

UNIVERSIDAD DE COSTA RICA
SISTEMA DE ESTUDIOS DE POSGRADO

IDENTIFICACIÓN DE UN POSIBLE GEN CAUSANTE
DE DISCAPACIDAD COGNITIVA NO SINDRÓMICA
CON HERENCIA AUTOSÓMICA RECESIVA

Tesis sometida a la consideración de la Comisión del Programa de
Estudios de Posgrado en Biología para optar al grado y título de
Maestría Académica en Biología con énfasis en Genética y Biología
Molecular

TANYA LOBO PRADA

Ciudad Universitaria Rodrigo Facio, Costa Rica

2013

DEDICATORIA

Esta tesis está dedicada a madres ejemplares:

A las mamás de los pacientes. A esas mujeres valientes que a pesar de las múltiples dificultades y necesidades le dieron vida y amor a niños especiales. La fortaleza, dedicación y amor con las que estas mujeres han sacado adelante sus hogares las hacen dignas de admiración y reconocimiento.

A las mamás que más quiero. A mi mamá por darme la vida, apoyar mis sueños y heredarme la pasión por tratar de comprender la complejidad de la mente humana. A Bela, mi mejor amiga y mi apoyo incondicional. A mi hermana, mi apoyo en todos estos años y mi segunda mamá.

AGRADECIMIENTOS

A Dios por que sin El nada es posible.

A las familias que participaron en esta investigación por su tiempo, paciencia y cooperación.

A mi mamá y German, a mis hermanos Silvia y Thoren, y a Belo y Bela por su apoyo incondicional y económico para los viajes a Alemania.

A mi profesor, tutor, mentor y amigo, el Dr.rer.nat. Alejandro Leal Esquivel por apoyarme y guiarme en cada paso de esta investigación.

A Dr.rer.nat Gabriela Chavarría Soley y M.Psic. Marcela Barguil Gallardo por corregir y leer la tesis.

Al Dr.rer.nat Jaime Fornaguera por su disponibilidad, asesoramiento y guía siempre que se le solicitó.

Al DAAD (Deutscher Akademischer Austausch Dienst) y al Sistema de Estudios de Posgrado por el apoyo económico en ambos viajes a Alemania.

A la Lic. Marisol Ortega, directora del Centro del Enseñanza Especial de Puntarenas por acogernos en su institución y hacer posible contactar a las familias.

A las psicólogas M.Psic. Silvia Prada, Lic.Psic. Norma Curcó, Lic.Psic. María Valle, L.Psic. Nathalie Vargas por su colaboración con el diagnóstico de los pacientes.

A la Dra. Sandra Silva y la Dra. Ivannia Altmétlla por abrirme las puertas de su laboratorio colaborar con la extracción de ADN de una de las familias, así como, con algunas pruebas moleculares.

A Federico Hernández Villalobos y Geovanni Vargas Cháves por estar siempre dispuestos a acompañarnos en las giras y sangrar a los pacientes.

A Carolina Ortiz por su apoyo en la gira a Puntarenas y su ayuda con la síntesis de algunos resultados.

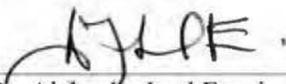
A Beatriz Willink por sus valiosas correcciones y enseñanzas a lo largo de mi carrera profesional.

A todo el personal Instituto de Genética Humana de Universitätsklinikum en Erlangen, Alemania. A todos los que me guiaron, asesoraron y ayudaron en esta investigación. Especialmente a Prof.Dr.med. Andre Reis por recibirme en el Instituto y permitirme llevar a cabo la investigación. Dr.rer.nat. Arif Ekici, BTA Angelika Diem y Dr.rer.nat. Steffen Uebe por su asistencia técnica y científica en el análisis de la secuenciación del exoma. A PD Dr. med. Rami Abou Jamra por siempre estar dispuesto a asesorarme especialmente con los análisis de ligamiento y mapeo por homocigosis. A Dipl. Biol. Petra Rothe y Dr.rer.nat. Arif Ekici por su guía y asesoramiento con los microarreglos de ADN. Dr.rer.nat Cornelia Kraus, PD Dr. med. Christiane Zweir, MTLA Andrea Eberwein, MTA Leonora Klassen, Dr. rer. nat. Mandy Krumbiegel por su asesoramiento con el diagnóstico de X-Frágil, MLPA, CNVs y exámenes de paternidad. A los estudiantes de doctorado Dipl. Biol. Diana Zahnleiter, Dipl. Mol. Med. Anne Gregor, Dipl. Mol. Med. Rebecca Buchert, Dipl. Mol. Med. Maria Apel por siempre estar dispuestas a ayudarme con mis dudas técnicas y científicas. A PhD.Dr.med. Georgia Vassiliou por hacer posible mi estadía en Alemania y siempre estar dispuesta a discutir sobre los pacientes conmigo. A MTLA Olga Zwenger por su ayuda técnica en el laboratorio.

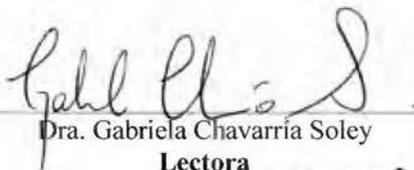
“Esta tesis fue aceptada por la Comisión del Programa de Estudios de Posgrado en Biología de la Universidad de Costa Rica, como requisito parcial para optar al grado y título de Maestría Académica en Biología con énfasis en Genética y Biología Molecular”



Dr. Gustavo Gutiérrez Espeleta
Representante del Decano
Sistema de Estudios de Posgrado



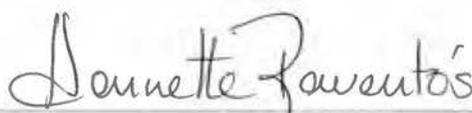
Dr. Alejandro Leal Esquivel
Director de Tesis



Dra. Gabriela Chavarría Soley
Lectora



MPsic. Marcela Banguil Gallardo
Lectora



Dra. Henriette Raventós Vorst
Representante Director
Programa de Posgrado en Biología



Tanya Lobo Prada
Candidata

INDICE

DEDICATORIA	ii
AGRADECIMIENTOS	iii
RESUMEN	vii
LISTA DE CUADROS	viii
LISTA DE FIGURAS	ix
LISTA DE ABREVIATURAS	xi
1. INTRODUCCION	1
1.1 Diagnóstico y prevalencia	1
1.2 Clasificación de la discapacidad cognitiva	1
1.3 Etiología	1
1.4 Genética y fisiopatología de la discapacidad cognitiva.....	1
2. JUSTIFICACIÓN	5
3. OBJETIVOS	6
3.1 Objetivo General.....	6
3.2 Objetivos específicos	6
4. MATERIALES Y MÉTODOS	6
4.1 Localización y escogencia de las familias.....	6
4.2 Aspectos éticos del proyecto	7
4.3 Diagnóstico clínico: caracterización del fenotipo de discapacidad cognitiva	7
4.4 Obtención de las muestras	8
4.5 Técnicas moleculares.....	8
4.5.1 Exámenes de paternidad.....	8
4.5.2 Descarte de síndromes asociados a discapacidad cognitiva	8
4.5.3 Microarreglos de ADN	11
4.5.4 Secuenciación del exoma	15
4.5.5 Secuenciación de Sanger.....	23
4.5.6 Modelaje proteico	24
5. RESULTADOS	25
5.1 Genealogías de las familias.....	25
5.2 Diagnóstico clínico	25

5.2.1	Familia PE.....	25
5.2.2	Familia ME	26
5.2.3	Familia PZ.....	27
5.2.4	Familia R	27
5.3	Técnicas moleculares.....	30
5.3.1	Exámenes de paternidad.....	30
5.3.2	Síndromes asociados a discapacidad cognitiva	30
5.3.3	Microarreglos de ADN	32
5.4	Cobertura del SOLiD4 con el IGV	39
5.5	Modelaje proteico	39
5.5.1	Dedos de zinc ZNF646 y ZNF629	39
5.5.2	GPT2.....	39
6.	DISCUSIÓN	41
6.1	Consideraciones del método	41
6.2	Genes candidatos	44
6.3	Implicaciones y direcciones futuras.....	46
7.	Conclusiones.....	47
8.	REFERENCIAS	49
9.	ANEXOS	55
9.1	Escala del desarrollo integral del niño (EDIN).....	55
9.2	Fórmula de consentimiento informado.....	57
9.3	Datos clínicos que conforman el expediente de los pacientes.....	60
9.4	Genealogía de la familia PZ utilizados para el análisis de ligamiento.....	61
9.5	Haplotipos de la familia PZ: para poder realizar el análisis de ligamiento se tuvo que partir la familia y aquí se presentan los haplotipos de las regiones que presentaron los puntajes LOD más altos.	62
9.6	Análisis de secuencias para los cambios genéticos en los genes SNTG2, RAB17 y ILKAP en todos los miembros de la familia ME.	68
9.7	Análisis de secuencias para los cambios genéticos , que no fueron validados, en los genes SMPD3 y SETD1A en los cuatro individuos afectados y TAOK2 y KIAA0554 en todos los miembros de la familia PZ.	69

RESUMEN

La discapacidad cognitiva no sindrómica (NS-ID) con herencia autosómica recesiva es una enfermedad heterogénea que afecta el sistema nervioso central e impide que los individuos afectados se valgan por sí mismos. Por medio de la secuenciación de nueva generación y regiones homocigotas idénticas por descendencia, es posible mapear genes causantes de NS-ID. Combinando estas técnicas moleculares en tres familias costarricenses con individuos afectados con discapacidad cognitiva severa se logró encontrar que una mutación no sinónima en el sitio activo del gen que codifica por la enzima GPT2 cosegrega con el fenotipo afectado en una familia. GPT2 es una transaminasa que cataliza una reacción irreversible en la que se produce glutamato. Es posible que un descenso en los niveles de glutamato impida la sincronización excitatoria de las neuronas durante procesos como el LTP (long term potentiation) involucrado en el aprendizaje y la memoria. Por lo tanto, GPT2 es un gen candidato de discapacidad cognitiva no sindrómica con herencia autosómica recesiva. Estos resultados contribuyen con el esclarecimiento de genes involucrados en el funcionamiento de la discapacidad cognitiva y fundamentan la búsqueda de alternativas terapéuticas para los individuos afectados.

LISTA DE CUADROS

Cuadro 1. Genes humanos mapeados que causan discapacidad cognitiva no sindrónica con herencia autosómica recesiva.	2
Cuadro 2. Criterios de inclusión utilizados para el reclutamiento de las familias candidatas para la investigación.	7
Cuadro 3. Síndromes asociados a discapacidad cognitiva detectados con el kit de MLPA P245-A2.	9
Cuadro 4. Microdeleciones reportadas de síndromes asociados a discapacidad cognitiva detectados con el kit de MLPA P297-B2.	9
Cuadro 5. Primers diseñados para las variantes de interés.	22
Cuadro 6. Programa de PCR utilizado para validar las mutaciones en genes candidatos.	23
Cuadro 7. Programa Exo_15min utilizado para la eliminación de exonucleasas en los productos de PCR.	23
Cuadro 8. Programa utilizado para la reacción de secuenciación.	24
Cuadro 11. Lista de mutaciones identificadas en la familia PZ.	39
Cuadro 12. Predicción realizada por los programas SIFT, PhyloP, PolyPhen y Mutation Taster para las mutaciones en genes candidatos de discapacidad idiopática no sindrónica en la familia PZ.	39

LISTA DE FIGURAS

- Figura 1.** Digestión del ADN genómico con las enzimas de restricción Nsp1 y Sty1 utilizado por el SNP 6.0. El Cytoscan HD utiliza exclusivamente la enzima Nsp1 para la digestión y, por lo tanto, consume la mitad del tiempo. (Imagen tomada de: http://www.asuragen.com/web2/Services/services/genotyping/affymetrix_SNP.aspx). 12
- Figura 2.** Marcaje utilizado por los microarreglos SNP 6.0 y CytoScan HD. Ambos utilizan un complejo de estreptavidina y ficoeritrina que reconocen la biotina unida a los extremos 3' de los fragmentos de ADN genómico. Esta imagen corresponde a microarreglos de ARN pero el proceso es el mismo para ADN. (Imagen tomada de: <http://www.ohsu.edu/xd/research/research-cores/gmsr/project-design/array-technology/affymetrix-genechip-arrays.cfm>). 13
- Figura 3.** Descripción general del enriquecimiento del exoma con el kit *SureSelect All Exon 50Mb*. Imagen tomada de: Agilent SureSelect Target Enrichment System Product Note. 17
- Figura 4.** Esquema del método de PCR por emulsión utilizado por el *SOLiD4*. Imagen tomada de: *ABi SOLiD 4 Templated Bead Preparation Guide*. 18
- Figura 5.** Esquema de la separación de las perlas o *beads* magnéticos unidos al ADN molde de ADN de interés luego de la centrifugación en un medio con glicerol al 60%. Imagen tomada de: *ABi SOLiD 4 Templated Bead Preparation Guide*. 19
- Figura 6.** Esquema de la reacción por ligación utilizada para la secuenciación del exoma por el *SOLiD4*. Imagen tomada de: http://www3.appliedbiosystems.com/AB_Home/applicationstechnologies/SOLiDSystemSequencing/OverviewofSOLiDSequencingChemistry/index.htm. 20
- Figura 7.** Esquema del método utilizado para filtrar las variantes detectadas por NGS. 22
- Figura 8.** Esquema general del proceso de limpieza de secuencias utilizando el kit AgencourtCleanSEQ® por medio del cual *beads* magnéticos separan los productos de secuenciación de los contaminantes. 24
- Figura 9.** Genealogía de las familias PE de Puntarenas. 28
- Figura 10.** Genealogía de la familia ME del Mojón de Esparza, Puntarenas. 29
- Figura 11.** Genealogía de la familia R del Roble, Puntarenas. 29
- Figura 12.** Genealogía de la familia PZ de Pérez Zeledón. 30
- Figura 13.** Electroferograma indicando las repeticiones CGG en el gen FMR1 características de una premutación del paciente R012. 31

Figura 14. Electroferograma indicando las repeticiones CGG en el gen FMR1 del individuo sano R015.	31
Figura 15. Puntajes LOD de todo el genoma luego de realizar el análisis de ligamiento utilizando un modelo de herencia recesivo en la familia ME.....	33
Figura 16. Puntajes LOD de todo el genoma luego de realizar el análisis de ligamiento utilizando un modelo de herencia recesivo en la familia PE.....	33
Figura 17. Puntajes LOD de todo el genoma luego de realizar el análisis de ligamiento con el modelo recesivo en la familia PZ (Pedigree D Anexo 7.4).	34
Figura 18. Vista del genome browser de UCSC donde se muestran las carreras de homocigosidad en chr2: 213824223-215284971 y, los genes dentro de ellas, compartidas por ambos afectados de la familia ME.	35
Figura 19. Vista del genome browser de UCSC donde se muestran las carreras de homocigosidad en chr12: 50322767-51626510 y, los genes dentro de ellas, compartidas por ambos afectados de la familia PE.	35
Figura 20. Vista del genome browser de UCSC donde se muestran las carreras de homocigosidad en chr16: 27157613-32483006 y, los genes dentro de ellas, compartidas por ambos afectados de la familia PZ.	37
Figura 21. Vista de la región candidata luego de realizar el mapeo por homocigosis en el genoma de los individuos de la familia PZ. La región oscila entre los marcadores rs2090296 y rs17821664 cuya posición genómica es chr16: 27177795-48202503.....	37
Figura 22. Validación de secuencias para los cambios genéticos en los genes <i>ZNF646</i> , <i>ZNF629</i> y <i>GPT2</i> en todos los miembros de la familia PZ (para <i>ZNF629</i> no se logró obtener la secuencia de la mamá).	38
Figura 23. Modelaje de la estructura proteica de GPT2. En verde, un análogo del sustrato, el inhibidor cicloserina (verde) y, en azul, la glicina 96 ubicada en el sitio activo de la enzima. Imagen producida por el de Dr.rer.nat Heinrich Sticht del Instituto de Bioquímica de la Universidad de Erlangen-Nuremberg.	40

LISTA DE ABREVIATURAS

ACCN2	Homo sapiens amiloride-sensitive cation channel 2
ADN	Ácido desoxirribonucleico
AFF2	Fragile X mental retardation 2
ANNOVAR	Annotation of genetic variants
ARN	Ácido ribonucleico
BAM	Bass alignment module
Ca	Calcio
CC2D1A	Coiled-coil C2 domain containing 1A
ChAS	Chromosome Analysis Suite
Chr	Chromosome
CI	Coeficiente Intelectual
CNV	Copy Number Variants
CRBN	Cereblon
db132	Data base 132
dbSNP	Data base of single nucleotide polymorfisms
DGV	Database of genomic variants
ePCR	Emulsion polymerase chain reaction
EDIN	Escala del desarrollo integral del niño
FASTQC	FAST quality control
FMR1	Fragile X mental retardation 1
FRAXA	Fragile X Polymorphism Assay
GATK	Genome Analysis Toolkit
GPT2	Glutamate pyruvate transaminase 2
GRIK2	Glutamate receptor ionotropic kainate 2
HapMap	Haplotype map project
hg19	Human genome 19
HpaII	Haemophilus parainfluenzae II
IBD	Identical by descent
ID	Intelligence Disability o Discapacidad Cognitiva
IGV	Integrative Genomics Viewers
IKFZ2	Ikaros family zinc finger 2
ILKAP	Integrin-linked kinase-associated serine/threonine phosphatase
InDels	Insertions deletions
K	Potasio
KAR	Kainate receptor
Kb	Kilobase
KIAA0556	Human uncharacterized protein 0556
KIAA1033	Human uncharacterized protein 1033
LOD	Logarithm of odds
LTP	Long term potentiation
MAN1B1	Mannosidase alpha class 1B member 1

Mb	Mega base
ME	Mojón de Esparza
MLPA	Multiplex ligand-dependent probe amplification
NCBI	National Center for Biotechnology Information
NfκB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NGS	Next generation sequencing
NHLBI	National Heart Lung Blood Institute
NIK	NF-kappa-B- inducing kinase
NRP2	Neuropilina 2
NS-ID	Nonsyndromic Intelligence Disability
OMIM	Online Mendelian Inheritance in Men
PCR	Polymerase chain reaction
PDB	Protein Data Bank
PE	Puntarenas
PolyPhen2	Polymorphism Phenotyping v2
POU6F1	Homo sapiens POU class 6 homeobox 1
PRSS12	Protease serine 12
PZ	Pérez Zeledón
R	Roble de Puntarenas
RAB17	Ras-related protein Rab-17
RefSeq	Reference Sequence
RM	Retardo Mental
ROH	Runs of homozygosity
S-ID	Syndromic Intelligence Disability
SAET	SOLiD Accuracy Enhancement Tool
SBK1	SH3-binding domain kinase 1
SETD1A	Set domain-containing protein 1A
SIFT	Sorting intolerant from tolerant software
SMPD3	Sphingomyelin Phosphodiesterase 3
SNC	Sistema nervioso central
SNP	Single nucleotide polymorphism
SNTG2	Syntrophin gamma 2
SOLiD	Sequencing by Oligonucleotide Ligation and Detection
SPAG16	Sperm associated antigen 16
TAOK2	TAO kinase 2
TAR	Trombocitopenia aplasia radial
TdT	Transferasa terminal de deoxinucleótidos
TECR	Trans-2,3-enoyl-CoA reductase
TRAPPC9	Trafficking protein particle complex 9
TUSC3	Tumor suppressor candidate 3
UCSC	University of California Santa Cruz
UTR	Untranslated region
WISC-IV	Weschler Intelligence Scale for Children-Fourth Edition

XIST	Homo sapiens X (inactive)-specific transcript
ZC3H14	Homo sapiens zinc finger CCCH-type containing 14
ZL3H14	Homo sapiens zinc finger CCCH-type containing 14
ZNF526	Zinc finger 526
ZNF629	Zinc finger 629

1. INTRODUCCION

1.1 Diagnóstico y prevalencia

La discapacidad cognitiva (ID por sus siglas en inglés) o retardo mental (RM) es una enfermedad que afecta el sistema nervioso central y se caracteriza por un reducido coeficiente intelectual (CI menor de 70), así como déficits en al menos dos comportamientos relacionados con funciones de adaptación (comunicación, cuidado personal, habilidad para vivir independiente, habilidades sociales o interpersonales, uso de servicios públicos, toma de decisiones, habilidades académicas, trabajo, salud y seguridad); y el inicio de la enfermedad debe ser antes de los 18 años (Asociación Americana de Psiquiatría 2000; Ropers & Hamel 2005). ID es la discapacidad más frecuente en niños y adultos jóvenes, está presente en alrededor de un 1-3% de la población y se encuentra en todas las clases sociales y culturas (Leonard & Wen 2002).

