH22 se ha asociado en ciertos estudios a los trastornos del espectro autista, de ahí que lo encontremos en la base de datos SFARI, de donde sale nuestro filtro; sin embargo, al tratarse de una variante intrónica no podemos ligarla con dichas publicaciones ya que desconocemos su efecto. Esto, sumado a la información que nos ofrece Franklin y otros predictores *in silico* de variantes intrónicas (splice AI, regSNP...), hace que no se pueda considerar la variante como causal del fenotipo clínico del paciente.

Por lo tanto, con la evidencia que se ha recopilado, la clasificación de esta variante se queda como de significado incierto (VUS, Clase 3) con cierta tendencia a ser probablemente benigna (Clase 2) y no se considera implicada en el fenotipo del paciente. Como consecuencia, en este caso, no se ha conseguido llegar a un diagnóstico genético. Este caso sería informado como no informativo, con una indicación de posibilidad de que el clínico solicite una reevaluación del genoma cuando él considere oportuno, bien sea por la aparición de nuevas características fenotípicas o simplemente porque los nuevos avances tecnológicos y de investigación pueden aportar más datos al estudio de este caso.

### 4.3 Probando 3

En este caso tenemos a un paciente con los siguientes datos clínicos: rasgos de comportamiento autista, trastorno por déficit de atención con hiperactividad, discapacidad intelectual, diabetes mellitus, retraso en el desarrollo del habla y del lenguaje, retraso en el desarrollo de la motricidad fina, manierismos repetitivos anormales y retraso grave del lenguaje expresivo. Ambos padres están sanos.

Aplicando nuestra búsqueda y priorización se han encontrado las siguientes variantes candidatas:

Gen	Nomenclatura	Efecto molecular	Heredada/ <i>De novo</i>	Zigosidad	Fenotipo asociado y herencia	Clasificación ACMG
FEZF2	FEZF2 (NM_018008.4) c.852+1G>A.	SNV. Mutación del sitio de <i>splicing</i> .	<i>De novo</i>	Heterocigoto	Se han asociado variantes deletéreas en FEZF2 a fenotipo asociado a trastornos del neurodesarrollo	<i>Likely Pathogenic</i>
FOXP1	NC_000003.1:g.(71579682-71584683) del.	CNV. Deleción de un tramo de la región intrónica 5' UTR.	<i>De novo</i>	Heterocigoto	Trastorno del desarrollo intelectual con trastornos del lenguaje con o sin rasgos autistas. También retraso en la adquisición del lenguaje.	Significado incierto (VUS)

-La variante NC\_000003.1:g.(71579682-71584683) del. ha sido identificada con el filtro *5' 3' UTR intronic de novo phenomatch*, se trata de una CNV tipo deleción en la región intrónica 5'UTR (antes del codón de iniciación del gen FOXP1. Vemos que el fenotipo encaja con el de nuestro caso ya que OMIM refleja variantes patogénicas en heterocigosis en este gen asocian retraso del desarrollo intelectual y del lenguaje con rasgos de comportamiento autista. Esto nos lleva a realizar la búsqueda de la variante en DECIPHER y en Franklin donde podemos ver que es una CNV de 4 kpb que afecta a la región 5 UTR del gen FOXP1 el cual tiene una alta puntuación PHaplo range (cerca de 1). Esto quiere decir que es un gen haploinsuficiente, con lo cual podría bastar una sola copia del gen alterada para provocar enfermedad. Sin embargo, vemos como su clasificación final sigue siendo de significado incierto (VUS, Clase 3) debido a la falta de evidencia clara de esta variante en el desarrollo del fenotipo (no afecta a la secuencia codificadora del gen en sí).

-Se encuentra una variante candidata empleando el filtro SFARI *de novo*, se trata de FEZF2 (NM\_018008.4) c.852+1G>A. Es una variante tipo SNV que afecta al sitio canónico de corte y empalme (*splicing*). La calidad de detección de la variante en Emedgene es alta y como se puede comprobar en la herramienta IGV (imagen adjunta) de la aplicación la variante existe y presenta una distribución alélica de 50:50. Los predictores *in silico* de la aplicación la clasifican como dañina, es decir, que es muy probable que altere el mecanismo de *splicing*. Este motivo unido al hecho de que se encuentra en un gen que aparece en SFARI (base de datos de genes asociados a autismo) hace que indagemos y realizamos una búsqueda en Franklin. Aquí, aparece clasificado como *Likely Pathogenic* (Clase 4), cumpliendo criterios de baja frecuencia poblacional y

variante *de novo*. Además es un gen con baja tolerancia a variantes de pérdida de función y alta haplosensibilidad. Los predictores *in silico* de splicing (*splice AI* y *SNV Ada*) y la consulta en el software Alamut, indican que afecta al *splicing* de una manera significativa y que afecta a un gen donde la pérdida de función es un mecanismo conocido de enfermedad. Al ser esta una variante que se encuentra un sitio canónico de *splicing* y que, según los predictores informáticos, podría condicionar la síntesis y procesamiento del RNAm, es considerada una variante de pérdida de función ya que el RNAm originado podría degradarse mediante el mecanismo NMD de *Nonsense Mediated Decay* (mecanismo celular de vigilancia del ARNm para detectar variantes tipo *nonsense* y evitar la expresión de proteínas truncadas o erróneas) o ser traducido a un producto proteico anormal. (33) Respecto al fenotipo, encontramos en la base de datos SFARI diversas publicaciones que relacionan mutaciones raras *de novo* de pérdida de función en el gen FEZF2 con el desarrollo de TEA y otros trastornos del neurodesarrollo. (34)