1.2 Clasificación de la discapacidad cognitiva

La discapacidad cognitiva puede ser clasificada en cinco categorías: leve (CI=50-70), moderada (CI=35-50), severa (CI=20-35) o profunda (CI<20) (Asociación Americana de Psiquiatría 2000). Sin embargo, en estudios epidemiológicos, la enfermedad se clasifica en leve cuando el CI=50-70 y severa (CI<50) (Ropers & Hamel 2005). Además, la discapacidad cognitiva puede clasificarse en sindrómico (S-ID), cuando los pacientes presentan otras manifestaciones clínicas asociadas al ID, o no-sindrómico (NS-ID), cuando la discapacidad cognitiva es la única característica clínica. El diagnóstico de la S-ID es bien definido, sin embargo, la clasificación de la discapacidad cognitiva no-sindrómica es más heterogénea y, por lo tanto, difícil de distinguir (Kaufman *et al.* 2010).

1.3 Etiología

La ID puede ser causada por factores genéticos y ambientales (Cuadro 1). Sin embargo, en un 80% de los casos, no es posible identificar la causa de la enfermedad (Rauch *et al.* 2006). Algunos factores ambientales como la exposición a teratógenos, virus, radiación, traumas en la cabeza o hipoxia cerebral, pueden causar la enfermedad en casos de NS-ID (Kaufman *et al.* 2010). Por otro lado, a partir de datos epidemiológicos, se estima que un 25-50% de los casos con ID podrían deberse a causas genéticas y, este porcentaje, aumenta proporcionalmente con la severidad (McLaren & Bryson 1987).

1.4 Genética y fisiopatología de la discapacidad cognitiva

La mayoría de los genes implicados en ID e identificados hasta el momento tienen una herencia ligada al cromosoma X (Vissers *et al.* 2010; Kaufman *et al.* 2010). Aproximadamente 40 genes se han encontrado

relacionados a NS-ID y 80% de estos se encuentran ubicados en el cromosoma X. Algunos de estos genes pueden causar S-ID o NS-ID dependiendo de la mutación y la interacción con otros factores genéticos. Sin embargo, se estima que la discapacidad cognitiva ligada al X corresponde, únicamente, a un 10% de los casos, por lo tanto, la causa subyacente en la gran mayoría de los casos debe ser autosómica (Ropers & Hamel 2005; Chelly *et al.* 2006; Philippe *et al.* 2009; Ropers 2010).

Un 25% de los pacientes con discapacidad cognitiva no sindrómica presentan un modo de herencia autosómico recesivo (Philippe *et al.* 2009). Sin embargo, los genes que causan NS-ID con herencia autosómica recesiva son difíciles de identificar dado que para ello se requieren familias grandes que son poco comunes pues los individuos con ID tienen una reproducción casi nula, y las familias con niños afectados en muchos casos dejan de reproducirse debido al intenso cuidado que demandan los hijos afectados (Vissers *et al.* 2010). Además, al tratarse de una enfermedad heterogénea—en la que muchos genes están causando el fenotipo—no existen criterios clínicos que permitan agrupar a las familias afectadas, según su etiología, para realizar mapeos genéticos apropiados—por ejemplo, a través de un análisis de ligamiento. Hasta la fecha se han reportado nueve genes involucrados en discapacidad cognitiva no sindrómica con herencia autosómica recesiva (*PRSS12*, *CRBN*, *CC2D1A*, *GRIK2*, *TUSC3*, *TRAPPC9*, *TECR*, *KIA1033*, *MAN1B1*) (Cuadro 1) Tan sólo tres de estos genes (*TUSC3*, *TRAPPC9* y *MAN1B1*) han sido identificados en más de una familia (Garshasbi *et al.* 2008, Mochida *et al.* 2009, Arshad-Rafiq *et al.* 2011). Además, recientemente, tres genes adicionales (*ZNF526*, *ST3GAL3* y *ZC3H14*) y 30 posibles loci se han reportado en congresos genéticos, pero todavía no se han publicado (Ropers 2010). Del mismo modo, combinando técnicas moleculares de alto rendimiento y cobertura, se han detectado 50 posibles genes nuevos involucrados en ID o síndromes asociados; de estos, 31 están relacionados con formas no sindrómicas. Sin embargo, se requieren estudios funcionales y más familias afectadas con mutaciones en esos genes, para comprobar la patogenicidad de las variantes (Najmabadi *et al.* 2011).

Cuadro 1. Genes humanos mapeados que causan discapacidad cognitiva no sindrómica con herencia autosómica recesiva.

Nombre del gen	Ubicación	OMIM	Tipo de mutación	Proteína	Función	Referencia
<i>PRSS12</i>	4q26	606709	Pérdida de sentido	Neurotripsina	Proteasa sináptica, plasticidad sináptica	Molinari <i>et al.</i> 2002
<i>CRBN</i>	3p26.3	607417	Pérdida de sentido	Cereblon	Expresión de los canales de potasio	Higgins <i>et al.</i> 2004. <i>Neurology.</i>
<i>CC2D1A</i>	19p13.12	608443	Pérdida de	Coiled-coil and C2	Regulador transcripcional de	Basel-Vanagaite <i>et al.</i>

			sentido	domain containing 1A	receptores de neurotransmisores, activador de la vía NF-kB	2005. <i>J Med Genet.</i>
<i>GRIK2</i>	6q21	611092	Pérdida de sentido	Receptor ionotrópico del glutamato, kainato 2	Subunidad del receptor sináptico del glutamato (KARs)	Motazacker <i>et al.</i> 2007. <i>Am J Hum Genet.</i>
<i>TUSC3</i>	8p22	611093	Pérdida de sentido	Supresor de tumores, candidato 3	Absorción de Mg ²⁺ , transferasa de oligosacáridos, glicosilación	Garshasbi <i>et al.</i> 2008. <i>Am J Hum Genet.</i>
<i>TRAPPC9</i>	8q24.3	611966	Pérdida de sentido	Proteína de unión a NIK y IKKB	Señalización neuronal NfκB, transporte vesicular.	Mochida <i>et al.</i> 2009. <i>Am J Hum Genet.</i>
<i>TECR</i>	19p13	610057	Cambio de sentido (no sinónima)	Glicoproteína sináptica	Síntesis de ácidos grasos importantes para mantener la homeostasis lipídica para el desarrollo adecuado del sistema nervioso	Caliskan <i>et al.</i> 2011. <i>Hum Mol Genet.</i>
<i>KIAA1033</i>	12q23-24	NM_015275	Cambio de sentido (no sinónima)	SWIP miembro del complejo WASH	Polimerización de actina y múltiples procesos involucrados en el transporte a los endosomas.	Ropers <i>et al.</i> 2011. <i>Hum Mol Genet.</i>
<i>MAN1B1</i>	9q34.3	604346	Cambio de sentido (no sinónima)	Manosil oligosacárido alpha 1,2-manosidasa	Maduración de péptidos (en la vía secretora) con glicanos unidos al extremo N-terminal. Desecho de glicoproteínas mal ensambladas a través de la degradación asociada al retículo endoplasmático	Arshad-Rafiq <i>et al.</i> 2011. <i>Am J Hum Genet.</i>

La mayoría de los genes mencionados anteriormente han sido identificados en familias grandes con alta consanguinidad, combinando técnicas moleculares como los microarreglos de ADN y el mapeo por homocigosis. Estas técnicas permiten detectar regiones de alelos homocigotas que comparten ambos

hermanos afectados. Actualmente, por medio de secuenciación de nueva generación (NGS) es posible secuenciar el exoma e identificar los genes que se expresen en el sistema nervioso central, entre los cuales posiblemente está el causante de la enfermedad. La combinación de estas técnicas es muy importante para el estudio de enfermedades raras como NS-ID con herencia autosómica recesiva (Kaufman *et al.* 2010). Además, la causa de muchas enfermedades recesivas, entre ellas el síndrome de Joubert, enfermedad de Charcot-Marie-Tooth, sordera no sindrómica y oligodontia, han sido exitosamente mapeadas con esta metodología (Bolino *et al.* 2000; Noor *et al.* 2008; Noor *et al.* 2009; Senderek *et al.* 2003; Verpy *et al.* 2001).

La gran cantidad de polimorfismos que se han detectado en el genoma han permitido detectar regiones extensas de autocigosidad en los individuos, conocidas como carreras de homocigosidad o ROH (runs of homozygosity, por sus siglas en inglés). Las ROHs son regiones que contienen genes homocigotas idénticos por descendencia (IBD-ambos alelos provienen de un ancestro común) como resultado de la reproducción de dos individuos emparentados. Sin embargo, también se han detectado ROHs comunes entre todos los individuos de una población (Nothnagel *et al.* 2010). Por lo tanto, es necesario identificar aquellos ROHs que compartan individuos afectados y que no sean comunes en controles, para descartar los ROHs no informativos. Con los ROHs y la información de secuenciación del exoma, se podrían identificar cuáles variantes raras están causando discapacidad cognitiva en una familia al reducirse significativamente la cantidad de información genética disponible para el análisis. Es por lo anterior, que esta nueva metodología (combinación de NGS-ROH) está pensada para familias nucleares con uno o más afectados y cuando se presupone que podrían haber alelos IBD—debido a una probable unión consanguínea algunas generaciones atrás.

La ID también está frecuentemente asociada con la presencia de variaciones en el número de copias (CNV por sus siglas en inglés) causadas por duplicaciones y deleciones en el genoma (Bolino *et al.* 2000; Noor *et al.* 2008; Senderek *et al.* 2003; Verpy *et al.* 2001). Sin embargo, los CNVs también se encuentran en individuos sanos en la población. Por lo tanto, es importante distinguir entre CNVs “benignos” de aquellos causantes de enfermedades (Webber *et al.* 2009). Una diferencia entre CNVs patogénicos con respecto a los benignos es que los benignos son regiones genómicas ricas en secuencias repetitivas mientras que los CNVs asociados a ID tienden a poseer un alto contenido de genes que conllevan a anormalidades específicas del sistema nervioso central (Hehir-Kwa *et al.* 2010). De este modo, se han podido asociar CNVs específicos a pacientes con discapacidad cognitiva, por medio de técnicas como los microarreglos de ADN (Webber *et al.* 2009). Por lo tanto, es importante descartar primero que la enfermedad está siendo causada por CNVs patogénicos.

Por otro lado, la técnica de citogenética molecular, MLPA (multiplex ligand-dependent probe amplification), ha sido utilizada para detectar rearrreglos teloméricos submicroscópicos en alrededor de un 5% de pacientes con ID (Flint *et al.* 1995; Knight *et al.* 1999; Koolen *et al.* 2004; De Vries *et al.* 2003). Estos resultados indican que los rearrreglos submicroscópicos en las regiones subteloméricas podrían ser más comunes en pacientes con ID, que rearrreglos cromosómicos visibles con técnicas de citogenética convencional (Bert de Vries *et al.* 2005). Por lo tanto, del mismo modo que con los CNVs, se puede utilizar MLPA para descartar que el fenotipo del individuo afectado se deba a un rearrreglo submicroscópico ya reportado en la literatura.

En el presente estudio se escogieron familias costarricenses con dos o más individuos afectados con discapacidad cognitiva no sindrómica, cuyos padres están sanos, para mantener la hipótesis de que la enfermedad se transmite de manera recesiva como la más plausible. Luego, se utilizaron las novedosas técnicas anteriormente mencionadas para tratar de identificar el o los genes candidatos de causar la discapacidad cognitiva.

2. JUSTIFICACIÓN

A un 80% de los pacientes con discapacidad cognitiva se les desconoce la etiología de la enfermedad (Webber *et al.* 2009). Se estima que existen causas genéticas en un 25-50% o más de los casos (McLaren & Bryson 1987). La mayoría de estudios se han enfocado en la búsqueda de mutaciones en genes con herencia ligada al cromosoma X (10-12% de hombres con ID) aunque se calcula que deben existir al menos 800-850 genes asociados a ID con herencia autosómica. Incluso es muy posible que existan miles de genes con herencia autosómica recesiva asociados a ID (Ropers 2010). Por lo tanto, usando la metodología propuesta en este estudio, que combina la secuenciación de nueva generación y las carreras de homocigosidad, podrían identificarse genes candidatos de ID con herencia autosómica recesiva.

Es necesario comprender mejor las causas genéticas de la NS-ID para un mejor entendimiento del papel del sistema nervioso central en procesos cognitivos. Los genes que participen en la patogénesis de la enfermedad tendrán un papel importante en los procesos involucrados en el neurodesarrollo o en el funcionamiento normal del sistema nervioso central. Aportes al entendimiento de las vías o mecanismos involucrados en el desarrollo de la cognición permitirán poner las bases para encontrar posibles tratamientos o para el alivio de síntomas asociados a la discapacidad cognitiva. Además, conocer las alteraciones en genes involucrados en estas vías permitirá construir el camino hacia un diagnóstico molecular e incluso la prevención del desarrollo de esta enfermedad (Kaufman *et al.* 2010; Ropers 2010).

Además,, comprender la genética de enfermedades complejas como la NS-ID es relevante en el asesoramiento genético que se le puede brindar a estas familias con varios individuos afectados especialmente cuando existe alta probabilidad de que los padres estén emparentados (Kaufman *et al.* 2010).

Finalmente, en Costa Rica es posible rastrear los matrimonios consaguíneos debido a que se cuenta con una extensa documentación de registros civiles y eclesiásticos, lo que permite reconstruir genealogías que son muy importantes para el estudio de enfermedades recesivas (Meléndez, 2004). Algunas enfermedades que se han estudiado en Costa Rica por estas características son Charcot Marie Tooth (Leal *et al.* 2001), ataxia telangiectasia (Telatar *et al.* 1998), distrofia muscular de Duchenne (Chavarría *et al.* 2002) y glaucoma primario congénito (Chavarría *et al.* 2004). Este hecho justifica el uso de la metodología propuesta para estudiar el origen genético de la NS-ID en familias costarricenses.

3. OBJETIVOS

3.1 Objetivo General

Determinar si hay mutaciones causantes de ID con herencia autosómica recesiva en algunas familias de la población costarricense.

3.2 Objetivos específicos

1. Identificar tres familias con al menos dos hijos afectados con ID, cuyos padres estén emparentados y sean sanos.
2. Caracterizar el fenotipo de ID en los afectados.
3. Determinar si los afectados tienen mutaciones en estado homocigoto, que no se encuentren en ese estado en los padres y otros parientes sanos, y que estén presentes en proteínas que se expresen en sistema nervioso central.
4. Definir los genes candidatos como causantes de ID en las familias.

4. MATERIALES Y MÉTODOS

4.1 Localización y escogencia de las familias

Se localizaron cuatro familias afectadas con NS-ID, por medio de bases de datos de Centro de Enseñanza Especial y una a través del Club de Leones de Costa Rica. Luego de una búsqueda exhaustiva en varios centros de enseñanza especial de Costa Rica, se escogió aquel en el que se encontraron más familias candidatas. Una vez localizadas las familias, se les solicitó la historia

médica de los niños afectados para descartar que la discapacidad cognitiva se debiera a factores ambientales. Luego, se entrevistó a los padres para reconstruir el árbol genealógico, y se determinó el modo de herencia de la ID. Mediante registros civiles y eclesiales se amplió el estudio genealógico de las familias escogidas (Fig. 1) y se estableció si los padres estaban emparentados. Seguidamente, se escogieron aquellas familias que tuvieran más de un hijo afectado con discapacidad cognitiva idiopática, que los padres estuvieran relacionados genealógicamente, aunque la relación fuera distante, y cuyos padres fueran sanos (para considerar la hipótesis de herencia autosómica recesiva como la más plausible). Ver en Cuadro 2 los criterios de inclusión que fueron utilizados para el reclutamiento de los pacientes.

Cuadro 2. Criterios de inclusión utilizados para el reclutamiento de las familias candidatas para la investigación.

-
- Familias con al menos dos miembros afectados con discapacidad cognitiva no sindrómica
 - Ambos padres emparentados y sanos
 - Detección de síntomas antes de 18 años (expresión temprana)
 - Cualquier grado de discapacidad cognitiva no sindrómica
 - Ausencia de otras patologías que puedan explicar la discapacidad cognitiva, incluyendo problemas durante el embarazo y período perinatal
 - Homogeneidad clínica
 - Exclusión de síndromes asociados a discapacidad cognitiva (por ejemplo: X-frágil)
-

4.2 Aspectos éticos del proyecto

Este proyecto fue presentado y aprobado por el Comité Ético Científico y la Vicerrectoría de Investigación de la Universidad de Costa Rica. Ver fórmula de consentimiento informado en Anexos 9.1. Se leyó y explicó el consentimiento informado a los pacientes para que conocieran los objetivos, los beneficios y los riesgos del proyecto. Los sujetos participantes mayores de edad autorizaron la extracción del ADN de una muestra de sangre periférica, así como, la utilización de estas en el extranjero por parte de otros investigadores. Del mismo modo, los padres de los menores de edad, autorizaron la participación de éstos en el estudio. Toda la información y las muestras se codificaron para proteger la confidencialidad de los sujetos.