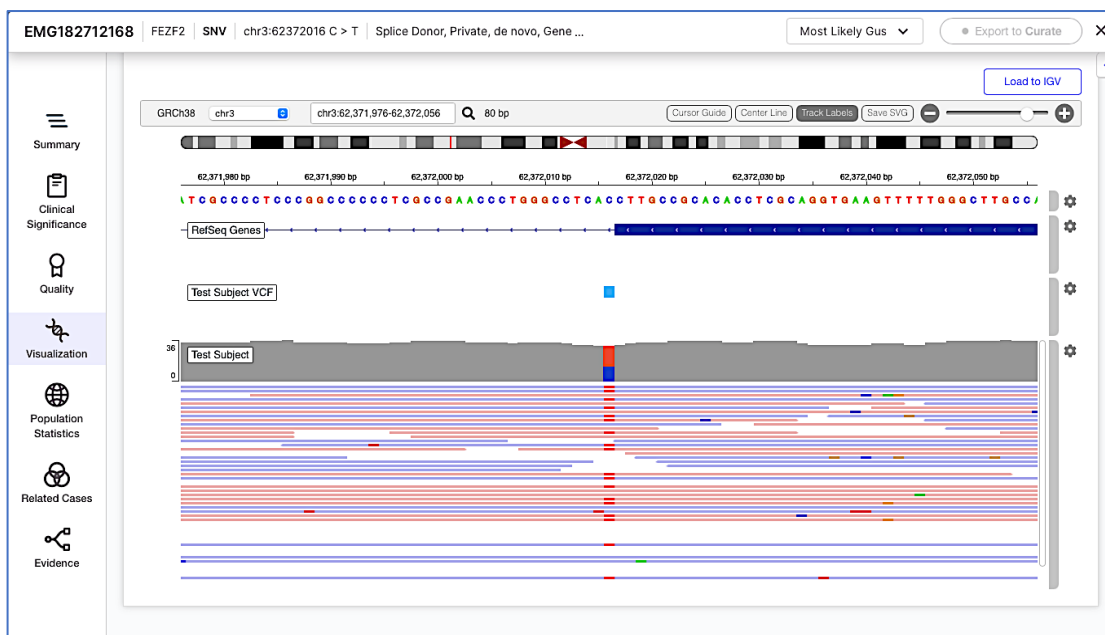


Imagen 3: Variante FEZF2 (NM\_018008.4)c.852+1G>A. en la herramienta IGV de Emedgene.

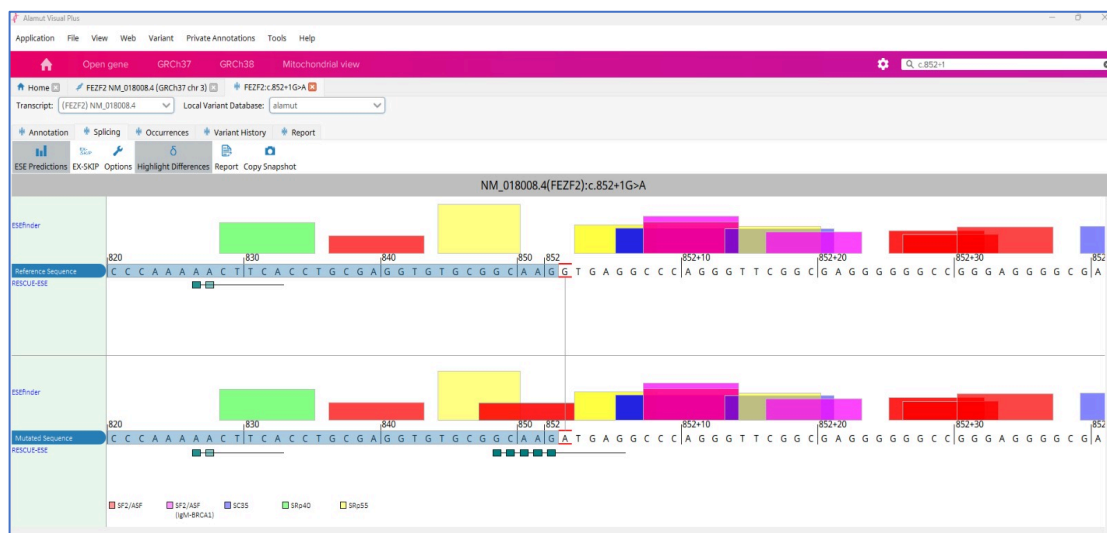


Imagen 4: Recreación de la variante FEZF2 (NM\_018008.4)c.852+1G>A. en la herramienta software Alamut.

En esta imagen, extraída de la herramienta web Alamut, podemos ver como se ha recreado la variante de nuestro probando (arriba tenemos la secuencia referencia y abajo la secuencia mutada). Las formas poligonales que aparecen en la imagen corresponden a factores de transcripción que se unen a la secuencia de ARNm para su procesamiento (corte de intrones y empalme de exones). Cuando introducimos la variante, vemos una clara diferencia, se une una nueva proteína en la secuencia mutada, mientras que en la secuencia referencia no se ha unido nada. Esto, como se ha explicado previamente, interfiere en el procesamiento de ARNm, lo que, en última instancia, puede llevar a la degradación de este o a la síntesis de una proteína no funcional, llegando a producir el fenotipo clínico.

Sabemos, por tanto, que FEZF2 (NM\_018008.4) c.852+1G>A. es una variante *de novo* cuyas predicciones *in silico* (tanto en Emedgene como en Franklin) son altamente indicativas de pérdida de función en un gen ampliamente conocido por su asociación con el TEA. Por ello podemos concluir con alta probabilidad que se trata de la variante causal de nuestro probando, completando así el diagnóstico genético de este caso.

#### 4.4 Probando 4

El paciente ha sido informado con las siguientes características clínicas: comportamiento autista, crisis epilépticas, hemiparesia y parálisis facial unilateral. Ambos progenitores están sanos.

Gen	Nomenclatura	Efecto molecular	Heredada/ <i>De novo</i>	Zigosidad	Fenotipo asociado y herencia	Clasificación ACMG
RFX3	RFX3 (NM_00128211) c.43_44del (p.Thr15Leufs*7)	SNV con efecto <i>frameshift</i>	Heredada (padre)	Heterocigoto	Trastorno del espectro autista	<i>Likely Pathogenic</i>
AUTS2, CT66	NC_000007.13: g.(69596387-69600390) del.	CNV. Delección de una secuencia codificadora y región 5' UTR	<i>De novo</i>	Heterocigoto	Trastorno del espectro autista por variantes deletéreas en AUTS2. También trastorno del desarrollo intelectual.  AD	Significado incierto (VUS)
RHOBTB2	RHOBTB2 (NM_015178.3): c.293G>A p.(Gly98Glu)	SNV con efecto <i>missense</i>	<i>De novo</i>	Heterocigoto	Encefalopatía epiléptica y del desarrollo (síndrome RHOBTB2)  AD	<i>Likely Pathogenic</i>

-La variante RFX3 (NM\_001282116.2) c.43\_44del (p.Thr15Leufs\*7) aunque si se relaciona con el TEA (existen publicaciones al respecto) no explica la totalidad del fenotipo y además se trata de un variante heredada por parte del padre (del cual no tenemos información de que presente o haya presentado rasgos autistas), con lo que, al aplicar este criterio a la clasificación de la variante, queda clasificada como probablemente benigna (Clase 2).