4.3 Diagnóstico clínico: caracterización del fenotipo de discapacidad cognitiva

Previo consentimiento informado, las familias participantes fueron visitadas por una psicóloga que aplicó la Escala Wechsler de Inteligencia para Niños, Cuarta Edición (WISC-IV) para los menores de 16 años (Wechsler 2003). Para los individuos mayores de 16 años se utilizó la Escala de

Wechsler de Inteligencia para Adultos, Tercera Edición (WAIS-III) (Wechsler 1997). Para calificar ambas pruebas se utilizaron las normas mexicanas, ya que este instrumento no ha sido validado ni estandarizado en Costa Rica. Se utilizó este instrumento para obtener de forma cuantitativa el coeficiente intelectual (CI), de cada paciente. Además, se pasó una escala de habilidades adaptativas para detectar si además de un CI bajo, los pacientes tenían dificultades en al menos dos habilidades adaptativas. Además, se pasó la Escala del Desarrollo Integral del Niño (EDIN) para aquellos individuos afectados menores de seis años. La EDIN evalúa los procesos cognoscitivos en niños de 2 a 6 años, así como, la motora fina y gruesa, el lenguaje, las áreas socioafectiva y cognoscitiva y los hábitos de salud y nutrición (Rodríguez Aragonés, 1987) (Anexo 9.2). Del mismo modo, se realizó un cuestionario con algunos datos clínicos para descartar causantes ambientales de discapacidad cognitiva (Anexo 9.3).

4.4 Obtención de las muestras

Se les solicitó a todos los familiares en primer grado donar una muestra de 5 mL sangre periférica del antebrazo. Este procedimiento fue realizado por un flebotomista del Hospital Monseñor Sanabria en Puntarenas o de la Facultad de Microbiología de la Universidad de Costa Rica. La sangre se utilizó únicamente para extraer el ADN.

La extracción del ADN se realizó por medio del método de extracción de fenol-cloroformo. El ADN se cuantificó con un espectrofotómetro, el NanoDrop 2000c, y se almacenó a 4°C.

4.5 Técnicas moleculares

4.5.1 Exámenes de paternidad

Se realizaron exámenes de paternidad a las tres familias por medio del PowerPlex16® de promega (Krenke et al. 2002). Este sistema utiliza 16 marcadores y una amplificación por PCR. Los resultados se obtuvieron del analizador genético ABI PRISM 3100®. Del mismo modo, para verificar los resultados—en aquellas familias con hijos hombres que no correspondieran al padre—se utilizó el Sistema PowerPlexY de promega que utiliza 12 marcadores en el cromosoma Y.

4.5.2 Descarte de síndromes asociados a discapacidad cognitiva

4.5.2.1 MLPA

Se utilizó la técnica de citogenética molecular MLPA (multiplex ligation-dependent probe amplification) de MRC-Holland® para descartar que la causa de la enfermedad se debiera a

microdeleciones, duplicaciones o translocaciones submicroscópicas asociadas a defectos cognitivos. Se utilizó el SALSA MLPA Kit P245-A2 Microdeletion Syndromes-1 para descartar simultáneamente múltiples síndromes comunes asociados a ID (Cuadro 3). También se utilizó el kit SALSA MLPA KIT P297-B2 para detectar otros síndromes asociados a ID que se deben a microdeleciones reportadas recientemente (Cuadro 4). Para detectar microdeleciones y duplicaciones de varios genes en regiones subteloméricas se utilizó el SALSA MLPA Kit P036-E1 Human Telomere-3, así como, el kit SALSA MLPA Kit P070-B2 Human Telomere-5, fabricados también por MRC-Holland®. La secuenciación de los fragmentos se obtuvo por medio del analizador genético ABI-3100 de Applied Biosystems y el análisis se realizó por medio del software Sequence Pilot®.

Cuadro 3. Síndromes asociados a discapacidad cognitiva detectados con el kit de MLPA P245-A2.

Microdeleciones		
síndrome deleción 1p36	Síndrome de Cri du Chat 5p15	Duplicación MECP2/Xq28
microdeleción 2p16	Síndrome de DiGeorge 22q11	Síndrome Rubinstein-Taybi
microdeleción 3q29	DiGeorge región 2/ 10p15	Síndrome Smith-Magenis
microdeleción 9p22.3	Síndrome de Langer-Giedion 8q	Síndrome Sotos 5q35.3
síndrome deleción 15q24	Síndrome de Miller-Dieker 17p	Síndrome Wagr
microdeleción 17q21	Síndrome microdeleción NF1	Síndrome Williams
Phelan-McDermid 22q13	Prader-Willi/Angelman	Wolf-Hirschhorn 4p16.3

Cuadro 4. Microdeleciones reportadas de síndromes asociados a discapacidad cognitiva detectados con el kit de MLPA P297-B2.

Microdeleción
1q21.1 síndrome de TAR
1q21.1 (diferente a TAR)
Síndrome deleción 3q29
7q36.1 Esquizofrenia/autismo
12p11.23 esquizofrenia
Síndrome microdeleción 15q13

Síndrome microdelección 15q24.1

Síndrome microdelección 16p11

Síndrome microdelección 17q12

18q21.2 Síndrome de Pitt-Hopkins

20p12.2 Esquizofrenia

4.5.2.2 Diagnóstico de X-Frágil

Se utilizó el kit SALSA MS-MLPA ME029-B1 FMR1/AFF2 para detectar si los pacientes tenían patrones de metilación anormales en los promotores de los genes FMR1 y AFF2 causantes del síndrome de X-frágil. Además, se realizó un análisis de fragmentos, utilizando el kit Fragile X Polymorphism Assay (FRAXA)[®] de Applied Biosystems para detectar el número de repeticiones CGG en el gen FMR1. La secuenciación de los fragmentos se obtuvo por medio del analizador genético ABI PRISM-3100[®] de Applied Biosystems y estos se analizaron por medio del software Sequence Pilot[®]. Debido a que el kit FRAXA no es recomendado para detectar mutaciones completas, aquellos pacientes que presentaron más de 61 repeticiones les realizamos un Southern Blot para detectar si eran portadores de la premutación o de la mutación completa. Para realizar el Southern Blot, el ADN fue cortado con dos enzimas de restricción (HindIII y EcoR1) y se incubó con una sonda marcada con P³² por una semana.

4.5.2.3 Análisis de inactivación del cromosoma X

Para aquellas familias, que en el momento del diagnóstico molecular, no se había encontrado un ancestro en común—y, por lo tanto, no se pudo establecer, con confianza, la hipótesis de herencia autosómica recesiva como la más plausible—se les realizó un análisis de inactivación no aleatoria del cromosoma X. La inactivación no aleatoria del cromosoma X es común en mujeres portadoras de rearrreglos cromosómicos, enfermedades ligadas al cromosoma X y mutaciones en el gen XIST (Lau *et al.* 1997). Debido a que existe una correlación entre la metilación del gen AR (receptor de andrógenos) y la inactivación del cromosoma X, se utilizó la enzima de restricción *HpaII*—que corta fragmentos de ADN sin metilaciones—para detectar los patrones de inactivación del cromosoma X de las madres (Lau *et al.* 1997). Se tomó como un patrón de inactivación del cromosoma X no aleatorio cuando la razón de cambio entre el área de los picos del electroferograma—que equivale al tamaño de los fragmentos—en presencia y ausencia de *HpaII* fue mayor a 90%. Los fragmentos se analizaron por medio del analizador genético ABI PRISM[®]-3100.

4.5.3 Microarreglos de ADN

Los microarreglos de ADN son superficies sólidas—la utilizada en esta investigación fue de vidrio—a las que se unen miles de oligonucleótidos de ADN, previamente digeridos con enzimas de restricción (Nsp1 y Sty1), que al hibridarse con sondas marcadas emiten fluorescencia en distintas intensidades y en lugares específicos, al mismo tiempo que son capturadas por una cámara. El microarreglo Genome-Wide Human SNP Array 6.0[®] de Affymetrics contiene más de 1.8 millones de marcadores genéticos, entre ellos 906.600 SNPs (tag SNPs, SNPs de cromosomas X y Y, SNPs mitocondriales, nuevos SNPs añadidos al dbSNP, SNPs en sitios calientes de recombinación), 945.826 sondas para CNV (202.000 sondas en el genoma que identifican 5677 CNV reportados en la base de datos de variaciones genómicas de la Universidad de Toronto (DGV), 744.000 sondas esparcidas a lo largo de todo el genoma); con un promedio de call rate de más de un 99% y la concordancia con HapMap excede 99.7%. Por otro lado, el nuevo Cytoscan HD Array[®], también de Affymetrics, contiene 2.6 millones de marcadores genéticos, entre ellos, 750.000 SNPs y es capaz de detectar con un 99% de confiabilidad cambios en el número de copias de 25-50kb. El promedio de la densidad de marcadores para el Cytoscan HD Array[®] es de 1/659pb en genes de OMIM[®], 1/880pb en genes de RefSeq[®] y 1/1737pb en regiones intergénicas (Consultado en: <http://www.affymetrix.com>).

En esta investigación, se realizaron microarreglos de ADN para todos los miembros (n=23) de las familias afectadas (ME, PE y PZ)—a excepción del individuo PZ012 que se excluyó del estudio por ser un caso de no paternidad (Fig. 9). Para las familias ME y PE se utilizó el Genome-Wide Human SNP Array 6.0[®] de Affymetrics y para la familia PZ se utilizó la versión más nueva de Affymetrics, el Cytoscan HD Array[®].

4.5.3.1 Preparación de las muestras

Se tomaron 250 ng de ADN de todos los individuos de las familias y se prosiguió a realizar la digestión con las enzimas Nsp1 y Sty1 para el SNP 6.0 (Fig. 1) y exclusivamente con Nsp1 para el Cytoscan HD. Luego, se ligaron adaptadores a las puntas pegajosas con ligasa de ADN T4. La amplificación de los productos se realizó con primers específicos para los adaptadores y la polimerasa Titanium Taq[®] de Promega. Para verificar que la amplificación del ADN se corrió un gel de agarosa al 2%, se prosiguió a limpiar los productos de PCR con perlas magnéticas para luego cuantificarlas con el espectrofotométricamente. Los demás pasos se ajustaron al protocolo original del fabricante. Los productos de PCR se fragmentaron con ADNasa1 para obtener tamaños de no

más de 125pb; el tamaño de los productos se verificó en un gel de agarosa al 4%. Los fragmentos, en los extremos 3', se biotinizaron con la transferasa terminal de deoxinucleótidos (TdT).

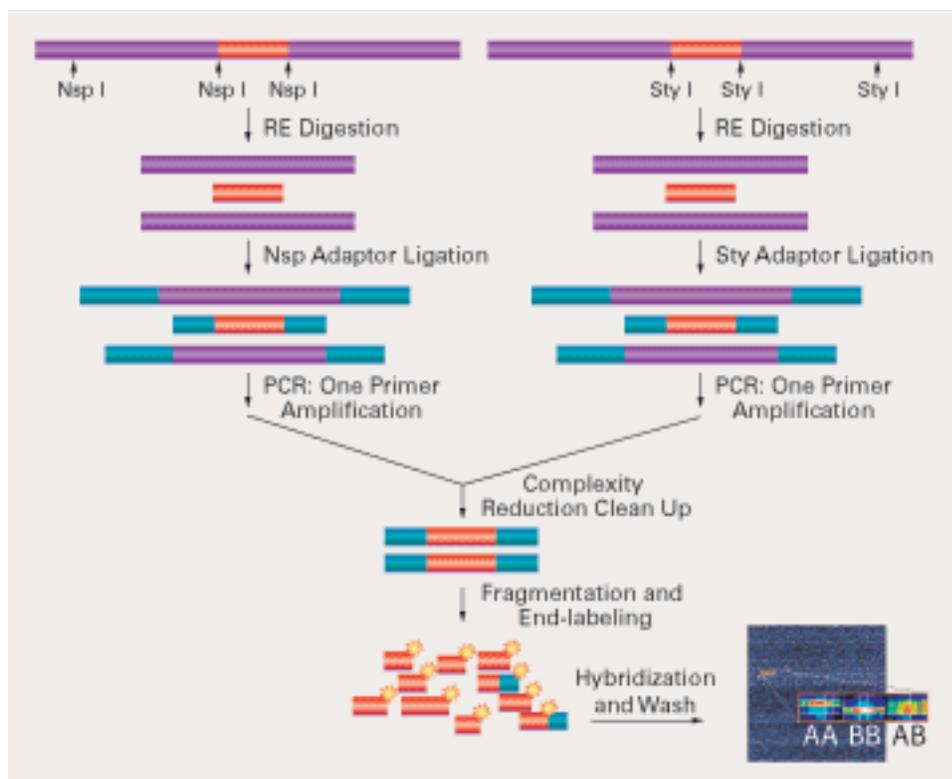


Figura 1. Digestión del ADN genómico con las enzimas de restricción Nsp1 y Sty1 utilizado por el SNP 6.0. El Cytoscan HD utiliza exclusivamente la enzima Nsp1 para la digestión y, por lo tanto, consume la mitad del tiempo. (Imagen tomada de: http://www.asuragen.com/web2/Services/services/genotyping/affymetrix_SNP.aspx).

4.5.3.2 Hibridación, lavado, marcaje y escaneo

Durante la hibridación los fragmentos de ADN se unen a sondas específicas, es decir, oligonucleótidos complementarios a las secuencias de los fragmentos. Las uniones no específicas son inhibidas por la solución tampón o buffer de hibridación. Los fragmentos no unidos luego se lavan y remueven por medio de la GeneChip® Fluidics Station 450 con diferentes niveles de astringencia. Después, el microarreglo es lavado con ficoeritrina unida a estreptavidina (SAPE) que reconocen y se unen a la biotina—previamente marcados con anticuerpos (Fig. 2). La ficoeritrina emite fluorescencia cuando se excita a 532nm y emite una longitud de onda de 575nm; esto es escaneado y guardado en formato .DAT por el GeneChip® Scanner 3000 7G de Affymetrics.

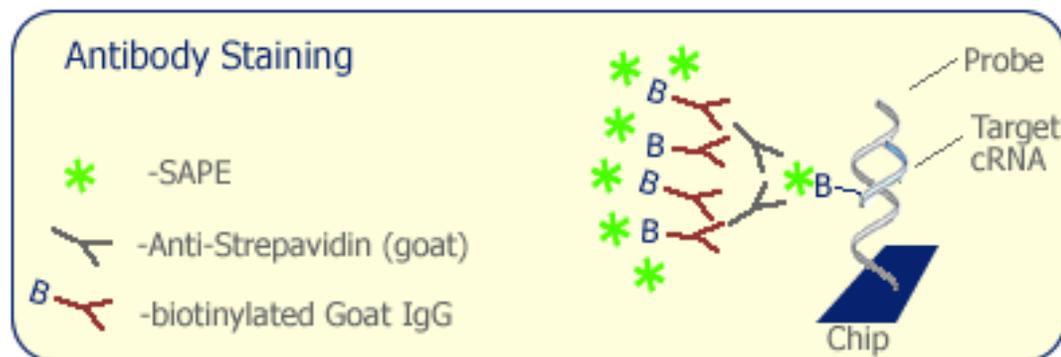


Figura 2. Marcaje utilizado por los microarreglos SNP 6.0 y CytoScan HD. Ambos utilizan un complejo de estreptavidina y ficoeritrina que reconocen la biotina unida a los extremos 3' de los fragmentos de ADN genómico. Esta imagen corresponde a microarreglos de ARN pero el proceso es el mismo para ADN. (Imagen tomada de: <http://www.ohsu.edu/xd/research/research-cores/gmsr/project-design/array-technology/affymetrix-genechip-arrays.cfm>).

4.5.3.3 Análisis de datos

Posteriormente, para el procesamiento de datos se utilizó el Affymetrix Chromosome Analysis Suite 1.2 (ChAS). Los parámetros utilizados para definir los niveles de intensidad fueron calculados utilizando el control interno de 224 microarreglos de pacientes del Instituto de Genética Humana de la Universidad de Erlangen-Nuremberg. De esta forma, se pueden eliminar, de mejor manera, los artefactos específicos de cada usuario—en este caso, la estación de trabajo del Instituto de Erlangen—normalizando los datos y mejorando la calidad de los resultados. Los CNVs, las regiones LOH (Loss of Heterozygosity) y los genotipos de los SNPs se almacenan en archivos *.CYCHP. Finalmente, los resultados se visualizan con el software ChAs.

4.5.3.4 Variantes en el número de copias

A partir del genotipo realizado con los microarreglos de ADN, se obtuvieron las variantes en el número de copias. Las variantes en el número de copias (CNV) son alteraciones estructurales en el genoma que producen duplicaciones o deleciones de segmentos de ADN. Los CNVs pueden ser heredados o provocados por mutaciones “de novo”. Debido a que se ha encontrado que CNVs raros son la causa de discapacidad cognitiva en algunas familias (Hehir-Kwa *et al.* 2010), se decidió descartar que en las familias PE, ME y PZ, la enfermedad se debiera a CNVs patogénicos y, por lo tanto, se realizó un tamizaje de CNVs en todos los cromosomas. Los CNVs (duplicaciones o deleciones) se consideraron como patogénicos cuando estaban presentes en los individuos afectados y no así en familiares ni en la base de datos de polimorfismos comunes. Además, se definieron

como CNVs aquellos segmentos que tuvieran una densidad de al menos 20 marcadores presentes en 100 kilobases.

4.5.3.5 Carreras de homocigosidad (ROHs) y mapeo por homocigosis

Se realizó un genotipo de los SNPs obtenidos a partir de los microarreglos de ADN para determinar los fragmentos genómicos que presentaban homocigosidad (runs of homozygosity, ROH) (Nothnagel et al. 2010). Para definir los ROHs se utilizó la herramienta bioinformática PLINK v.1.07. (Purcell *et al.* 2007, <http://pngu.mgh.harvard.edu/purcell/plink/>). Los ROHs se definieron como aquellas regiones genómicas en las que habían 100 marcadores en homocigosis en al menos 1000 kb (1 Mb) y con una densidad mínima de 1 SNP por cada 50 kb. Después, se buscó en todos los cromosomas, ROHs que compartieran los individuos afectados mas no así los parientes sanos. Aquellos ROHs compartidos entre los individuos afectados se consideraron como regiones de interés para la búsqueda de genes candidatos. También se utilizó el programa HomozygosityMapper® para realizar el mapeo por homocigosis para obtener las coordenadas de los marcadores que delimitan la región en homocigosis en los individuos afectados.

4.5.3.6 Análisis de ligamiento

Se tomaron 10.000 SNPs de los microarreglos de ADN y se utilizó el software EasyLINKAGE-Plus v.5® (Hoffman & Lindner 2005) para realizar el análisis de ligamiento y verificar la existencia posibles *loci* candidatos. Se tomó en cuenta como locus de interés aquel cuyo puntaje LOD fueran mayores a 2. Al observar la segregación de la enfermedad en las familias, se decidió realizar análisis de ligamiento bajo la suposición de un modelo autosómico recesivo, codominante, penetrancia del 100%, frecuencia de la enfermedad de 0.0001, con un tasa de asignación de 95% y frecuencias alélicas distribuidas uniformemente en todo el genoma. Sin embargo, para la familia ME—en la cual no se logró establecer un ancestro común entre padre y madre—se decidió correr el análisis suponiendo un modelo dominante y uno recesivo. Para la familia PE también se decidió correr ambos modelos (dominante y recesivo) a pesar de que se logró unir la genealogía. No se justificó realizar un modelo con todas las familias juntas porque no habían evidencias clínicas ni en el comportamiento similares. Debido a limitaciones en la capacidad de almacenaje de las computadoras utilizadas, la familia PZ (n=13) se dividió en cinco análisis diferentes para poder correr el programa. En todos los análisis se tomaron en cuenta los cuatro afectados pero varió la combinación de hermanos sanos incluidos (Anexo 9.4).