-Aplicando el filtro *5' 3' untranslated región* se ha encontrado una CNV de 4003 kbp, NC\_000007.13: g.(69596387-69600390) del. Se trata de una delección desde la posición 69 59 63 87 hasta la posición 69 60 03 90 del brazo largo del cromosoma 7. Esta delección solapa parcialmente con los genes AUTS2 (afecta a su región 5'UTR) y CT66. Esta CNV llama la atención debido a que, como se muestra en Emedgene, OMIM describe la existencia de delecciones exónicas en el gen AUTS2 que se asocian a un fenotipo sindrómico que incluye discapacidad intelectual, autismo, baja estatura, microcefalia, parálisis cerebral y malformaciones faciales. Sin embargo, encontramos varios problemas: por una parte, no se trata de una variante exónica, sino que afecta a la región 5'UTR; tampoco explica completamente el fenotipo (la hemiparesia es un dato muy característico y tampoco se habla de crisis en OMIM) y además presenta un parámetro de calidad moderado en Emedgene debido a una baja profundidad de lectura.

Realizamos una búsqueda en DECIPHER, donde encontramos una CNV con coordenadas y tamaño semejante, y si bien aparece asociada a los TEA, se clasifica la variante como VUS. Teniendo en cuenta estos criterios, la clasificación de la variante la sitúa como probablemente benigna (Clase 2) o de significado incierto (Clase 3), pero esto sumado a que los parámetros de calidad indican que la calidad es moderada, hace que la descartemos como variante causal.

-Se ha localizado, empleando el filtro *rare de novo*, la variante RHOBTB2 (NM\_015178.3): c.293G>A p. (Gly98Glu) en heterocigosis. El fenotipo asociado a esta variante es característicamente semejante al de nuestro caso ya que como se describe en OMIM se asocia a crisis tónico-clónicas en el primer año de vida, junto con un retraso en el neurodesarrollo y retraso mental. Además, describe que algunos pacientes presentan hemiparesia junto con las crisis. Esto nos hace sospechar, por lo que se lleva a cabo una búsqueda en Franklin, donde se clasifica la variante como probablemente patogénica (Clase 4) y vemos que las predicciones *in silico* de la herramienta, indican una alta probabilidad de ser una variante deletérea para la función del gen. Se trata de un gen con baja tolerancia a las variantes tipo *missense* como se puede comprobar en el browser de DECIPHER. Además, existen numerosas publicaciones que indican la asociación con una encefalopatía epiléptica y del neurodesarrollo. En la base de datos [www.rarechromo.org](http://www.rarechromo.org) está descrito el síndrome RHOBTB2, donde se indica que está provocado por variantes *missense* en una de las copias del gen en el cromosoma 8 (AD). El fenotipo descrito en esta página es muy similar al de nuestro probando (epilepsia, retraso en el neurodesarrollo, retraso en el crecimiento, hemiparesia...).



Imagen 5: Variante RHOBTB2 (NM\_015178.3): c.293G>A p. (Gly98Glu) en la herramienta IGV de Emedgene.

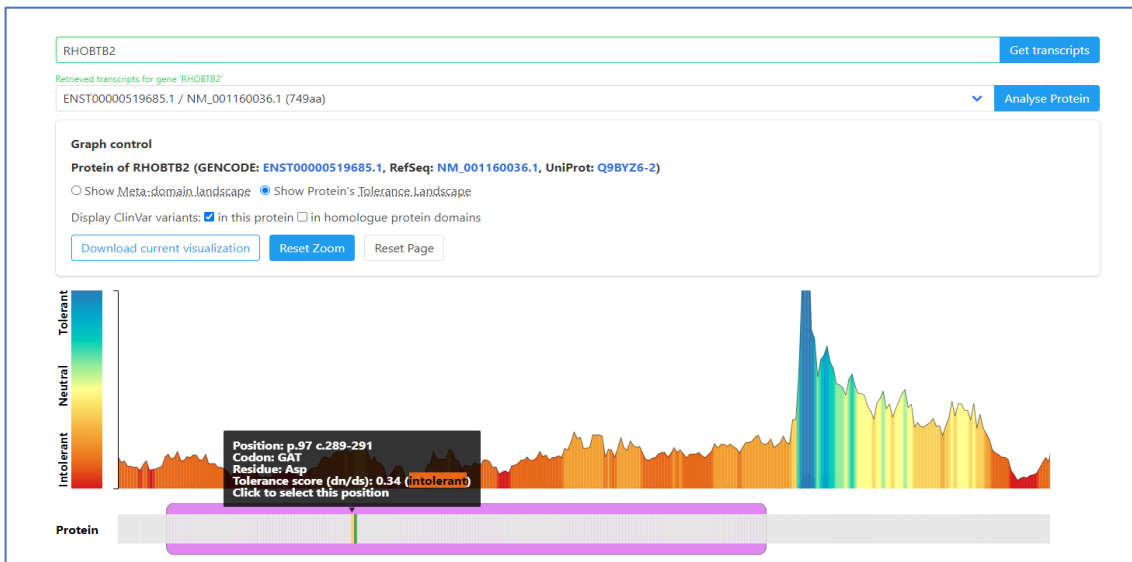


Imagen 6: Visualización del gen RHOBTB2 en la herramienta MetaDome, donde vemos que se trata de una región genética intolerante para variantes tipo *missense*.

Con esta evidencia, se concluye que la variante RHOBTB2 (NM\_015178.3): c.293G>A p. (Gly98Glu) tiene altas probabilidades de ser la causal de la patología del paciente y por lo tanto confirmaría el diagnóstico genético del probando número 4: Encefalopatía epiléptica y del neurodesarrollo cuyo fenotipo OMIM corresponde al código numérico 618004.

## 4.5 Probando 5

Este último caso se trata de un paciente diagnosticado de TEA, retraso en la adquisición del lenguaje y que presenta crisis epilépticas no filiadas. Ambos progenitores están sanos.

En este caso hemos empleado los sets de filtros preestablecidos para este trabajo sin encontrar apenas variantes candidatas. Sin embargo, damos con una CNV de gran tamaño localizada en el cromosoma X (filtro *CNV > 1Mb*) que debido a que se trata de una duplicación de 3063788 Kbp, abarca más de 30 genes afectados por la variante y por ello múltiples fenotipos posibles asociados (entre ellos trastornos del neurodesarrollo y TEA).