Los haplotipos se calcularon con el software Allegro® (Gudbjartsson et al. 2000) y, luego, se visualizaron con la herramienta HaploPainter® (Thiel y Nürnberg 2007).

4.5.4 Secuenciación del exoma

Se utilizó el *Solid4* de Applied Biosystems para secuenciar el exoma—la región del genoma que codifica para proteínas—así como, algunos miARN y pequeñas regiones intrónicas (NGS) (Montenegro et al. 2011). Se realizó la secuenciación para un único paciente afectado en las familias ME y PE, y dos afectados de la familia PZ. Luego se buscaron mutaciones no sinónimas, con pérdida o ganancia de función, así como, inserciones o deleciones que fueran homocigotas y se ubicaran dentro de los ROH que compartían únicamente los pacientes. Se descartaron polimorfismos comunes que se encontraran reportados en la base de datos dbSNP132, así como, en 1000 genome project (febrero del 2012), 5400 Exome sequencing project (NHLBI) y en 230 controles (133 se secuenciaron fragmentos y 97 ambas hebras del ADN) de pacientes secuenciados en el Instituto de Genética Humana de la Universidad de Erlangen-Nuremberg en Alemania. Se dio especial atención a variaciones en genes que se expresan en el sistema nervioso central. Sin embargo, cualquier variación que cumpliera los requisitos mencionados anteriormente fue considerada como una posible candidata. Además, se utilizaron las herramientas bioinformáticas para predecir patogenicidad como PolyPhen2 (Adzhubei et al. 2010) y SIFT (Ng y Henikoff 2003) que predicen si el cambio en un aminoácido altera la estructura o función de la proteína. Se utilizó la herramienta PhyloP (Spiel et al. 2006) para medir el grado de conservación de SNPs en una posición determinada en diferentes especies. *Mutation Taster* (Schwarz et al. 2010)—que predice si un cambio genético podría ser patogénico. Finalmente, por medio del software IGV v.2.2® (*Integrative Genomics Viewers*) se verificó la cobertura de las lecturas del secuenciador en cada uno de los genes del ROH candidato. Esto para asegurarse que no se dejara por fuera ningún exón de ningún gen candidato. Se tomaron como cubiertos aquellos exones que tuvieran al menos una lectura que abarcara todo el exón.

4.5.4.1 Preparación de la biblioteca de fragmentos de ADN

Se utilizó el ultrasonificador Covaris—que utiliza energía acústica ultrasónica—para cortar el ADN en fragmentos de 150-180pb. Para verificar el tamaño, la calidad y la cantidad de los fragmentos de ADN se utilizó el kit de alta sensibilidad de Agilent 2100 Bioanalyzer®. Luego, se utilizó el sistema automatizado SPRIworks system III® de Beckmann Coulter para la creación de la biblioteca de ADN con tamaños y limpieza específicos. En el proceso de creación de la biblioteca

se ligaron adaptadores en los extremos del ADN fragmentado, también, por medio del SPRIworks system III[®] de Beckmann Coulter. Luego, se obtuvo ADN simple banda por medio del *Nick translation* y se realizó la amplificación de la biblioteca. 450µl de los productos de PCR de la biblioteca se limpiaron por medio de *beads* magnéticos de AMPureXP[®] y 1mL de etanol al 70%. Luego se cuantificó la pureza por medio del kit 1000 para ADN que utiliza la plataforma Agilent 2100 Bioanalyzer[®].

4.5.4.2 Bead Enrichment con el kit SureSelect All Exon 50Mb Kit

Para capturar exclusivamente todos los exones, se utilizó el sistema conocido como Agilent SureSelect Target Enrichment System[®] (Porreca *et al.* 2007), por medio de sondas de ARN biotinilado. El kit utilizado fue el Agilent SureSelect[®] Human All Exon 50Mb que utiliza 635,000 sondas diferentes, de manera que captura una región de 51.6 Mpb, es decir, 300.040 exones de 23.031 transcritos de la secuencia codificante (CCDS) y 669 miARN (Sulonen *et al.* 2011).

Las sondas de ARN y la biblioteca de ADN preparada previamente se suspendieron en un buffer de hibridación por 24 horas a 65°C. Luego se tomaron 5µL de la biblioteca *enriquecida* y se mezclaron con 2 µL de ARNasa. Luego, se utilizaron *My One Streptavidin T1 Magnetic Beads* o perlas magnéticas cubiertas con estreptavidina—de Invitrogen—que reconocen y se adhieren a las sondas de ARN. La fracción de ADN al que no se unieron las perlas magnéticas fue descartada (Fig. 3). Al ADN de la biblioteca “madre” unida a los *beads* se le añadieron códigos de barras con el fin de secuenciar diferentes ADNs en una celda de flujo al mismo tiempo. Posteriormente, el ADN se amplificó utilizando 4µl de SureSelect Barcode Primers. 120µl del producto de PCR se purificó con 216µl de AMPure XP Beads. La pureza de las muestras se cuantificó por medio del kit de alta sensibilidad de Agilent 2100 Bioanalyzer[®].

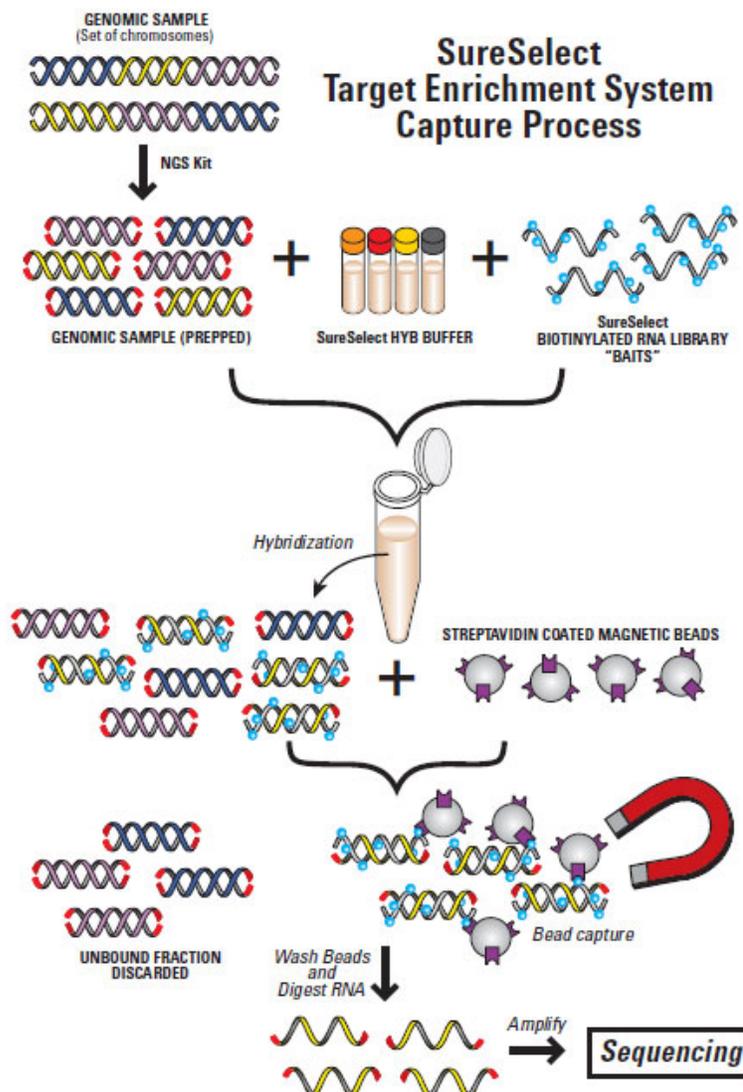


Figura 3. Descripción general del enriquecimiento del exoma con el kit *SureSelect All Exon 50Mb*. Imagen tomada de: Agilent SureSelect Target Enrichment System Product Note.

4.5.4.3 PCR por emulsion/Enrichement de las hebras molde

La secuenciación de nueva generación requiere de un método de PCR que amplifique masivamente la biblioteca de ADN de interés. El SOLiD® utiliza el PCR por emulsión (ePCR) o SOLiD EZ Bead Emulsifier®. Con este método se generan poblaciones de perlas magnéticas en una emulsión de agua en aceite. La emulsión la contienen microreactores en los que además se encuentran perlas magnéticas, ADN molde, componentes del PCR y primers (Fig.4). La

concentración de las perlas magnéticas y del ADN se escogen de manera tal, que se favorezca la formación de productos monoclonales.

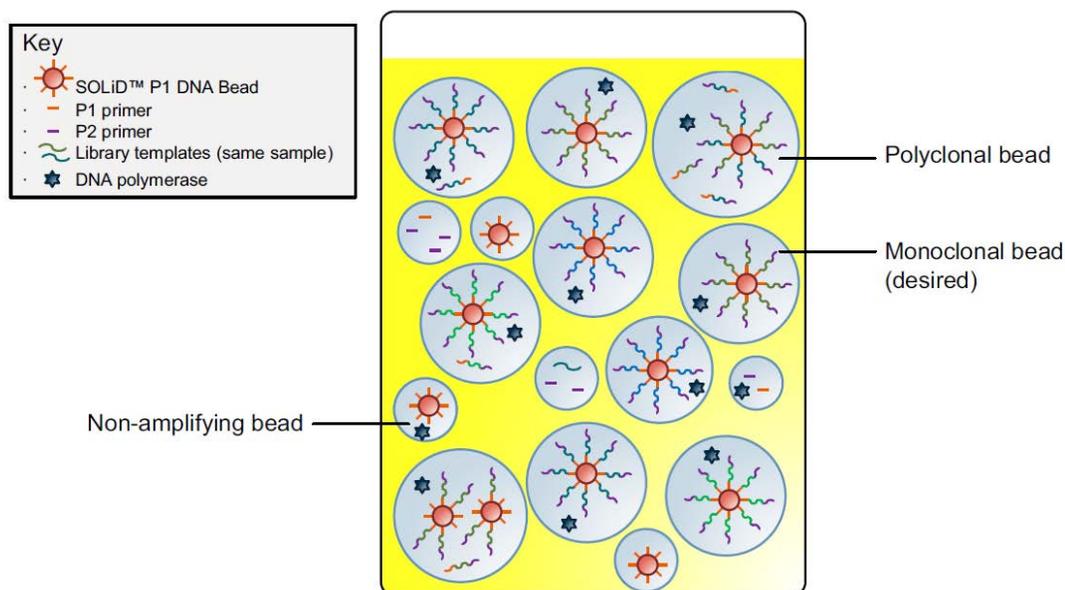


Figura 4. Esquema del método de PCR por emulsión utilizado por el *SOLiD4*. Imagen tomada de: *ABi SOLiD 4 Templated Bead Preparation Guide*.

Los ADN molde de ADN se llevan concentración de 0.5pM para maximizar la probabilidad de que, en una misma cámara de reacción, se obtengan alrededor de 30.000 copias de ADN. En este proceso, la polimerasa reconoce el primer con el adaptador P1 luego de que el ADN molde se ha unido a las perlas magnéticas que contienen secuencias complementarias para el primer con el adaptador P1. La secuencia complementaria al ADN molde se extiende fuera de la superficie de la perla magnética y, finalmente, se disocia del ADN molde.

Después del ePCR, los ADN molde se desnaturalizan y se realiza el *bead enrichment* con perlas magnéticas de poliestireno, cubiertas con adaptadores P2, que capturan ADN molde unidos a beads con P2, recuperándose alrededor de 200 millones de beads, es decir, un 80% del total de beads con P2 (Tomado del: *ABi SOLiD 4 Templated Bead Preparation Guide*). Los beads unidos a ADN molde se separan de los beads vacíos utilizando un gradiente de glicerol al 60% (Fig. 5).

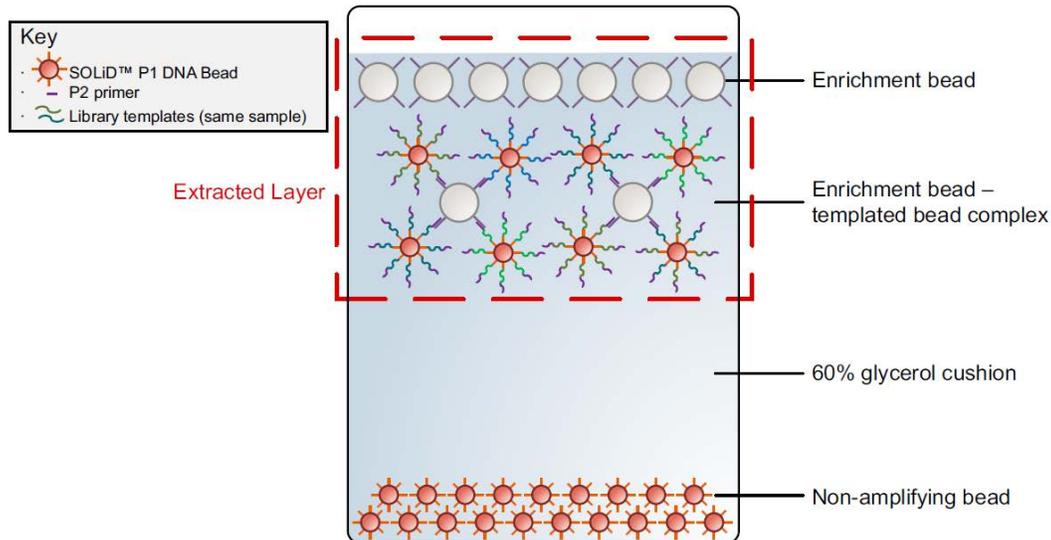


Figura 5. Esquema de la separación de las perlas o *beads* magnéticas unidas al ADN molde de ADN de interés luego de la centrifugación en un medio con glicerol al 60%. Imagen tomada de: *ABi SOLiD 4 Templated Bead Preparation Guide*.

Posteriormente, los ADN molde unidos a los beads se modifican en el extremo 3', con una transferasa de desoxinucleótidos por 4 horas. De esta forma, los ADN molde pueden unirse covalentemente a la superficie de vidrio—que ha sido modificada químicamente—de la celda de flujo.

4.5.4.4 Secuenciación por ligación

El *SOLID4*—a diferencia de otros métodos de secuenciación que utilizan la ADN polimerasa—utiliza el método de secuenciación por ligación. En este método, un primer universal se une al adaptador P1 del ADN molde. Luego, se añaden 1084 oligonucleótidos (octámeros) con las dos primeras bases marcadas con fluorocromos, estos compiten por unirse a la secuencia desconocida. El octámero se une a la cadena molde—por medio de una ligasa de ADN—cuando los primeros cinco nucleótidos del extremo 3' son totalmente complementarios. Luego, una cámara de alta resolución toma imágenes de la fluorescencia emitida, permitiendo conocer la identidad de los nucleótidos. La especificidad se da cuando dos del oligonucleótido que están marcadas compiten cada primera y segunda base en cada reacción de ligación. Además, cada oligonucleótido es construido de manera tal que posee un sitio de ruptura, el cual, una vez unida la secuencia

complementaria, se desprende del oligonucleótido, dejando el extremo 5' de la hebra elongada libre para la próxima reacción de ligación (Fig. 6). Luego de una serie de ciclos de ligación, el producto elongado es retirado y la hebra molde es elongada por otros 5 sets de primers diferentes. Al final de la reacción de secuenciación, cada base es analizada en al menos dos reacciones de ligación independientes por dos primers diferentes.

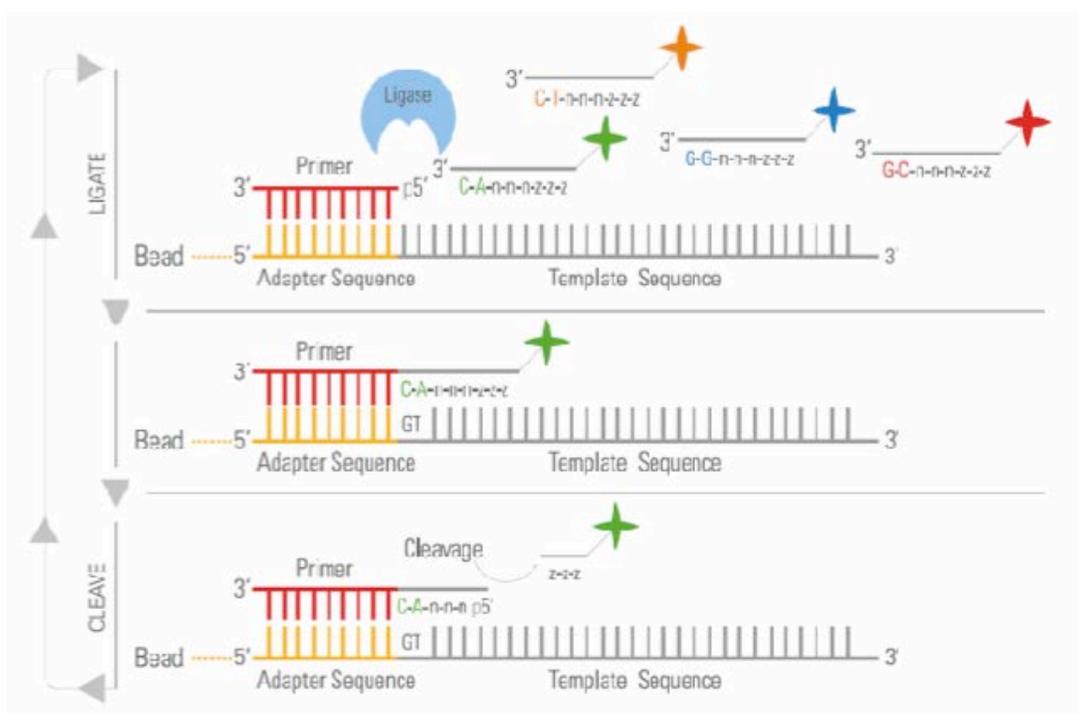


Figura 6. Esquema de la reacción por ligación utilizada para la secuenciación del exoma por el *SOLiD4*. Imagen tomada de: http://www3.appliedbiosystems.com/AB_Home/applicationstechnologies/SOLiDSystemSequencing/OverviewofSOLiDSequencingChemistry/index.htm.

4.5.4.5 Genotipo y mapeo de variantes

Para el genotipo y mapeo de las lecturas realizadas por el secuenciador se utilizó el SOLiD™ Accuracy Enhancement Tool (SAET), una herramienta bioinformática que permite asignar de manera precisa SNPs y aumentar la lectura de los contigs. Luego, las secuencias se alinearon al genoma de referencia hg19 (UCSC) por medio del software LifeScope 2.5. Se utilizaron dos programas bioinformáticos para detectar variaciones en el exoma. GATK (Genome Analysis Toolkit) (McKenna *et al.* 2010) para corregir la calidad de la asignación de bases (base quality

score recalibration), realignment de indels, remoción de duplicaciones, así como, la identificación de SNP e Indels y el genotipo en las cuatro muestras simultáneamente utilizando los parámetros estándar del programa (DePristo *et al.* 2011); y Samtools (Li *et al.* 2009). Además, se utilizó la herramienta FASTQC para evaluar la calidad de los archivos BAM y FASTQ que permiten conocer la calidad de los resultados de la secuenciación.