Gen	Nomenclatura	Efecto molecular	Heredada/ <i>De novo</i>	Zigosidad	Fenotipo asociado y herencia	Clasificación ACMG
CLCN5, DGKK, XAGE1A. ...	NC_000023.10, X:g.(49604544- 52668332) dup.	CNV. Duplicación que afecta a la secuencia de 40 genes.	<i>De novo</i>	Heterocigoto	Múltiples (seis genes con fenotipo patológico asociado en Emedgene)	Patogénica

-Se encuentra la variante NC\_000023.10, X:g.(49604544-52668332) dup. Utilizando el filtro CNV >1 Mb. La puntuación de calidad agregada en Emedgene es alta y se trata de una variante *de novo* e heterocigosis. Decidimos profundizar en esta variante realizando su búsqueda en Franklin, donde añadiendo la información de ser una variante *de novo* (dato clave que nos proporciona el hecho de trabajar con tríos) se clasifica como Patogénica con una puntuación de 1 (la más alta). Realizamos una búsqueda por triplosensibilidad (sensibilidad o probabilidad de función alterada en genes en los que existe más de dos copias) en los genes y regiones afectados por la CNV y encontramos la región Xp11.22-p11.23 que contiene al gen SHROOM4 con una alta puntuación en triplosensibilidad, que indican que los genes incluidos en esa región son poco tolerantes a las copias adicionales, lo que implica que una duplicación en esa zona puede provocar un fenotipo clínico. Para hacer una comprobación del fenotipo asociado, se hace una búsqueda en OMIM de dicha región y encontramos una entrada que caracteriza un síndrome específico de esta región (*Chromosome Xp11.23-p11.22 Duplication syndrome, 300801*) cuyos rasgos fenotípicos concuerdan con los de nuestro probando: rasgos autistas, obstinación, retraso mental de leve a grave, mala articulación del habla, crisis de ausencia subclínica, descargas paroxísticas difusas y focales y descargas continuas de ondas en espiga durante el sueño.

Sabemos por lo tanto que se trata de una duplicación de gran tamaño (en el TEA las CNVs de gran tamaño tienen una importancia patogénica causal muy relevante) *de novo*, que Franklin clasifica como Patogénica y además abarca una región con una alta triplosensibilidad descrita y que se asocia a un síndrome conocido de alteraciones del neurodesarrollo y crisis epilépticas (cuadro perfectamente superponible al probando número cinco). Es por ello que, juntando toda esta evidencia, se puede afirmar que NC\_000023.10, X:g.(49604544-52668332) dup. es la variante causal del cuadro clínico del paciente y por tanto confirmaría el diagnóstico genético de *Chromosome Xp11.23-p11.22 duplication syndrome* cuyo fenotipo OMIM corresponde al código 300801.

## 5. Discusión

El TEA es un trastorno del neurodesarrollo que aparece típicamente en la primera infancia y produce dificultades en interacción social y patrones conductuales anormales en los niños que lo padecen. Se trata de una condición cuyo diagnóstico se realiza en base a una serie de criterios clínicos establecidos por la DSM y con una incidencia claramente ascendente a lo largo de los últimos años.

Como su nombre indica, el TEA, consiste en un amplio espectro de condiciones clínicas con las características comunes que se acaban de describir. Esto implica que la discapacidad que produce es muy variable en función de la gravedad del cuadro. Es por ello por lo que tanto un diagnóstico temprano como una estimación pronóstica se hacen necesarias para una buena actuación terapéutica. Además, la mayoría de los casos se asocian a algún tipo de comorbilidad (psiquiátrica o no psiquiátrica), que por su puesto también necesitan tratamiento.

Hoy en día, aunque sabemos que existen factores de riesgo ambientales, se conoce que el TEA es un trastorno predominantemente genético, con una heredabilidad calculada de en torno al 83%, que es una de las más altas de los trastornos neuropsiquiátricos.

La práctica habitual de estudio genético en estos casos es hacer un análisis de exoma (WES) en el probando únicamente. Como hemos mencionado en la introducción, esto, junto con el CMA y el estudio del cariotipo, nos ofrecen una rentabilidad diagnóstica del 25%, según datos de la Fundación Pública de Medicina Xenómica, mientras que se estima que la variación rara (la que buscamos en este tipo de estudios y pruebas genéticas) representa el 45% de los casos de TEA. Esta diferencia del 20% (variación perdida o *missing variation*) es la que tratamos de abordar en este trabajo, llevando a cabo algo de mucha más envergadura (WGS en tríos) para intentar conseguir un diagnóstico y valorar si esta estrategia puede ser usada en la práctica clínica habitual.

El uso de tríos en este tipo de estudios es muy útil pues permite saber inmediatamente que variantes genéticas son heredadas y cuáles son *de novo*; esto nos permite saber si las variantes de significado incierto (la gran mayoría) o desconocidas son variantes benignas (si están presentes en alguno de los progenitores, es decir, son heredadas, y estos están sanos, eso indicaría, salvo en variantes de penetrancia incompleta, que son variantes que no tienen efecto, y por lo tanto las descartaríamos) o son causales (si una variante incierta en un gen candidato es *de novo*, es decir, no está presente en ninguno de los progenitores, es muy probable que sea la variante causal).

Como se puede comprobar en el apartado de resultados se ha conseguido llegar, mediante el análisis y clasificación de variantes obtenidas por secuenciación de genoma completo, al diagnóstico genético en cuatro de los cinco probandos afectados que se han incluido en este trabajo. Es importante señalar que las variantes causales a las que se ha llegado durante el análisis son de naturaleza distinta entre sí en lo que respecta su efecto molecular sobre el genoma, demostrando que cualquier tipo de variante puede participar en la etiología del TEA.

Por un lado, en el probando 3, tenemos una SNV que afecta al sitio de *splicing* del gen FEZF2. Como hemos visto mediante predictores computacionales esta variante, hace que se una un factor de transcripción anómalo interfiriendo en el procesamiento de ARNm y dando lugar, en última instancia, a una pérdida de función en la proteína por mecanismos como el NMD.

En el probando 4 se ha llegado, guiándonos esta vez por el fenotipo característico del paciente, a una SNV *de novo* con efecto *missense* en un gen con baja tolerancia a variantes de esta clase y cuya pérdida de función da lugar a un síndrome clínico ampliamente descrito en la literatura. (35) (36)

En lo que respecta al probando 5, la búsqueda y clasificación de variantes ha sido, quizás, un poco más compleja. Tras emplear los filtros predefinidos en Emedgene emparejados por fenotipo (son en principio estos con los que esperamos encontrar las variantes causales con más probabilidad) no se había dado con ninguna candidata. Sin embargo, empleando un filtro de CNVs de gran tamaño *de novo* se ha encontrado una duplicación *de novo*, que abarca más de treinta regiones codificadoras en el genoma del paciente. Entre ellas, una, a la que hemos llegado gracias a una búsqueda por triplosensibilidad en la web Franklin, cuya alteración se asocia a un síndrome clínico que coincide claramente con fenotipo de nuestro paciente. (37)

El probando 1 cobra especial importancia en este trabajo ya que se trata de un paciente con un estudio previo de exoma (WES) cuyo resultado no había conseguido dilucidar cuál era la variante causal y, como se expone al comienzo de esta discusión, en este trabajo pretendemos comprobar si pudiera alcanzarse una mejora diagnóstica con WGS. Además, se sabía que su madre, presentaba un fenotipo similar (rasgos autistas y retraso del lenguaje) algo que nos podía llevar a pensar que existía algún componente genético heredado. Como se expone en los resultados gracias a los filtros predefinidos se ha conseguido encontrar una CNV (microdelección de 133 131 kbp en el cromosoma 2) que se solapa parcialmente con el gen de la neurexina 1 (NRXN1). Ahora bien, la pregunta que ha de plantearse es ¿por qué no se ha conseguido detectar esta variante en un análisis de exoma? La búsqueda de la variante en la web Franklin nos da una clave para responder a ello. Como vemos en la figura, nuestra variante abarca los exones 6, 7 y 8 del gen, pero comenzando y acabando en el interior de dos regiones intrónicas de mucha mayor longitud en comparación con los exones. Esto hace que muy probablemente la variante haya pasado desapercibida para una prueba en que no se secuencian las regiones intrónicas, ya que, para el cálculo de diferencia de dosis génica, los softwares utilizan las regiones secuenciadas, y en un exoma, al abarcar la delección solo dos exones, al software le es mucho más difícil detectarlas y a veces, como en este caso, no es capaz de hacerlo.

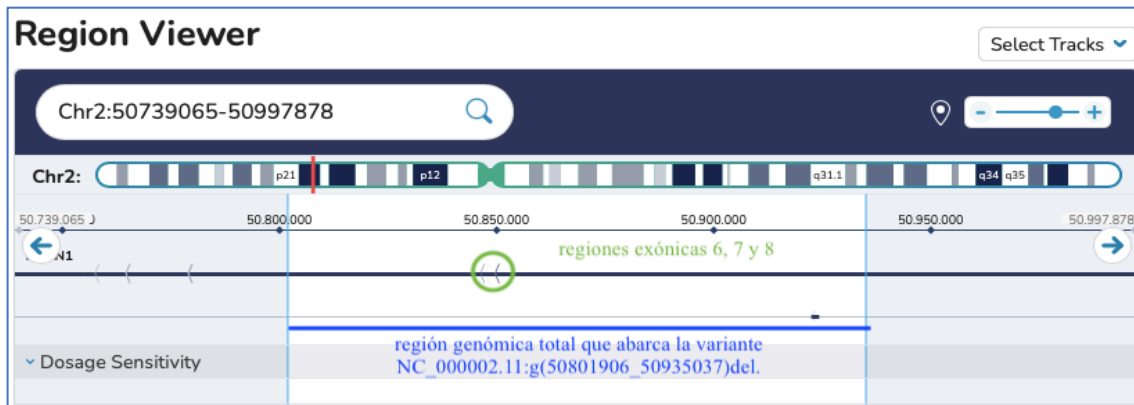


Imagen 6: Visualización de la variante causal del probando 1 en la herramienta Region Viewer de Franklin. En verde, destacadas las regiones exónicas. En azul, la totalidad de la variante en el cromosoma 2.

Esta que se acaba de mostrar, es una de las ventajas del WGS. Como hemos visto en la introducción, la posibilidad que ofrece WGS en lo que respecta a la captura de *indels*, variantes del ADN mitocondrial (ADNmt), variante en regiones reguladoras no codificantes (como son las variantes en 5'3'UTR) y variantes estructurales (SV, incluyendo variantes de número de copias [CNV], inversiones, inserciones mayores, disomías uniparentales [UPD] y expansiones de repeticiones en tándem o *short tandem repeats* [STR o TRE]), es diferencial respecto al análisis de exoma. De todas estas variantes, muchas se encuentran en regiones no codificantes del genoma y estas son las que conforman el llamado “ADN oscuro”. Una hipótesis estudiada en proyectos de investigación del grupo de la Fundación Pública de Medicina Xenómica, es que este “ADN oscuro” sea el responsable de la variación perdida (aquella no capturada con los recursos bioinformáticos y pruebas genéticas diagnósticas actuales) y que nos impide llegar al diagnóstico en aproximadamente un 20% de los casos. (38)

Durante el análisis de variantes realizado en este trabajo, se han identificado diversas variantes candidatas en los probandos, que se encontraban en regiones no codificantes del genoma (el denominado “ADN oscuro”). Algunas de ellas son:

- Variantes del sitio de *splicing*. Se trata de una región de corte y empalme de exones, cuya alteración puede dar lugar a transcripción anómala y finalmente a alteraciones en la función de las proteínas. Entre ellas, por ejemplo, tenemos variante causal del probando 3, FEZF2 (NM\_018008.4) c.852+1G>A.
- Variantes en la región 5'3'UTR. Son regiones inmediatamente anteriores a las codificadoras que se conoce que tienen un papel relevante en la regulación de la transcripción.
- Variantes en regiones intrónicas, cuyo análisis es ciertamente difícil al carecer de información en los recursos bioinformáticos acerca de su implicación con la enfermedad.
- Variantes en ADN mitocondrial.
- Variantes tipo STR, repeticiones en tándem que se pueden distribuir a lo largo de todo el genoma. Hoy en día, sí se pueden estudiar específicamente con técnicas dirigidas a ellas; lo que no se puede hacer es un estudio masivo de ellas en todo el genoma.