4.5.4.6 Anotación de variantes

Se utilizó el software ANNOVAR (Wang *et al.* 2010) para clasificar las variantes de acuerdo al locus—en exónicas, intrónicas, sitios de splicing, 5'UTR, 3'UTR, corriente arriba, corriente abajo o intergénicas—, así como, clasificar las variantes exónicas de acuerdo a la función (sinónimas, no sinónimas, inserción/delección que alteran el marco abierto de lectura, pérdida de sentido, etc).

4.5.4.7 Validación de variantes

4.5.4.7.1 Diseño de primers

Se diseñaron primers para aquellas mutaciones homocigotas, no sinónimas o con pérdida de sentido, en exones dentro de los ROH o región de ligamiento y que no estuvieran reportadas en bases de datos de SNPs ni en 230 controles internos (Fig. 7). Se utilizó el programa Primer3 v.0.4.0[®] para diseñar los primers (Cuadro 5). La secuencia de ADN se obtuvo del genome browser del UCSC[®] tomando en cuenta 300pb antes y después de la mutación de interés. Para verificar que los primers estuvieran bien diseñados y que no hubieran SNPs reportados en la secuencia genómica donde los primers se adhieren se realizó un PCR *in silico* utilizando el genome browser de UCSC[®].

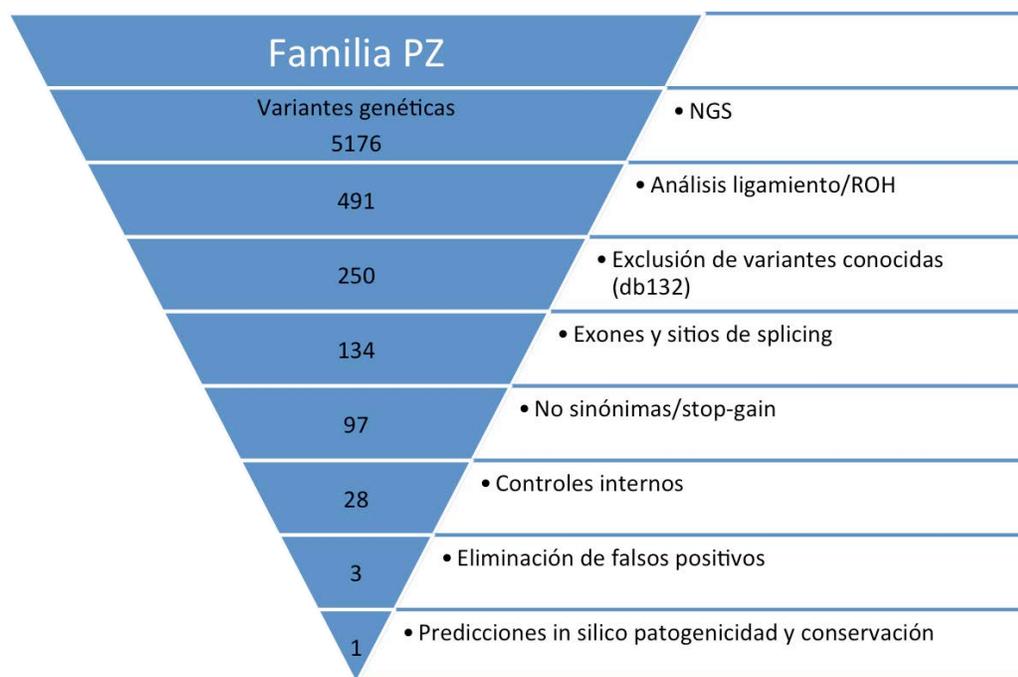


Figura 7. Esquema del método utilizado para filtrar las variantes detectadas por NGS.

Cuadro 5. Primers diseñados para las variantes de interés.

Familia	Gen	Primer F	Primer R	Amplicón (pb)	Posición	Exón	Genotipo referencia	Variante	Falso positivo
PZ	SMPD3	CGTCACCCTTGTACTCCACA	TCTGCACTGCCCTCTTCAC	474	68405685	3	G	T	Sí
PZ	ZNF646	GCTGGGAAAACCAGACAAAA	CTCATGGGACTGCTGGTGTT	545	31088336	2	C	T	No
PZ	ZNF629	TGCAGATGTAGGGCTTCTCC	AGAGGTGAGGGGCACAAGAA	543	30795441	3	G	C	No
PZ	SETD1A	GTGGACCACGACACCATCAT	GGCACCAGGGGAATAAATA	371	30995220	19	A	T	Sí
PZ	GPT2	CCTCATCTGTCAAATGGGATT	AGCAACCCCAAAACAGGTCT	497	46931602	5	G	A	No
PZ	GPT2	CCTCATCTGTCAAATGGGATT	AGAGCAACCCCAAAACAGGT	499	46931602	5	G	A	No
PZ	SBK1	TGCTAGAGAGGGAGCTGGAG	GCTCCAGGGCCAGTAAGC	553	28,331,080-28,332,513	4	Exón4	Exón4	Sí
PZ	SBK2	CCTTCTCGAGGAGTTCGTG	CGAGGTGGCTCCATTCC	582	28,331,080-28,332,514	4	Exón4	Exón4	Sí
PZ	TAOK2	CAGGCCATAGGTGAGAAGGA	GGGCCAAGGAGAAAGAAGAC	532	29997004	15	C	S	Sí
PZ	KIAA0556	TGTAGTCAATGTCCTCATAACAAG	CTGGGCTCAAAGTGATCCTC	400	27642430	5	A	W	Sí
ME	SNTG2	TGCATTTTAACCTAAACTTGGTTTG	CAGAGGGATCAGCAGAGGTT	463	1094080	4	A	T	Sí
ME	RAB17	CCTAATGCAGCACCAGGAAC	GCTCCATCTCCAGAAGAACT	575	238486675	3	G	T	Sí
ME	ILKAP	ACCCACCAGATGCCAGTAAC	GGGGTATTGCTTATTGTTTGC	497	239096813	5	C	G	Sí

4.5.4.7.2 Estandarización del PCR

Se realizó un PCR de prueba en controles internos para verificar que los primers y el programa de PCR funcionaran adecuadamente. La polimerasa (WinTaq) utilizada fue fabricada en el Instituto donde se realizaron los experimentos. Una vez estandarizado el programa de PCR se prosiguió a secuenciar a toda la familia en estudio (Cuadro 6).

Cuadro 6. Programa de PCR utilizado para validar las mutaciones en genes candidatos.

	T	Tiempo (min)	Ciclos
Desnaturalización	94°C	03:00	x9
Annealing	94°C	00:20	
	65°C*	01:00	
	68°C	01:00	
Elongación	94°C	00:20	
	55°C	01:00	x29
	68°C	01:00	
Elongación final	68°C	10:00	
Final hold	15°C	05:00	

*Touchdown

La limpieza de los productos de PCR se realizó con exonucleasas. Para eliminar las enzimas, posterior a la limpieza, se utilizó el programa Exo_15min (Cuadro 7).

Cuadro 7. Programa Exo_15min utilizado para la eliminación de exonucleasas en los productos de PCR.

T	Tiempo (min)
37°C	15:00
80°C	15:00
15°C	10:00

4.5.5 Secuenciación de Sanger

Se utilizó la reacción de secuenciación correspondiente al programa BigDye55_2min (Cuadro 8). La limpieza de las secuencias se realizó por medio de beads magnéticos del kit AgencourtCleanSEQ® de la casa comercial Beckman Coulter, para esto se utilizó un estación de trabajo automatizada conocida como Beckman Coulter Biomek® NXP (Fig.8). La secuenciación se realizó por medio del analizador genético ABI-3730® de Applied Biosystems de 48 capilares. Las

secuencias se analizaron en el software Windows 32 SeqMan v.5.03[®]. Aquellas mutaciones homocigotas en los afectados, heterocigotas en ambos padres y ausentes en los hermanos sanos se tomaron como mutaciones candidatas y posiblemente responsables del fenotipo de discapacidad cognitiva en la familia.

Cuadro 8. Programa utilizado para la reacción de secuenciación.

T	Tiempo (min)	Ciclos
96°C	00:10	
55°C	00:10	x25
60°C	02:00	
15°C	10:00	

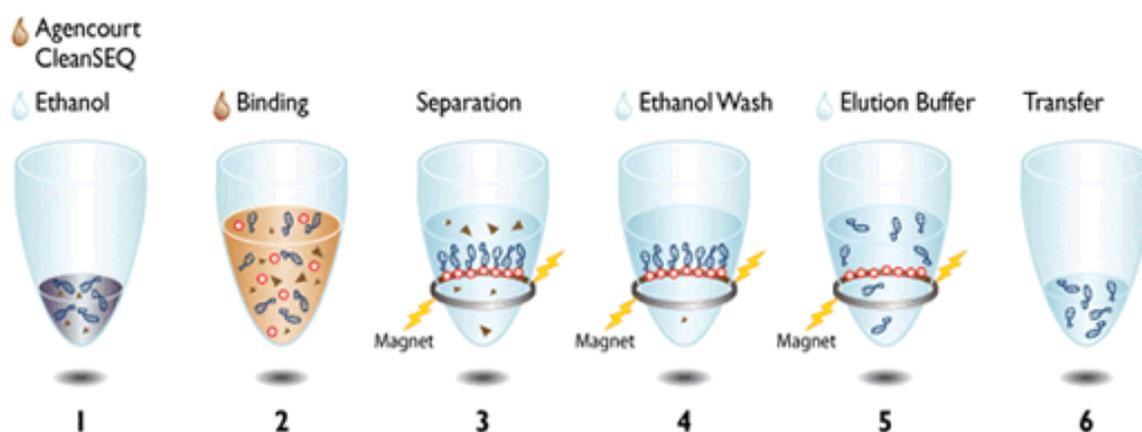


Figura 8. Esquema general del proceso de limpieza de secuencias utilizando el kit AgencourtCleanSEQ[®] por medio del cual beads magnéticos separan los productos de secuenciación de los contaminantes.

4.5.6 Modelaje proteico

Se utilizó el programa Modeller v.9.9 para modelar la proteína del gen candidato. Las estructuras de las proteínas unidas a PLP (cofactor esencial del sustrato) y cicloserina (inhibidor análogo al sustrato) se obtuvieron del Protein Data Bank (PDB).

5. RESULTADOS

5.1 Genealogías de las familias

Se encontraron cuatro familias (R, PE, ME y PZ) con más de un hijo afectado y se pudo identificar un ancestro común para las familias PE y PZ. En la familia PE ambos padres, abuelos y bisabuelos de los afectados provenían de la provincia de Puntarenas y, por lo tanto, fue posible unir la genealogía ocho generaciones arriba. La familia PZ, oriunda del mismo pueblo en Pérez Zeledón, se trataba de un matrimonio entre primos hermanos. En la familia ME ambos padres, ambos abuelos paternos y la abuela materna provenían de Puntarenas, sin embargo, el abuelo y bisabuelos maternos provenían de Heredia y el bisabuelo paterno provenía de Nicaragua. Para la familia ME no fue posible identificar un ancestro común por la presencia de hijos naturales reportados para esta familia en los registros civiles y eclesiásticos. Tampoco fue posible establecer un ancestro común para la familia R pues las familias de ambos padres provenían de localidades diferentes: el padre provenía de Guanacaste, la madre de Puntarenas, los bisabuelos maternos de Alajuela y los bisabuelos paternos eran de origen nicaragüense (Fig.9).

5.2 Diagnóstico clínico

Las individuos afectados de PE, ME, PZ y R carecían de malformaciones físicas, tenían problemas para comunicarse, dependían de otros para sus necesidades básicas como alimentación, vestimenta, aseo personal y carecían de un diagnóstico establecido previamente. Para las cuatro familias se descartó, luego de revisar los expedientes clínicos, que la discapacidad cognitiva se debiera a problemas prenatales, perinatales o postnatales, traumas en la cabeza u otro accidente.

5.2.1 Familia PE

Los individuos PE001 y PE004 fueron referidos por la Escuela de Enseñanza Especial de Puntarenas debido a dificultades en el aprendizaje y problemas en el lenguaje. Ambos embarazos fueron normales y sin ninguna complicación, sin embargo, el parto de PE004 fue cesárea debido a que se trataba de un parto pélvico. El tamaño de la cabeza de PE004, a simple vista, se observó más grande de lo normal con respecto al tamaño del cuerpo, no levantó la cabeza hasta ni se sentó hasta cumplir un año de edad, empezó a caminar a los 2.7 años y, previamente, había sido diagnosticado por las psicólogas con retraso general del desarrollo. Según la escala de EDIN, PZ004 de 4 años presentó habilidades motoras (finas y gruesas) y cognitivas de un niño de 1.5 a 2 años,

desarrollo del lenguaje de un niño menor de 1 año, hábitos de salud y nutrición correspondientes a niños de 1-1.5 años y habilidades socioafectivas de un niño de 2 a 2.5 años. Por otro lado, PZ001 fue referida por problemas de lenguaje (hablaba rápido y enredado) exclusivamente. PZ003 también presentaba problemas en el aprendizaje, la memoria y dificultades para leer y escribir, sin embargo, no se pudo determinar la causa del retardo mental como genético debido a que sufrió un evento de hipoxia cuando tenía pocos meses de nacida. PE001 presentó un CI<50 y los demás parientes fenotípicamente sanos obtuvieron un CI>70 (Cuadro 9).

Cuadro 9. Puntajes de coeficiente intelectual obtenidos para la familia PE.

Familia PE	CI
P011	109
PE002	70
PE003	76
PE001	72
PE004	†

† No se pudo evaluar porque tenía 4 años

5.2.2 Familia ME

Los individuos afectados ME005 y ME009 habían sido diagnosticados en el Hospital Nacional de Niños, con autismo, retardo mental moderado, conductas atípicas, déficit atencional e hiperactividad problemas en el desarrollo psicomotor y lenguaje. ME005 presentó conductas repetitivas, ejecutaba órdenes y era capaz de vestirse y bañarse solo. ME009 es afectivo con personas que conoce, depende completamente de alguien para su aseo personal y no logra comunicarse ni atender órdenes. Ambos partos fueron vaginales y sin ninguna complicación. ME005 presentó un CI<50 y los demás parientes fenotípicamente sanos obtuvieron un CI>70 (Cuadro 10).

Cuadro 10. Puntajes de coeficiente intelectual obtenidos para la familia ME.

Familia ME	CI
ME006	†
ME008	92
ME010	70
ME007	84
ME005	40
ME009	*

† No se pudo evaluar

*ME009 no se pudo evaluar porque no cumplía con las habilidades mínimas para aplicar la prueba

5.2.3 Familia PZ

Los cuatro individuos afectados de esta familia eran incapaces de comunicarse, dependían completamente de alguien más para su aseo personal y habilidades adaptativas. PZ0006 y PZ007 presentaron movimientos espásticos en los miembros inferiores. Por sus dificultades en la comunicación, no fueron evaluados con las pruebas de inteligencia ya que no cumplían con las habilidades verbales mínimas para su aplicación. Por lo tanto, se tomó en cuenta el criterio clínico de una psicóloga para diagnosticar discapacidad cognitiva profunda en los cuatro individuos afectados, además, de que han sido admitidos en el CAIPAD (Centro de Atención Integral para la Persona Adulta con Discapacidad) de Pérez Zeledón. No se observaron malformaciones físicas de ningún tipo.

5.2.4 Familia R

R012 y R1016 presentaron orejas de implantación baja, cara alargada y comportamientos agresivos. A ambos les costó aprender a caminar, son incapaces de atender órdenes y de realizar su aseo personal. Fueron referidos por poseer discapacidad cognitiva y problemas de aprendizaje.

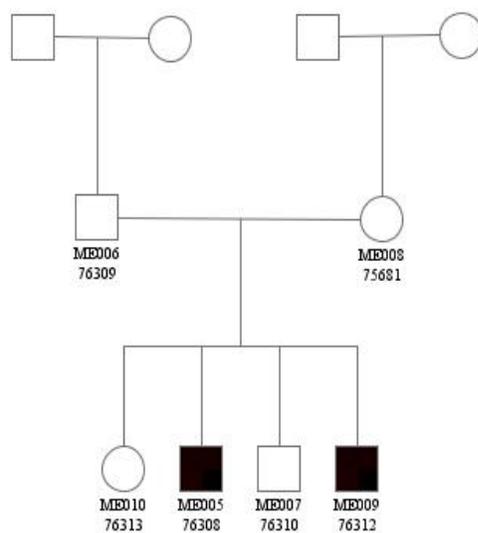


Figura 10. Genealogía de la familia ME del Mojón de Esparza, Puntarenas.

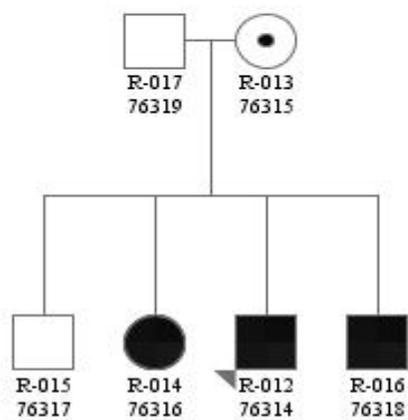


Figura 11. Genealogía de la familia R del Roble, Puntarenas.

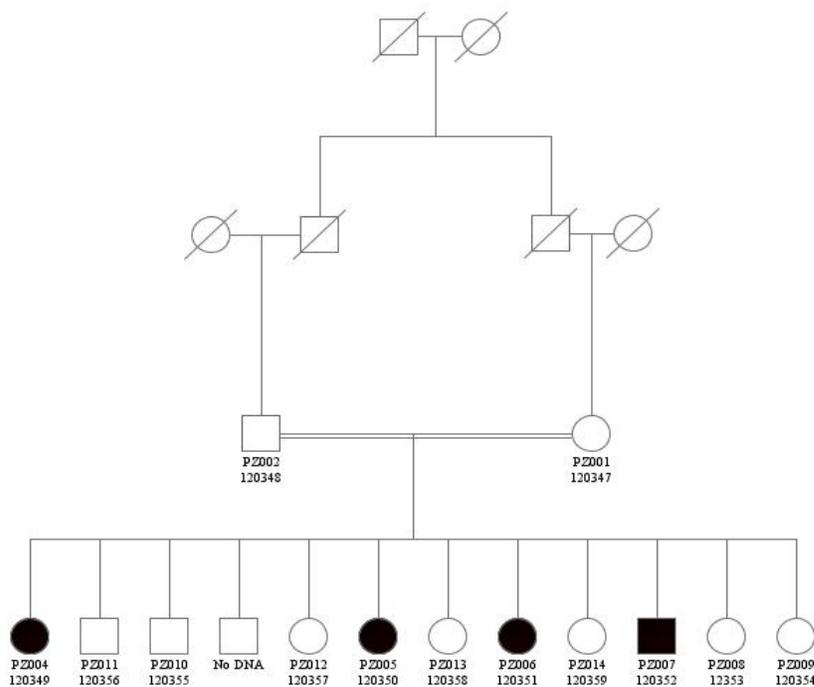


Figura 12. Genealogía de la familia PZ de Pérez Zeledón.