El problema con la mayoría de estas variantes (a excepción de las de *splicing* cuyo papel patogénico en el TEA es conocido y se pueden identificar mediante estudios de exoma) es que apenas existen estudios a cerca de su implicación o posible asociación con el TEA, ya que estas serían imposibles de identificar de no ser por un análisis de genoma completo (es una prueba novedosa y ni mucho menos disponible en todos los centros, debido a su complejidad técnica y su coste económico). Es por ello, que sería casi imposible informar alguna de estas como causal en un diagnóstico genético. Sin embargo, todo ello ofrece un amplísimo campo de investigación para tratar de buscar lo que conocemos como “variación perdida” y por tanto un futuro prometedor en el avance científico y conocimiento del TEA.

Además, situaciones como la que se ha dado en el probando 1 de este trabajo, podrían ser responsables de la mejora diagnóstica que ya ha empezado a demostrar el WGS en estudios de gran calibre como el de Lowther et al. (26) donde veíamos que ya ha demostrado una ligera superioridad diagnóstica frente al uso combinado de CMA y análisis de exoma (WES). Debido a la novedad de la técnica, su uso todavía está por establecerse con unos parámetros de secuenciación y bioinformáticos comunes que permita hacer más estudios como este e incluso superiores.

Por último, respecto al probando 2, no hemos encontrado ninguna variante causal. Este caso representa algo que es una realidad en el día a día del diagnóstico genético en TEA, pues como hemos dicho, aún con técnicas tan novedosas como el WGS, que usamos en este trabajo, la capacidad diagnóstica máxima estimada se encuentra en el 45%. Esto se debe a que, aunque tenemos la capacidad de “leer” el genoma completo, no sabemos que consecuencia tienen esas variantes más allá de lo que se pueda estimar con predictores bioinformáticos. Los estudios de investigación que ya se están llevando a cabo en el presente son los que nos aportarán esta información que podrá ser aplicada en el diagnóstico genético en el futuro inmediato. Debido a esto, una de las recomendaciones en los estudios genéticos no informativos es la de que el facultativo que sigue al paciente solicite una reevaluación del estudio pasado un tiempo para que el genetista pueda comprobar sin los estudios de investigación y la mejora tecnológica realizados en ese periodo de tiempo han aportado información contrastada que pueda permitir que se llegue a un diagnóstico genético en ese caso.

Para estos pacientes existen, además, otras posibilidades más allá de los estudios genéticos llevados a cabo en su área de referencia, en caso de que estos no fuesen esclarecedores, y es entrar en algún proyecto de investigación básica que pueda ayudar a arrojar luz sobre la causa de su enfermedad, aunque sea dentro del campo de la investigación. Por ejemplo, el Consorcio Internacional de enfermedades raras (IRDiRC), del que España es signatario, propone que se intente el diagnóstico de potenciales afectos de enfermedades raras genéticas. En un año a partir de los síntomas y en caso de que den negativo en las pruebas realizadas en su área de referencia, el IRDiRC propone que entren en programas de investigación conectados internacionalmente.

Para los pacientes con enfermedades raras no diagnosticadas de origen genético hay dos iniciativas en España, una es el programa ENOD del CIBERER y a mucha más escala el programa de EERR de IMPaCT Genómica, una infraestructura del Instituto de Salud Carlos III (ISCIII) para el desarrollo de la Medicina de precisión en España. La Fundación Pública de Medicina Xenómica participa activamente en ambas iniciativas con diversos profesionales y grupos de investigación involucrados. (39) (40)

## 6. Conclusiones

El objetivo de este trabajo era el de realizar el proceso de estudio y diagnóstico genético de cinco probandos afectados de TEA mediante el análisis de variantes de su genoma completo (*whole genome sequencing*). A través de este primer objetivo se ha querido dar una visión de futuro a cerca de la posibilidad de que el análisis de variantes por WGS pueda mejorar la capacidad diagnóstica en el TEA respecto a las pruebas genéticas empleadas en la práctica clínica habitual, el CMA y WES.

De los cinco probandos incluidos en este trabajo, se ha conseguido llegar al diagnóstico genético en cuatro de ellos. Los pacientes incluidos eran fenotípicamente diferentes entre sí. Por un lado, los probandos 1, 2 y 3 tenían un cuadro de TEA más típico (con comorbilidades más frecuentes como retraso de la adquisición del lenguaje y discapacidad intelectual) mientras que los dos últimos presentaban una clínica más florida asociando crisis epilépticas y hemiparesia en el probando 4. Con el análisis de variantes mediante WGS se han encontrado variantes causales con distintas características en cada uno de ellos.

- El probando 2 ha sido negativo para la prueba sin que esto descarte completamente que exista una causa genética. Los siguientes pasos a seguir en el estudio de este caso serían el seguimiento clínico del paciente con posibilidad de repetir la prueba en un futuro y su inclusión en programas nacionales e internacionales de investigación de enfermedades raras no diagnosticadas de origen genético.
- En el probando 3 se ha encontrado una variante causal *de novo* del sitio de splicing deletérea para la función de la proteína en un gen con baja tolerancia a variantes de pérdida de función, alta haplosensibilidad y asociado al TEA, que explica el fenotipo del paciente (TEA con retraso en el lenguaje y discapacidad intelectual).
- En el probando 4 se ha dado con una SNV *de novo* en un gen con baja tolerancia a las variantes tipo *missense* en la región en concreto en la que se encuentra la variante. Alteraciones en este gen dan lugar a un fenotipo clínico sindrómico de encefalopatía epiléptica con alteraciones del movimiento.
- En el probando 5 se ha dado con una duplicación de gran tamaño *de novo*, que abarcaba entre otras una región sensible para este tipo de variantes. Su alteración da lugar un síndrome genético de alteraciones en el neurodesarrollo y crisis epilépticas además de otros hallazgos que podrían darse en el futuro del paciente.
- Crucialmente, se ha conseguido llegar al diagnóstico en el probando 1. Este paciente tenía un estudio previo con WES que no había conseguido dar con la variante causal. Además, se conocía que la madre presentaba un cuadro clínico similar, con rasgos del comportamiento autista y retraso del lenguaje, por lo que las sospechas de una causa genética eran altas.

Investigando con diferentes herramientas bioinformáticas web, se ha visto que se trataba de una CNV que abarca fundamentalmente regiones intrónicas del genoma (razón por la cual posiblemente no se haya detectado con una secuenciación de exoma). Es por ello, que este trabajo va en consonancia con algunos estudios que ya se publicado demostrando la mayor eficacia del WGS como prueba de primera línea frente al WES y al CMA (incluso combinados), para el diagnóstico genético del TEA.

Además, gracias al empleo de WGS se han encontrado durante la realización de este trabajo, diversas variantes candidatas en regiones no codificantes del genoma, cuya contribución al riesgo de la enfermedad es difícilmente cuantificable con los recursos disponibles hoy en día, pero que suponen un amplio campo de investigación en la estructura genética y por tanto en el progreso científico del TEA.

## 7. Referencias bibliográficas

1. Autismo [Internet]. Who.int. [citado 14 de mayo de 2024]. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/autism-spectrum-disorders>
2. Khan NZ, Gallo LA, Arghir A, Budisteanu B, Budisteanu M, Dobrescu I, et al. Autism and the grand challenges in global mental health. *Autism Res* [Internet]. 2012 [citado 5 de octubre de 2023];5(3):156-9. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/22605618/>
3. Manoli DS, Matthew W. State. Autism spectrum disorder genetics and the search for pathological mechanisms. *Am J Psychiatry* [Internet]. 2021 [citado 5 de octubre de 2023];178(1):30-8. Disponible en: <http://dx.doi.org/10.1176/appi.ajp.2020.20111608>
4. Masi A, DeMayo MM, Glozier N, Guastella AJ. An overview of autism spectrum disorder, heterogeneity and treatment options. *Neurosci Bull* [Internet]. 2017 [citado 15 de mayo de 2024];33(2):183-93. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/28213805/>
5. Martín Del Valle F, García Pérez A, Losada R, Pozo D. Trastornos del espectro del autismo [Internet]. *Aeped.es*. [citado 13 de octubre de 2023]. Disponible en: <https://www.aeped.es/sites/default/files/documentos/08.pdf>
6. Brignell A, Harwood RC, May T, Woolfenden S, Montgomery A, Iorio A, et al. Overall prognosis of preschool autism spectrum disorder diagnoses. *Cochrane Libr* [Internet]. 2022 [citado 15 de mayo de 2024];2022(9). Disponible en: <https://pubmed.ncbi.nlm.nih.gov/36169177/>
7. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5 (R))*. 5.a ed. Arlington, TX, Estados Unidos de América: American Psychiatric Association Publishing; 2013.
8. 10th revision of the international classification of diseases , chapter V (F): Mental and behavioural disorders (including disorders of psychological development. Ginebra, Suiza: World Health Organization; 1990.
9. Maenner MJ, Warren Z, Williams AR, Amoakohene E, Bakian AV, Bilder DA, et al. Prevalence and characteristics of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, United States, 2020. *Morb Mortal Wkly Rep Surveill Summ* [Internet]. 2023 [citado 2 de enero de 2024];72(2):1-14. Disponible en: <https://www.cdc.gov/mmwr/volumes/72/ss/ss7202a1.htm>
10. Lyall K, Croen L, Daniels J, Fallin MD, Ladd-Acosta C, Lee BK, et al. The changing epidemiology of autism spectrum disorders. *Annu Rev Public Health* [Internet]. 2017;38(1):81-102. Disponible en: <http://dx.doi.org/10.1146/annurev-publhealth-031816-044318>
11. Loomes R, Hull L, Mandy WPL. What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *J Am Acad Child Adolesc Psychiatry* [Internet]. 2017 [citado 13 de octubre de 2023];56(6):466-74. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/28545751/>

12. Tubío-Fungueiriño M, Cruz S, Sampaio A, Carracedo A, Fernández-Prieto M. Social camouflaging in females with autism spectrum disorder: A systematic review. *J Autism Dev Disord* [Internet]. 2021 [citado 15 de mayo de 2024];51(7):2190-9. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/32926304/>
13. Hervás Zúñiga A, Maraver García N. Los trastornos del espectro autista [Internet]. *Pediatriaintegral.es*. [citado 13 de octubre de 2023]. Disponible en: [https://www.pediatriaintegral.es/wp-content/uploads/2020/xxiv06/03/n6-325e1-21\\_AmaiaHervas.pdf](https://www.pediatriaintegral.es/wp-content/uploads/2020/xxiv06/03/n6-325e1-21_AmaiaHervas.pdf)
14. Baio J, Wiggins L, Christensen DL, Maenner MJ, Daniels J, Warren Z, et al. Prevalence of autism spectrum disorder among children aged 8 years — autism and developmental disabilities monitoring network, 11 sites, United States, 2014. *Morb Mortal Wkly Rep Surveill Summ* [Internet]. 2018 [citado 13 de octubre de 2023];67(6):1-23. Disponible en: <http://dx.doi.org/10.15585/mmwr.ss6706a1>
15. Valverde Gómez M, Raventós Šimić J, editores. Autismo y trastorno de la comunicación social. Síndromes clínicos en psiquiatría infanto-juvenil. Trastorno del espectro autista. Aula-Salud
16. Trastornos del espectro autista. Detección precoz, herramientas de cribado [Internet]. *Pap.es*. [citado 15 de mayo de 2024]. Disponible en: <https://pap.es/articulo/1076/trastornos-del-espectro-autista-deteccion-precoz-herramientas-de-cribado>
17. Jones W, Klaiman C, Richardson S, Aoki C, Smith C, Minjarez M, et al. Eye-tracking-based measurement of social visual engagement compared with expert clinical diagnosis of autism. *JAMA* [Internet]. 2023 [citado 15 de mayo de 2024];330(9):854. Disponible en: <https://jamanetwork.com/journals/jama/fullarticle/2808996>
18. Lord C, Elsabbagh M, Baird G, Veenstra-Vanderweele J. Autism spectrum disorder. *Lancet* [Internet]. 2018 [citado 20 de octubre de 2023];392(10146):508-20. Disponible en: [http://dx.doi.org/10.1016/s0140-6736\(18\)31129-2](http://dx.doi.org/10.1016/s0140-6736(18)31129-2)
19. Wang L, Wang B, Wu C, Wang J, Sun M. Autism spectrum disorder: Neurodevelopmental risk factors, biological mechanism, and precision therapy. *Int J Mol Sci* [Internet]. 2023 [citado 20 de octubre de 2023];24(3):1819. Disponible en: <http://dx.doi.org/10.3390/ijms24031819>
20. Taylor LE, Swerdfeger AL, Eslick GD. Vaccines are not associated with autism: An evidence-based meta-analysis of case-control and cohort studies. *Vaccine* [Internet]. 2014;32(29):3623-9. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0264410X14006367>
21. Sandin S, Lichtenstein P, Kuja-Halkola R, Hultman C, Larsson H, Reichenberg A. The heritability of autism spectrum disorder. *JAMA* [Internet]. 2017 [citado 24 de noviembre de 2023];318(12):1182. Disponible en: <http://dx.doi.org/10.1001/jama.2017.12141>
22. Carracedo Á. «Avances en genética y TEA» [Internet]. 17 de Noviembre de 2018. Disponible en: <http://aetapi.org/videos-xix-congreso-nacional-de-autismo/>