5.3 Técnicas moleculares

5.3.1 Exámenes de paternidad

El individuo R12 (DNA#76314) no compartía 13/15 marcadores con el padre R17 (DNA#76319). Del mismo modo, PZ012 fue excluida del análisis por no compartir 11/15 marcadores con el padre PZ002. Los demás individuos de las familias ME, PE, R y PZ compartieron todos los marcadores con sus progenitores.

5.3.2 Síndromes asociados a discapacidad cognitiva

5.3.2.1 MLPA

En ninguna familia se encontró ninguna microdelección, duplicación o translocación submicroscópica asociadas a síndromes comunes de discapacidad cognitiva contenidos dentro de los cuatro kits de MLPAs utilizados.

5.3.2.2 Diagnóstico de X-frágil

Descartamos el síndrome de x-frágil en todas las familias a excepción de la familia R. El individuo R-012 presentó 97 repeticiones CGG en el exón 1 del gen FMR1, lo que es característico de una premutación (59-200 repeticiones) (Fig. 13). El resultado del Southern-Blot mostró la mutación completa pero ninguna premutación. Por lo tanto, R-012 presenta un mosaicismo de células con la mutación completa y una pequeña fracción de células con la premutación. Por otro lado, de R-016 no se obtuvo ningún producto de PCR, sin embargo, los resultados del Southern-Blot mostraron una expansión de repeticiones característica de una mutación completa. R-014 presentó un número de repeticiones normal (19 repeticiones) en un alelo, después del PCR, sin embargo, con el Southern-Blot se evidenció que presentaba la mutación completa en el otro alelo. Los patrones de metilaciones observados en R-014 mostraron que la paciente presenta tanto el alelo normal metilado (inactivo) y no metilado (activo). Esto confirma el diagnóstico de X-frágil en los individuos R-012, R-014 y R-016. En los individuos R-015 (Fig. 14) y R-017 no se encontraron expansiones de las repeticiones CGG.

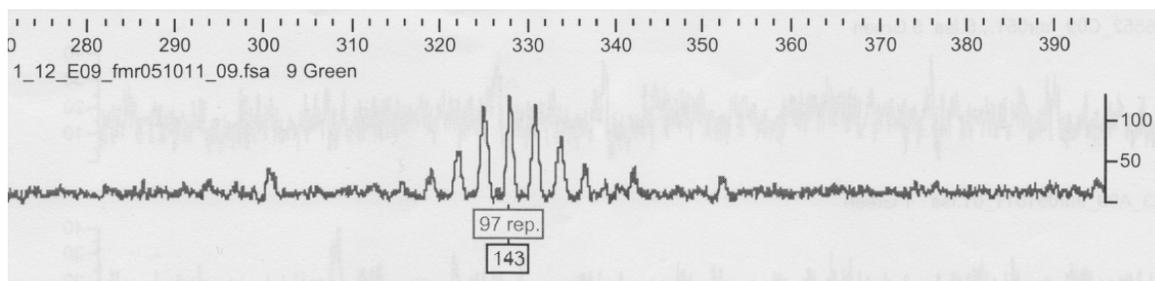


Figura 13. Electroferograma indicando las repeticiones CGG en el gen FMR1 características de una premutación del paciente R012.

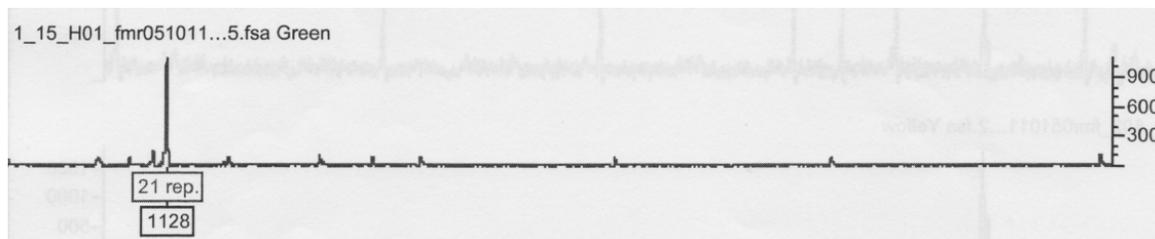


Figura 14. Electroferograma indicando las repeticiones CGG en el gen FMR1 del individuo sano R015.

El diagnóstico de X-Frágil en esta familia se explica porque la madre R-013 presentó—luego de analizar los resultados del Southern-Blot—un alelo con la mutación completa, sin embargo, el alelo normal no se encuentra metilado y, por lo tanto, se cree que está activo en las células analizadas. R-013 es portadora asintomática de la mutación completa causante de X-Frágil.

5.3.2.3 Inactivación del cromosoma X

No encontramos patrones no aleatorios de inactivación del cromosoma X que pudieran indicar que las familias presentaran una enfermedad con herencia ligada al cromosoma X.

5.3.3 Microarreglos de ADN

5.3.3.1 Variantes en el número de copias (CNV)

No se encontraron deleciones o duplicaciones patogénicas en regiones genómicas mayores a 100kb para ninguna de las tres familias.

5.3.3.2 Análisis de ligamiento

En las familias ME y PE no se encontraron puntajes LOD superiores a 2—indicativos de ligamiento—en ningún cromosoma, tanto para el modelo recesivo como para el modelo dominante. Sin embargo, para ME el puntaje LOD más alto correspondió al cromosoma 1 (LOD=0.85) con el modelo recesivo (Fig. 15). PE presentó el mayor puntaje LOD con el modelo recesivo en el cromosoma 1 (LOD=0.60 dominante y LOD=0.73 recesivo) (Fig. 16). Luego de correr los cinco análisis, se comprobó que la familia PZ tenía una región candidata estadísticamente significativa en el cromosoma 16 (LOD=3.4) (Fig. 17). En el cromosoma 2 se detectó otra región candidata estadísticamente significativa (LOD=3.1). Los haplotipos para los análisis de ligamiento de la familia PZ se pueden observar en el Anexo 9.5.

A. ME recesivo

Project:	TANYA	Marker	CHR	cM	LOD	ALPHA	HLOD
Family name:	TOTALS	1. SNPAB495626	1	197.10	0.8490	0.0000	0.8490
Used map:	AFFY 1M deCODE 02122007 v6 (sex-averaged)	2. SNPAB495626 - SNPAB666397	1	197.11	0.8490	0.0000	0.8490
Marker positions:	132901 cM / 0.7 / 0 outside	3. SNPAB666397	1	197.12	0.8490	0.0000	0.8490
Allele frequencies:	Codominant	4. SNPAB666397 - SNPAB551606	1	197.13	0.8490	0.0000	0.8490
Scoring / Weight:	pairs / Equal	5. SNPAB551606	1	197.14	0.8490	0.0000	0.8490

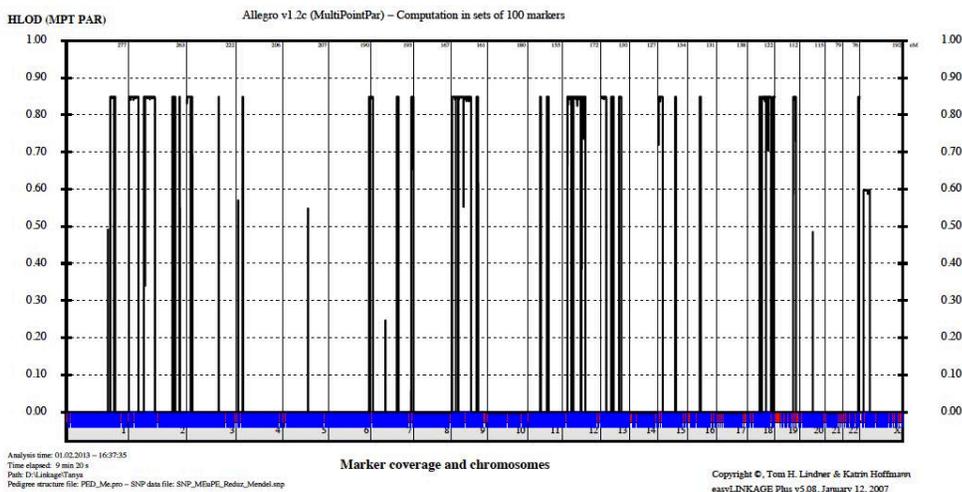


Figura 15. Puntajes LOD de todo el genoma luego de realizar el análisis de ligamiento utilizando un modelo de herencia recesivo en la familia ME.

Project:	TANYA	Marker	CHR	cM	LOD	ALPHA	HLOD
Family name:	TOTALS	1. SNPAB8315688	1	0.88	0.7256	0.0000	0.7256
Used map:	AFFY 1M deCODE 02122007 v6 (sex-averaged)	2. SNPAB2115098	1	2.03	0.7256	0.0000	0.7256
Marker positions:	132901 cM / 0.7 / 0 outside	3. SNPAB678738	1	2.29	0.7256	0.0000	0.7256
Allele frequencies:	Codominant	4. SNPAB678738 - SNPAB317896	1	2.30	0.7256	0.0000	0.7256
Scoring / Weight:	pairs / Equal	5. SNPAB317896	1	2.32	0.7256	0.0000	0.7256

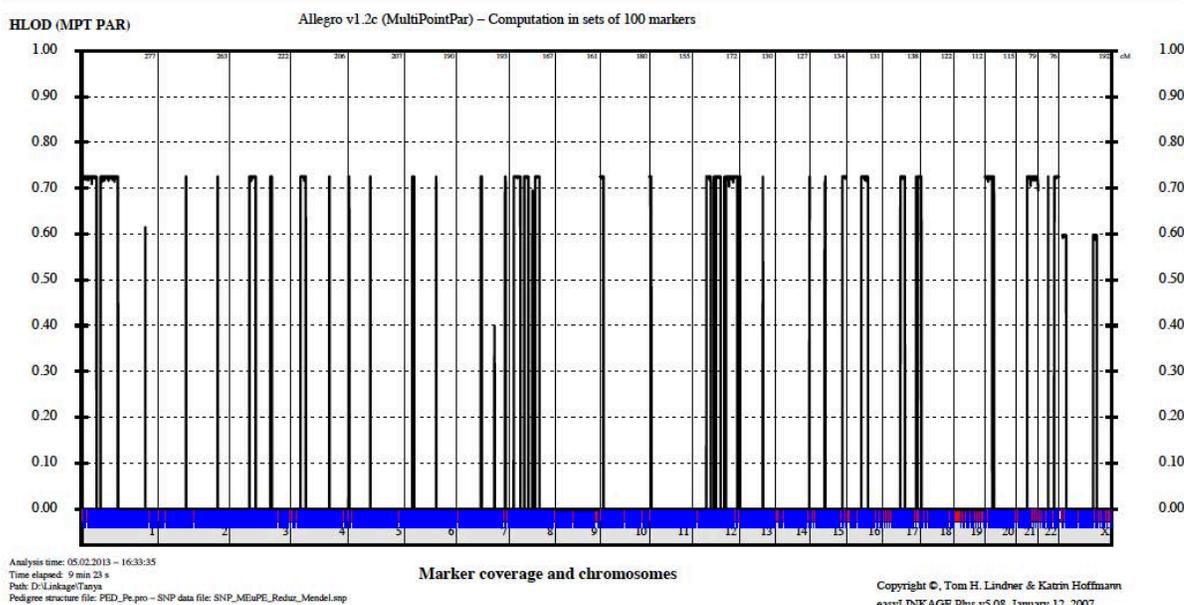


Figura 16. Puntajes LOD de todo el genoma luego de realizar el análisis de ligamiento utilizando un modelo de herencia recesivo en la familia PE.

Project:	TANYA	Marker	CHR	cM	LOD	ALPHA	HLOD
Family name:	TOTALS	1. SNPA8426033	16	48.07	3.3803	0.0000	3.3803
Used map:	AFFY 1M deCODE 02122007 v6 (sex-averaged)	2. SNPA2206771	16	49.44	3.3803	0.0000	3.3803
Marker positions:	65241 ok / 0 ? / 0 outside	3. SNPA2301872	16	53.39	3.3803	0.0000	3.3803
Allele frequencies:	Codominant	4. SNPA2301872 - SNPA2110369	16	53.41	3.3803	0.0000	3.3803
Scoring / Weight:	pairs / Equal	5. SNPA2110369	16	53.43	3.3803	0.0000	3.3803

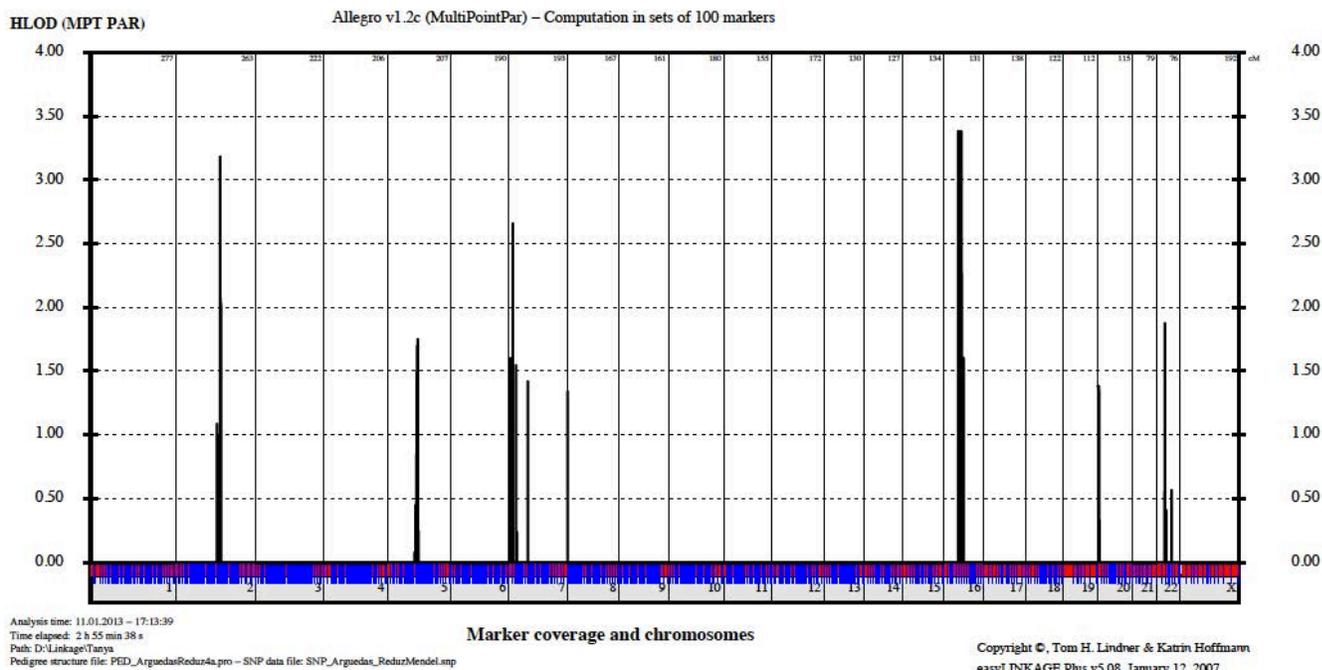


Figura 17. Puntajes LOD de todo el genoma luego de realizar el análisis de ligamiento con el modelo recesivo en la familia PZ (Pedigree D Anexo 9.4).

5.3.3.3 Carreras de homocigosidad (ROH), mapeo por homocigosidad y genes candidatos

5.3.3.3.1 Familia ME

Ambos individuos afectados, ME005 y ME009, compartieron únicamente un ROH de alrededor de 1Mb en el brazo largo del cromosoma 2 (2q34). En este ROH encontramos dos genes candidatos: IKZF2 (proteína dedo de zinc) y SPAG16 (Homo sapiens sperm associated antigen 16) (Fig. 18). Ambos genes se descartaron como genes candidatos debido a que las variaciones encontradas en IKZF2 fueron SNPs en intrones y en SPAG16 encontramos dos variaciones, una mutación no sinónima en un exón y otra en un intrón, sin embargo, ambas han sido reportadas previamente en bases de datos de dbSNP, así como, en los 230 controles internos del Instituto de Genética Humana de la Universidad de Erlangen-Nuremberg. Al revisar los InDels de todo el genoma, se identificó una deleción en el cromosoma 2 que altera el marco de lectura del gen neuropilina 2 (NRP2) en el

individuo afectado ME009 (Fig. 19). Además, se encontraron otras variaciones homocigotas, no sinónimas, cercanas al ROH y relacionadas al funcionamiento del SNC, en los genes SNTG2 (Sintrofina Gamma-2), RAB17 (Ras-associated protein) e ILKAP (Homo sapiens integrin-linked kinase-associated serine/threonine phosphatase). Sin embargo, después de la validación se comprobó que todas eran falsos positivos (Anexo 9.5).

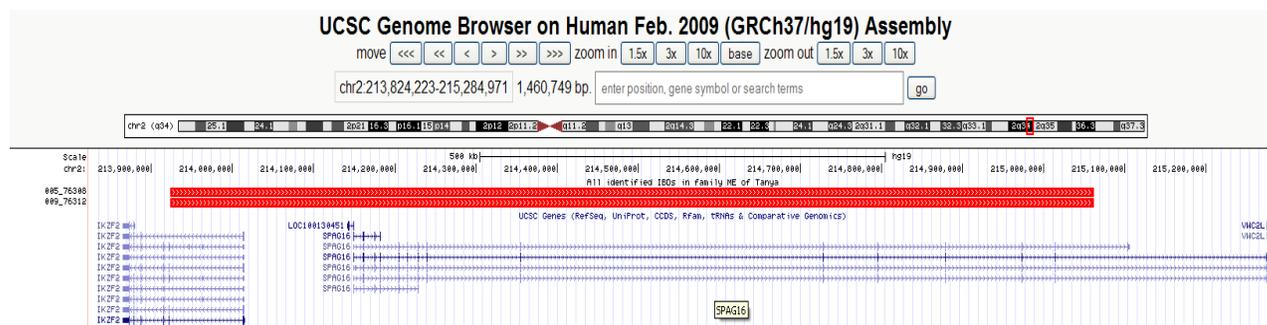


Figura 18. Vista del genome browser de UCSC donde se muestran las carreras de homocigosidad en chr2: 213824223-215284971 y, los genes dentro de ellas, compartidas por ambos afectados de la familia ME.

5.3.3.3.2 Familia PE

Ambos individuos afectados, PE001 y PE004 compartieron únicamente un ROH de alrededor de 1 megabase en el brazo largo del cromosoma 12 (12q13). En esta región se encontraron 26 genes de los cuales dos se encontraban directamente relacionados con funciones neurológicas: neurotransmisión (*ACCN2-Homo sapiens amiloride-sensitive cation channel 2*) y neurodesarrollo (*POU6F1- Homo sapiens POU class 6 homeobox 1*). Sin embargo, estos genes no tenían mutaciones en homocigosis. No se encontraron inserciones o deleciones candidatas.