23. Genovese A, Butler MG. Clinical assessment, genetics, and treatment approaches in autism spectrum disorder (ASD). *Int J Mol Sci* [Internet]. 2020 [citado 3 de enero de 2024];21(13):4726. Disponible en: <https://www.mdpi.com/1422-0067/21/13/4726>
24. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. 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. *Genet Med* [Internet]. 2015 [citado 26 de febrero de 2024];17(5):405-24. Disponible en: <https://www.nature.com/articles/gim201530>
25. Whole-genome sequencing (WGS) [Internet]. Illumina.com. [citado 2 de enero de 2024]. Disponible en: <https://emea.illumina.com/techniques/sequencing/dna-sequencing/whole-genome-sequencing.html>
26. Lowther C, Valkanas E, Giordano JL, Wang HZ, Currall BB, O’Keefe K, et al. Systematic evaluation of genome sequencing for the diagnostic assessment of autism spectrum disorder and fetal structural anomalies. *Am J Hum Genet* [Internet]. 2023 [citado 3 de enero de 2024];110(9):1454-69. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/37595579/>
27. Steri M, Idda ML, Whalen MB, Orrù V. Genetic variants in mRNA untranslated regions. *Wiley Interdiscip Rev RNA* [Internet]. 2018 [citado 9 de abril de 2024];9(4):e1474. Disponible en: <http://dx.doi.org/10.1002/wrna.1474>
28. Diccionario de genética del NCI [Internet]. Instituto Nacional del Cáncer. 2012 [citado 5 de marzo de 2024]. Disponible en: <https://www.cancer.gov/espanol/publicaciones/diccionarios/diccionario-genetica/def/haploinsuficiencia>
29. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* [Internet]. 2020 [citado 9 de abril de 2024];581(7809):434-43. Disponible en: <https://www.nature.com/articles/s41586-020-2308-7>
30. Whiffin N, Minikel E, Walsh R, O’Donnell-Luria AH, Karczewski K, Ing AY, et al. Using high-resolution variant frequencies to empower clinical genome interpretation. *Genet Med* [Internet]. 2017 [citado 1 de marzo de 2024];19(10):1151-8. Disponible en: <http://dx.doi.org/10.1038/gim.2017.26>
31. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of chromosomal imbalance and phenotype in humans using ensembl resources. *Am J Hum Genet* [Internet]. 2009;84(4):524-33. Disponible en: <http://dx.doi.org/10.1016/j.ajhg.2009.03.010>
32. Liu Z, Xin B, Smith IN, Sency V, Szekely J, Alkelai A, et al. Hemizygous variants in protein phosphatase 1 regulatory subunit 3F (PPP1R3F) are associated with a neurodevelopmental disorder characterized by developmental delay, intellectual disability and autistic features. *Hum Mol Genet* [Internet]. 2023 [citado 28 de marzo de 2024];32(20):2981-95. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/37531237/>
33. Garber A, Weingarten LS, Abreu NJ, Elloumi HZ, Haack T, Hildebrant C, et al. Rare predicted deleterious *FEZF2* variants are associated with a neurodevelopmental phenotype. *Am J Med Genet A* [Internet]. 2024 [citado 1 de abril de 2024]; Disponible en: <https://pubmed.ncbi.nlm.nih.gov/38425142/>

34. Cheng MM, Cao YY. The NMD escape mechanism and its application in disease therapy. Yi Chuan [Internet]. 2020 [citado 13 de mayo de 2024];42(4). Disponible en: <https://pubmed.ncbi.nlm.nih.gov/32312704/>
35. Straub J, Konrad EDH, Grüner J, Toutain A, Bok LA, Cho MT, et al. Missense variants in RHOTB2 cause a developmental and epileptic encephalopathy in humans, and altered levels cause neurological defects in Drosophila. Am J Hum Genet [Internet]. 2018 [citado 14 de mayo de 2024];102(1):44-57. Disponible en: <http://dx.doi.org/10.1016/j.ajhg.2017.11.008>
36. Rarechromo.org. [citado 14 de mayo de 2024]. Disponible en: <https://rarechromo.org/media/singlegeneinfo/Single%20Gene%20Disorder%20Guides/RHOTB2%20syndrome%20QFN.pdf>
37. Entry - #300801 - CHROMOSOME Xp11.23-p11.22 DUPLICATION SYNDROME - OMIM [Internet]. Omim.org. [citado 14 de mayo de 2024]. Disponible en: <https://omim.org/entry/300801?search=Chromosome%20Xp11.23-p11.22%20duplication%20syndrome&highlight=%28chromosomal%7Cchromosome%29%2C%28syndrome%7Csyndromic%29%2Cduplication%2Cxp1123p1122>
38. Rodríguez Fontenla MC, Carracedo Álvarez Á, Fuentes Pita P, Carballo Pacoret P. El papel del ADN “oscuro” en la base genética de los Trastornos del Espectro Autista (TEA): caracterización de secuencias altamente repetitivas del genoma mediante estrategias de secuenciación de lectura larga. Fundación Alicia Koplowitz; 2024
39. CIBERER-ENOD [Internet]. Ciberer.es. [citado 15 de mayo de 2024]. Disponible en: <https://enod.ciberer.es>
40. IMPaCT – Infraestructura de Medicina de Precisión asociada a la Ciencia y la Tecnología [Internet]. Iscii.es. [citado 15 de mayo de 2024]. Disponible en: <https://impact.iscii.es>

## AGRADECIMIENTOS

Al Proyecto del Instituto de Salud Carlos III (ISCIII)/PI1900809/Cofinanciado FEDER Acuerdo de colaboración Illumina-Fundación Pública Galega de Medicina Xenómica (2023-2024) para el diagnóstico de los TEA.

A Illumina, Inc. por permitirme acceder al soporte tecnológico sin el cual este trabajo no hubiese sido posible.

A mis tutores, el Dr. Carracedo y el Dr. Rodríguez, así como a la Fundación Pública Galega de Medicina Xenómica por abrirme sus puertas y permitirme aprender de ellos y de su trabajo.



A mis padres, por su apoyo.