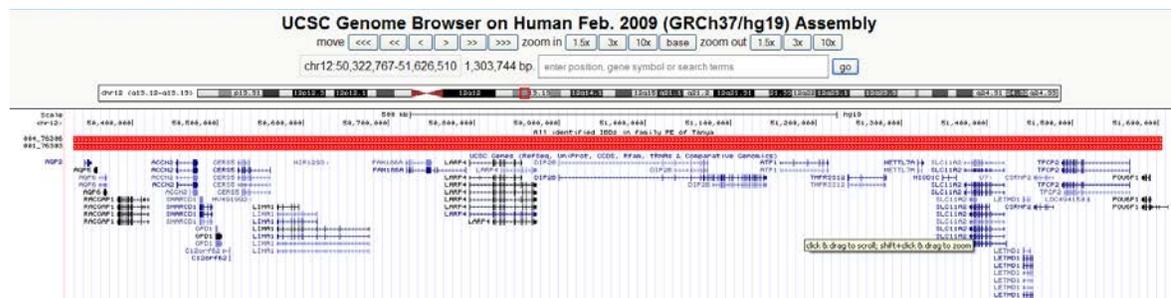


Figura 19. Vista del genome browser de UCSC donde se muestran las carreras de homocigosidad en chr12: 50322767-51626510 y, los genes dentro de ellas, compartidas por ambos afectados de la familia PE.

5.3.3.3 Familia PZ

En esta familia se encontró que los cuatro individuos afectados: PZ004 (DNA#120349), PZ005 (DNA#120350), PZ006 (DNA#120351), PZ007 (DNA#120352), compartían ROH únicamente en el cromosoma 16. Se encontró un ROH que compartían los cuatro afectados en chr16: 27157613-32483006 (Fig. 20). Sin embargo, luego del mapeo por homocigosis, la región de interés resultó ser más extensa: chr16: 27177795-48202503 y contenía 226 genes de los cuales 56 no son codificantes (miARN, pseudogenes) y 11 se han reportado expresados en el cerebro (Fig. 21). Después de filtrar los datos del NGS, en esta región, se encontraron mutaciones no sinónimas en los genes *ZNF646* (Zinc finger 646), *ZNF629* (Zinc finger 629) y *GPT2* (Glutamate pyruvate transaminase 2) (Fig. 22). Los tres cambios genéticos segregaban con la enfermedad en la familia (Cuadro 9). Sin embargo, para *ZNF646* y *ZNF629*, los cambios genéticos se encontraron reportados en dbSNP137 con una frecuencia alélica de 1:2000. También se secuenciaron cambios genéticos en los genes *SMPD3* (Sphingomyelin Phosphodiesterase 3), *SETD1A* (Set domain-containing protein 1A), *TAOK2* (TAO kinase 2) y *KIAA0556* (ARNm no caracterizado), que tenían mutaciones heterocigotas pero se encontraban dentro del ROH de interés. Del mismo modo, se secuenció el exón 4 del gen y el exón 4 de *SBKI* (Homo sapiens SH3-binding domain kinase 1) que no fue cubierto por el NGS y se trataba de un gen relacionado con el neurodesarrollo. Todas las variantes candidatas resultaron ser falsos positivos, es decir, no se detectaban los cambios genéticos luego de la validación por el método de Sanger (Anexo 9.7).

Los programas PhyloP, PolyPhen, SIFT y Mutation Taster predijeron que el cambio genético en *GPT2* c.G286A se daba en una región altamente conservada y su efecto era patogénico. Por el contrario, para *ZNF646* (c.C691T) predijeron que era un cambio genético en una región poco conservada y con un efecto tolerable o benigno y para *ZNF629*: c.C208G no se obtuvieron resultados (Cuadro 10).

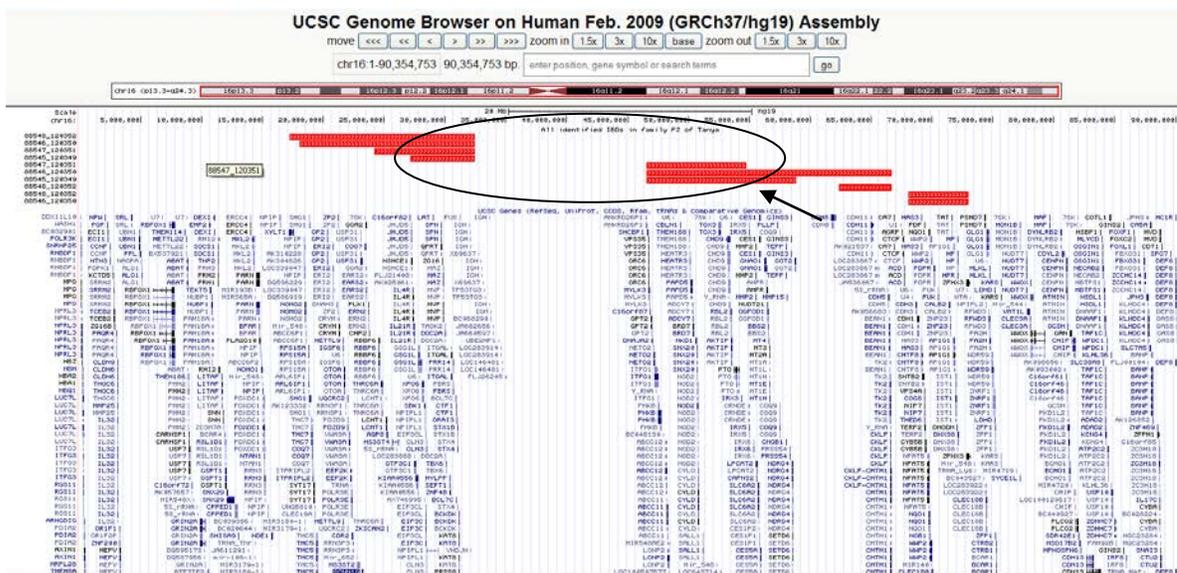


Figura 20. Vista del genome browser de UCSC donde se muestran las carreras de homocigosidad en chr16: 27157613-32483006 y, los genes dentro de ellas, compartidas por ambos afectados de la familia PZ.

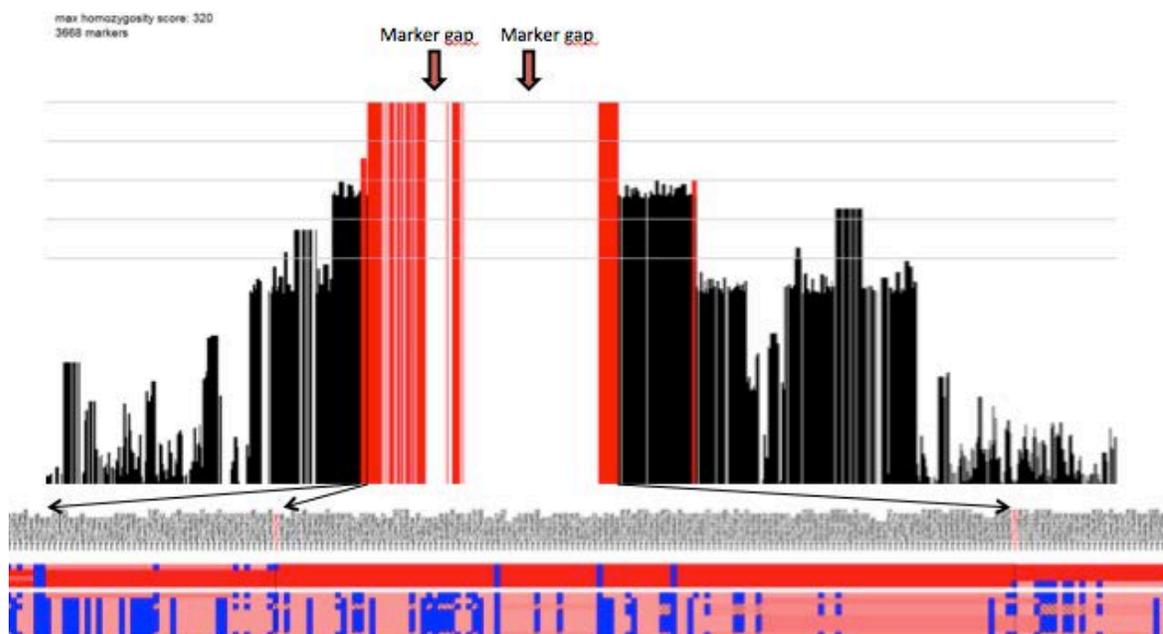


Figura 21. Vista de la región candidata luego de realizar el mapeo por homocigosidad en el genoma de los individuos de la familia PZ. La región oscila entre los marcadores rs2090296 y rs17821664 cuya posición genómica es chr16: 27177795-48202503.

Cuadro 11. Lista de mutaciones identificadas en la familia PZ.

Gen	Cromosoma	Posición	Exón	Cambio genético	Efecto
<i>ZNF646</i>	16	31088336	2	C/T	p.P231S
<i>ZNF629</i>	16	30795441	3	C/G	p.P70A
<i>GPT2</i>	16	46931602	5	G/A	p.G96R

Cuadro 12. Predicción realizada por los programas SIFT, PhyloP, PolyPhen y Mutation Taster para las mutaciones en genes candidatos de discapacidad idiopática no sindrómica en la familia PZ.

Gen	SIFT	Puntaje	PhyloP	Puntaje	PolyPhen	Puntaje	Mutation Taster	Puntaje
ZNF646	T	0.79	N	0.23	B	0	N	0.023
GPT2	D	1	C	0.99	D	0.99	D	0.99

- **T=tolerable, N=normal, B=benigno, D=deletéreo, C=conservado**

5.4 Cobertura del SOLiD4 con el IGV

La cobertura del SOLiD4 de la región candidata para la familia PZ (chr16:27000000-48000000) fue cercana a un 80%, es decir, de 256 exones pertenecientes a 171 genes codificantes, 46 exones no fueron cubiertos por en NGS y de esos 10 pertenecieron a 14 genes que se expresan en el cerebro. Las regiones candidatas de las familias ME y PE fueron cubiertas en la totalidad de sus exones.

5.5 Modelaje proteico

5.5.1 Dedos de zinc ZNF646 y ZNF629

Al observar el efecto de ambas mutaciones en la estructura de la proteína no se encontró ninguna relación con la patogénesis de la enfermedad. Ambas mutaciones se localizaron en regiones fuera de la estructura globular de la proteína. Por lo tanto, mutaciones en estas posiciones tienen poco efecto sobre la estructura 3D de la proteína.

5.5.2 GPT2

La mutación G96R se localizó dentro de un bucle donde se une el sustrato al sitio activo de la enzima GPT2 (Modeller v9.9). Por medio del modelaje proteico se logró determinar que sin la mutación G96R la proteína es flexible y permite la unión del sustrato al sitio activo, sin embargo,

después de la mutación se observó *in silico* que el sustrato no es disuelto por la enzima. Estos podrían ser los primeros indicios de que la glicina es necesaria para la flexibilidad del bucle y, por ende, el buen funcionamiento de la enzima. Esto concuerda con las características estructurales de los aminoácidos que intervienen, ya que, la glicina es un aminoácido no polar y pequeño que confiere movilidad a la proteína y, debido a la mutación, es reemplazado por la arginina, un aminoácido cargado positivamente y polar, lo cual, le otorga mayor rigidez a GPT2 (Lee & Cox 2008).

La estructura cristalina de una proteína homóloga a GPT2 se modeló utilizando el programa Modeller v.9.9 (Sali & Blundell 1993) en presencia de PLP (PDB: AP:3IHJ)—cofactor esencial que se une al sitio activo independientemente del sustrato o inhibidor—y cicloserina (PDB: AP:3TCM)—un análogo al sustrato que actúa como inhibidor. *In silico*, la cicloserina se unió covalentemente a PLP inhibiendo los siguientes pasos de la reacción enzimática. La estructura proteica permitió mostrar que la glicina 96 se encuentra en extrema proximidad al inhibidor y, por lo tanto, está directamente relacionada con la unión del sustrato a la enzima. Por lo tanto, se cuenta con evidencia preliminar de que una arginina en lugar de la glicina altera la conformación del bucle, impidiendo la unión del sustrato al sitio activo de la enzima (Fig. 23).

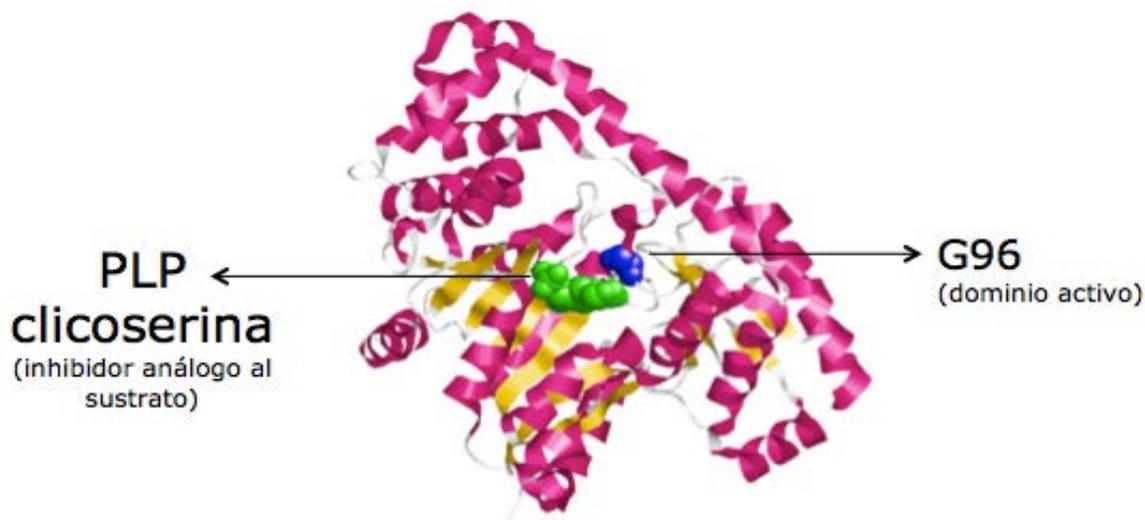


Figura 23. Modelaje de la estructura proteica de GPT2. En verde, un análogo del sustrato, el inhibidor cicloserina (verde) y, en azul, la glicina 96 ubicada en el sitio activo de la enzima. Imagen producida por el de Dr.rer.nat Heinrich Sticht del Instituto de Bioquímica de la Universidad de Erlangen-Nuremberg.

6. DISCUSIÓN

En esta investigación fue posible elucidar cambios genéticos posiblemente causantes de discapacidad cognitiva no sindrómica, en familias consanguíneas y con más de un individuo afectado, utilizando la combinación de técnicas moleculares de microarreglos de ADN (ROH) y secuenciación de nueva generación. Con la metodología antes expuesta se encontró una mutación no sinónima en el sitio activo de un gen candidato (GPT2) de causar discapacidad cognitiva que nunca antes ha sido relacionado con el fenotipo de discapacidad cognitiva no sindrómica ni con ninguna otra enfermedad. Sin embargo, se hacen necesarios análisis funcionales que demuestren que esta variante es verdaderamente causante de la enfermedad.

6.1 Consideraciones del método

La secuenciación de nueva generación ha permitido elucidar decenas de variaciones genéticas causantes de enfermedades, sin embargo, aún es un método imperfecto, cuya cobertura suele ser alrededor de un 80% de todos los exones (Bamshad *et al.* 2011). Por lo tanto, es posible que en las regiones no cubiertas por el secuenciador se encuentren otras mutaciones causantes de enfermedades. Además, la secuenciación del exoma cubre el 1% del genoma que es codificante y, es posible, que algunas enfermedades podrían estar causadas por mutaciones en regiones no codificantes pero que, por medio de algún mecanismo aún desconocido, regulen la expresión genética. Finalmente, el modelaje proteico por homología a pesar de ser el método de predicción más utilizado sigue siendo imperfecto y se hace indispensable realizar pruebas funcionales para la definición de genes candidatos (Launay y Simonson, 2008). A pesar de estas limitaciones, en esta investigación se logró encontrar, en una de las familias, una mutación en un gen para el que existen buenas razones para pensar que podría ser causante de discapacidad cognitiva no sindrómica.

La metodología utilizada es una aproximación válida para familias con varios individuos afectados en los que se sospecha un modo de herencia autosómico recesivo, cuyos padres están emparentados y en los que se desconoce la causa genética de la enfermedad. Un panorama ideal sería cubrir mediante secuenciación de Sanger todos aquellos exones de los genes que no fueron cubiertos por el NGS en la región candidata. De cualquier forma, se hace imperativo localizar más familias para aumentar las posibilidades de detectar otras mutaciones en el mismo gen o en otros genes candidatos. Está claro que esta metodología no será exitosa para encontrar todas las causas de enfermedades con herencia autosómica recesiva con pocos individuos afectados, pero en otros casos esta estrategia si podría funcionar.

Como se observó en la figura 15 la región candidata para la familia PZ incluye el centrómero del cromosoma 16 y, por lo tanto, se trata de una región problemática, ya que, las regiones centroméricas se caracterizan por contener secuencias altamente repetitivas, secuencias no codificantes y ricas en heterocromatina (Grady *et al.* 1992). Esto podría explicar el hecho de que el ROH que comparten los cuatro individuos afectados inicie en 27Mb termine en 32Mb y luego inicie de nuevo en 46Mb para únicamente tres de los cuatro de los afectados en esta familia. Es válido suponer—y fue comprobado por el análisis de ligamiento—que se trata de un solo ROH que abarca toda la región y que, por lo tanto, es compartido por todos los individuos afectados. Por lo tanto, es importante tomar en cuenta que regiones candidatas cercanas a regiones centroméricas podrían dificultar el mapeo de genes candidatos al tratarse de regiones difíciles de secuenciar.

En el caso de la familia PE en la que se logró comprobar que ambos padres eran parientes algunas generaciones arriba, no se lograron encontrar mutaciones potencialmente causantes de enfermedades. Esto pudo deberse, por ejemplo, a un problema de cobertura como se explicó anteriormente, o a que la causa genética no se encuentra en una secuencia codificante. Además, en esta familia fue difícil comprobar que ambos hermanos estuvieran afectados por la misma enfermedad debido a que uno de los individuos afectados eran todavía muy pequeño para realizar pruebas de IQ y aún no hablaba (problemas en la caracterización clínica).

Para la familia ME no se logró unir la genealogía—debido a que en los registros civiles se encontraron muchos hijos naturales dentro de esta familia, es decir, los apellidos se transmitieron de forma ilegítima o no matrimonial, sin un padre oficial—lo cual complicó identificar un ancestro en común. Además, es posible que los padres—si estaban emparentados—fuera muchas generaciones arriba, ya que, se encontró que ambos individuos afectados compartían únicamente un ROH de 1Mb con únicamente dos genes en los que no se encontraron variantes de candidatas. Además, el análisis de ligamiento se dificultó por tratarse de una familia pequeña y, por lo tanto, con la obtención de los puntajes LOD—utilizando un modelo de herencia autosómico recesivo—no fue posible identificar regiones candidatas en el genoma. Por otro lado, al realizar el análisis utilizando un modelo de herencia autosómico dominante se lograron detectar sólo dos regiones sugerentes de ligamiento—en los cromosomas 2 y 9—sin embargo, no se encontraron variantes genéticas de interés. En esta familia contar con los datos de la secuenciación del exoma de exclusivamente uno de los afectados dificultó el análisis y, es recomendable, cuando se trabaja con familias pequeñas en

las que no se logra unir la genealogía, secuenciar ambos individuos afectados para buscar variantes genéticas que tengan en común.

La familia R se resolvió en los primeros análisis de diagnóstico molecular, al detectarse expansiones anormales del triplete CGG en el gen FMR1 causante del síndrome de X-frágil (Pieretti *et al.* 1991). Esta familia ejemplifica la importancia de descartar síndromes comunes asociados a discapacidad cognitiva antes de realizar la secuenciación del exoma o los microarreglos de ADN. Las recomendaciones que se podrían dar para esta familia son de tipo consejería genética, especialmente, para las mujeres portadoras de la mutación completa que tienen un 50% de probabilidades de tener un hijo afectado (Sherman *et al.* 1985).

Un aspecto que debe tomarse en cuenta es que para enfermedades como la discapacidad cognitiva suelen buscarse mutaciones en genes directamente relacionados con la actividad de las neuronas o las sinapsis. Sin embargo, Najmabadi y colaboradores (2011) encontraron 50 genes nuevos relacionados con ID utilizando NGS, de los cuales la mayoría de los genes estaban expresados ubicuamente e involucrados en procesos celulares importantes (transcripción y traducción del ADN, degradación proteica, *splicing* del ARN, metabolismo celular, síntesis de ácidos grasos y plegamiento de proteínas) pero no directamente relacionados con la actividad neuronal o sináptica. Aún se desconoce el mecanismo por el cual defectos en procesos celulares tan diversos puedan estar afectando al cerebro. Una hipótesis es que los efectos de mutaciones en genes involucrados en procesos celulares generales—y no sólo del SNC—pueden ocasionar fenotipos de ID al alterar procesos celulares que desencadenan cambios en el funcionamiento del SNC. De esta forma, al investigar enfermedades heterogéneas—como discapacidad cognitiva no sindrómica—debe prestarse atención a cualquier mutación no sólo directamente relacionada con el cerebro sino también en genes que se expresan ubicuamente e involucrados en procesos celulares importantes (Najmabadi *et al.* 2011).

Finalmente, el abordaje de familias grandes, con más de un individuo afectado con discapacidad cognitiva no sindrómica—en las que se presupone un modo de herencia autosómico recesivo—es necesario, primero, descartar con pruebas moleculares todos los síndromes asociados a ID reportados. Luego, identificar un ancestro común cercano—dos o tres generaciones arriba—entre ambos padres sanos, ya que, se facilita la obtención de regiones homocigotas grandes, sin embargo, familias grandes en las que no se logra unir la genealogía podrían estudiarse probando diferentes modelos de análisis de ligamiento (recesivo, dominante, ligado al cromosoma X) y buscando

regiones candidatas en cromosomas con loci que tengan altos puntajes de ligamiento. Detectar un locus candidato es indispensable para iniciar la filtración de datos de la secuenciación del exoma. Una vez identificada la región candidata se procede a secuenciar el exoma de al menos dos individuos afectados y—si se cuenta con recursos suficientes—secuenciar a todos los individuos afectados sería ideal. Luego de la secuenciación, es importante revisar mediante algún programa—como IGV browser—la cobertura y efectividad de la secuenciación en la región de interés. Si la cobertura fue igual o mayor a un 80% se puede proceder a filtrar las variantes. Al proceder con la filtración de las variantes se recomienda iniciar con parámetros de alta astringencia y si no se logran identificar variantes candidatas se puede proceder a relajar la astringencia y ser más laxo con los parámetros de filtración. Como se observa en la figura 7, el mejor método de filtración de variantes genéticas para familias de este tipo—en las que se desconoce la causa genética y se presupone un modo de herencia—es recomendable iniciar excluyendo las variantes reportadas previamente en bases de datos de polimorfismos frecuentes en la población (dbSNP132, 1000 genome project, controles internos de pacientes o controles que se hayan secuenciado previamente, entre otros). Luego, filtrar exclusivamente variantes que se encuentren en regiones exónicas o sitios de splicing y excluir las variantes sinónimas. En un primer paso de filtración de alta astringencia es recomendable descartar variantes que sean identificadas como heterocigotas, sin embargo, si se es más estricto y, especialmente, si no se logran detectar variantes de interés, se deberían validar aquellas que variantes poco comunes que se encuentran en heterocigosis en la región de interés para descartar falsos negativos. Además, con algunos datos de las variantes como la cobertura, la cantidad de lecturas realizadas por el SOLiD y si la secuenciación se dio en ambas direcciones (forward y reverse), se pueden descartar algunos falsos positivos. En este último paso, si no se tiene certeza de que una variante es realmente un falso positivo, es recomendable realizar la validación mediante el método de secuenciación de Sanger. Finalmente, los programas que predicen la patogenicidad y conservación de una variante sirven como guía y apoyan los resultados pero no deben ser un criterio estricto para descartar variantes candidatas pues podría tratarse de algún error en los cálculos del programa o que no se dispone de información suficiente para dicha variante en bases de datos.

6.2 Genes candidatos

GPT2 es un enzima que cataliza la reacción reversible que transfiere un grupo alfa-amino del alfa-oxoglutarato a la L-alanina para producir piruvato y L-glutamato y, por lo tanto, tiene un papel importante en la gluconeogénesis y el metabolismo de aminoácidos. Además, GPT2 está

involucrada en la producción de glutamato—un neurotransmisor importante para el cerebro (Peng *et al.* 1991). Los niveles de expresión de la enzima van desde altos en músculo, tejido adiposo, riñón y cerebro hasta bajos en hígado y mamas (Yang *et al.* 2002). Altos niveles de GPT se han relacionado con daño hepático (Sherman 1991), obesidad (Strauss 2000) y enfermedades musculares (Rutledge *et al.* 1985, Lin *et al.* 1999), sin embargo, aún se desconoce el papel funcional de GPT2 en procesos fisiopatológicos (Yang *et al.* 2002).

El glutamato es el aminoácido más abundante en el cerebro y ha sido muy estudiado por sus funciones como neurotransmisor excitatorio en el SNC. Además, se ha relacionado con una serie de procesos característicos del buen funcionamiento cerebral: la cognición, el aprendizaje y la memoria (McEntee y Crook 1993). En algunos estudios se ha demostrado que no sólo el glutamato sino también su precursor metabólico, el alfa-oxoglutarato, se encuentran involucrados en procesos neurológicos importantes. Peng y colaboradores (1991) encontraron que los astrocitos—células gliales más abundantes en el cerebro que llevan a cabo funciones de nutrición, reparación, y mantenimiento del tejido nervioso—tienen un papel importante al suplir de glutamato a neuronas glutaminérgicas no sólo a partir de glutamina sino también a partir de alfa-cetoglutarato (alfa-oxoglutarato) y alanina. Sin embargo, a pesar de que el glutamato producido a partir de alfa-cetoglutarato está disponible para ser utilizado como neurotransmisor, no constituye la principal fuente de glutamato intracelular. Esto porque se ha encontrado que—al menos en rata—tanto la glutamina como el 2-oxoglutarato, son precursores importantes del glutamato consumido en el cerebro (Shank *et al.* 1989).

La mutación G96R en GPT2 podría estar causando alteraciones en la concentración de glutamato o su precursor el alfa-oxoglutarato. Debido a que esta enzima cataliza una reacción reversible, se podría afectar el equilibrio químico en ambas direcciones, es decir, produciéndose un exceso de glutamato o, por el contrario, disminuyéndose los niveles de glutamato. En el primer escenario, un exceso de glutamato podría convertirse en un agente tóxico que podría provocar apoptosis, llevar a la neurodegeneración (Niciu *et al.* 2012) y, por lo tanto, dañar zonas del cerebro involucradas en procesos de aprendizaje o memoria. Por otro lado, una disminución en la síntesis de glutamato podría afectar mecanismos sinápticos importantes como el LTP (long term potentiation) que se ha relacionado con el aprendizaje y la memoria ((Bliss y Collingridge 1993). Una evidencia de que los niveles de glutamato son esenciales para la adecuada homeostasis en el cerebro son las numerosas alteraciones en los niveles de glutamato o en receptores de glutamato que se han encontrado en

pacientes con enfermedades neurodegenerativas como Parkinson (Plaitakis y Shashidharan 2000), Alzheimer (Hynd *et al.* 2004), epilepsia (Meldrum 1994), enfermedad de Huntington (Behrens *et al.* 2002), esclerosis múltiple (Srinivasan *et al.* 2005) y en NS-ID (Motazacker *et al.* 2007). *GRIK2* codifica para un receptor de glutamato que se encontró relacionado a NS-MR y fue el primer gen causante de NS-ID involucrado directamente con el SNC (Motazaker *et al.* 2007).

Si bien es cierto que el componente genético tiene un papel importante en modelar y regular el neurodesarrollo temprano y en la vida postnatal, la estimulación extrínseca es esencial para el modelaje de las conexiones en la vida postnatal. En este sentido, estímulos externos pueden conducir a la alteración del circuito nervioso para recordar experiencias que permitan que el ser humano se adapte mejor a su ambiente y responda mejor a estímulos sensoriales. De esta forma, en muchos de los pacientes con alteraciones neurológicas estos cambios en la conectividad neuronal se ven interrumpidos (Flavell y Greenberg 2008). Bajo esta premisa, podría darse que en individuos con mutaciones en genes indispensables para el adecuado mantenimiento del SNC—como podría ser G96R—que, además, se encuentran sujetos a escasa estimulación temprano en el desarrollo, y en la vida adulta, puedan presentar un fenotipo más severo de la enfermedad que padecen. Es posible que una combinación entre la alteración de los niveles en glutamato en el cerebro, así como, poca estimulación sensorial en los individuos afectados de PZ esté provocando un fenotipo severo de discapacidad cognitiva en esta familia.

6.3 Implicaciones y direcciones futuras

La complejidad de enfermedades heterogéneas como la discapacidad cognitiva no sindrómica exige la elaboración de experimentos que proporcionen evidencias para realizar la correlación genotipo-fenotipo. Por ello, el efecto G96R en GPT2 y su relación con el funcionamiento del SNC requiere de evidencias funcionales. Debido a que la mutación se localiza en el sitio activo de GPT2, se podría comprobar si la mutación encontrada tiene un efecto fenotípico midiendo la actividad enzimática de GPT2 normal y afectada. En un ensayo enzimático *in vitro* mediante la creación de un vector con G96R que se ponga a crecer en un cultivo bacteriano y, al agregar un análogo del sustrato, se podría comprobar si GPT2 realiza la catálisis apropiadamente. Al mismo tiempo, se podría obtener un cultivo bacteriano control, es decir, un vector con la enzima silvestre. Entonces, al agregar el sustrato al cultivo con la enzima mutada, la producción de glutamato debería verse disminuida o aumentada con respecto a la producción de glutamato de la enzima silvestre. Por otro lado, es posible examinar la expresión de GPT2 durante el desarrollo embrionario de un organismo

modelo como el ratón, para observar si GPT2 se expresa en zonas del cerebro—como el hipocampo y la corteza—involucradas en procesos sinápticos como el LTP. Otras pruebas que podría realizarse para estudiar de manera indirecta el efecto de G96R en zonas específicas del cerebro involucran la observación de anomalías morfológicas en el cerebro mediante resonancias magnéticas de los individuos afectados, medición de los distintos aminoácidos—especialmente el glutamato y el alfa-oxoglutarato—en el líquido cefalorraquídeo, sangre u orina; y la creación de un ratón knockout para GPT2 que permita estudiar efecto de esta mutación sobre el comportamiento, memoria y habilidades cognitivas complejas.

Dependiendo del efecto de G96R en GPT2 podría desarrollarse una terapia, ya que, GPT2 en condiciones normales, podría funcionar como un barredor del exceso de glutamato en el cerebro. El efecto neuroprotector de la enzima en condiciones normales podría ser de interés terapéutico para todas aquellas enfermedades en las que un exceso de glutamato esté dañando al cerebro (Teichberg *et al.* 2009). Por otro lado, si GPT2 fuera incapaz de sintetizar glutamato, podría pensarse en administrar bajas dosis de glutamato temprano en el desarrollo, a los pacientes con G96R, para lograr recuperar la homeostasis en los fluidos cerebrales.

7. CONCLUSIONES

1. La combinación de técnicas moleculares como la secuenciación del exoma y carreras de homocigosidad para detectar genes causantes de discapacidad cognitiva no sindrómica con herencia autosómica recesiva es una aproximación eficaz en familias con más de un individuo afectado y cuyos padres están emparentados.
2. Se encontró la mutación G96R en GPT2, candidato a ser causante de NS-ID en la familia PZ. Esta mutación altera el sitio activo de la enzima, impidiendo que el sustrato se una adecuadamente y, por lo tanto, afectando la función catalítica.
3. La mutación en GPT2 podría estar causando una acumulación o un descenso en los niveles de glutamato o su precursor el 2-oxoglutarato.
4. Para comprobar que GPT2 es un gen nuevo asociado a retardo mental se requiere de pruebas funcionales, por ejemplo, se puede realizar un ensayo enzimático en el que se compruebe el efecto de la mutación en la actividad enzimática y su relación con la producción de glutamato.
5. Las familias a las que no se les encontró la causa genética de discapacidad cognitiva no sindrómica podría explicarse porque la mutación se encuentre en la región no codificante

del genoma (intrones, regiones intergénicas, regiones no traducibles, entre otras), en el 20% del exoma que no fue cubierto por el NGS o que se trate de una mutación sinónima.

8. REFERENCIAS

- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat Methods*. 7: 248-249.
- Bamshad, M. J., Ng, S. B., Bigham, A. W., Tabor, H. K., Emond, M. J., Nickerson, D. A., & Shendure, J. 2011. Exome sequencing as a tool for Mendelian disease gene discovery. *Nature Rev Genet*. 12: 745-755.
- Behrens, P. F., Franz, P., Woodman, B., Lindenberg, K. S., & Landwehrmeyer, G. B. 2002. Impaired glutamate transport and glutamate–glutamine cycling: downstream effects of the Huntington mutation. *Brain*, 125: 1908-1922.
- Bolino, A., M. Muglia, F. Conforti, E. Leguern, M. Salih, D. Georgiou, K. Christodoulou, I. Hausmanowa, P. Mandich, A. Schenone, A. Gambardella, F. Bono, A. Quattrone, M. Devoto & A. Monaco. 2000. Charcot-Marie-Tooth type 4B is caused by mutations in the gene encoding myotubularin-related protein-2. *Nature genetics*. 25: 17-19.
- Bliss, T. V. & G.L. Collingridg. 1993. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature*. 361: 31-39.
- Chavarría, G., A. Reis & J. Azofeifa. 2002. Determinación indirecta, mediante marcadores de ADN, del estado de portadoras de distrofia muscular de Duchenne (DMD) en una familia costarricense. *Acta Pediátrica Costarricense*. 16: 32-38.
- Chavarría, G., B. Rautenstrauss & J. Azofeifa. 2004. Glaucoma in Costa Rica: Initial approaches. *Rev. Biol. Trop*. 52: 507-520.
- Chelly, J., M. Khelifaoui, F. Francis, B. Cherif & T. Bienvenu. 2006. Genetics and pathophysiology of mental retardation. *Eur. J. Hum. Genet*. 14: 701-713.
- DePristo M, Banks E, Poplin R, Garimella K, Maguire J, Hartl C, Philippakis A, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell T, Kernysky A, Sivachenko A, Cibulskis K, Gabriel S, Altshuler D and Daly, M. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*. 43:491-498.
- DeVries, B., R. Winter, A. Schinzel & C. van Ravenswaaij. 2003. Telomeres: a diagnosis at the end of the chromosomes. *J. Med. Genet*. 40: 385-398.
- DeVries, Bert, R. Pfundt, M. Leisink, D.A. Koolen, L. Vissers, I. Janssen, S. Reijmersdal, W. Nillesen, E. Huys, N. Leeuw, D. Smeets, E. Sijm, T. Feuth, C. van Ravenswaaij, A.G. van Kessel, E. Schoenmakers, H. Brunner & J. Veltman. 2005. Diagnostic genome profiling in mental retardation. *Am. J. Hum. Genet*. 77: 606-616.

- Flint, J., A. Wilkie, V.J. Buckle, R.M. Winter, A.J. Holland & H.E. Mcdermid. 1995. The detection of subtelomeric chromosomal rearrangements in idiopathic mental retardation. *Nature genetics*. 9: 132-140.
- Grady, D. L., Ratliff, R. L., Robinson, D. L., McCanlies, E. C., Meyne, J., & Moyzis, R. K. 1992. Highly conserved repetitive DNA sequences are present at human centromeres. *P Ntl Acad Sci USA*. 89: 1695-1699.
- Gudbjartsson, D. F., Jonasson, K., Frigge, M. L., & Kong, A. 2000. Allegro, a new computer program for multipoint linkage analysis. *Nature genetics*. 25: 12.
- Hehir-Kwa, J. Y., N. Wieskamp, C. Webber, R. Pfundt, H.G. Brunner, C. Gilissen, B. de Vries, C. Pontig & J. Veltman. 2010. Accurate distinction of pathogenic from benign CNVs in mental retardation. *PLoS Comput. Biol.* 6: 1-9.
- Hernández, R & Rodríguez, S (1987) Manual Operativo para la Evaluación y Estimulación del Crecimiento del Niño. San José, Costa Rica: EUNED.
- Higgins, J. J., J. Pucilowska, R. Lombardi & J. P. Rooney. 2005. A mutation in a novel ATP-dependent Lon protease gene in a kindred with mild mental retardation. 63: 1927-1931.
- Hildebrandt, F., S.F. Heeringa, F. Rüschen-dorf, M. Attanasio, G. Nürnberg, C. Becker, D.D. Schoeb, S. Allen, V. Vega, E. Wise, H. Harville, R. Lyons, J. Washburn, J. Macdonald, P. Nürnberg & E. Otto. 2009. A systematic approach to mapping recessive disease genes in individuals from outbred populations. *PLoS Genet.* 5: 1-10.
- Hoffman, K. & TH Lindner. 2005. easyLINKAGE plus -- Automated linkage analysis using large-scale SNP data. *Bioinformatics*. 21: 3565-3567.
- Hu, W.H., J.S. Pendergast, X.M. Mo, R. Brambilla, V. Bracchi-Ricard, F. Li, W. Winston, B. Blits, L. He, S. Schaal & J. Bethea. 2005. NIBP, a novel NIK and IKK(beta)-binding protein that enhances NF-(kappa)B activation. *J. Biol. Chem.* 280: 29233-29241.
- Hynd, M. R., Scott, H. L., & Dodd, P. R. (2004). Glutamate-mediated excitotoxicity and neurodegeneration in Alzheimer's disease. *Neurochemistry international*, 45(5), 583-595.
- Kaufman, L., M. Ayub & J.B. Vincent. 2010. The genetic basis of non-syndromic intellectual disability: a review. *J. Neurodevelop. Disord.* 2: 182-209.
- Knight, S., R. Regan, A. Nicod, S. Horsley, L. Kearney, T. Homfray, R. Winter, M. Robin, P. Bolton & J. Flint.. 1999. Subtle chromosomal rearrangements in children with unexplained mental retardation. *Lancet*. 354: 1676-1681.
- Koolen, D., W. Nillesen, M. Versteeg, G. Merckx, N. Knoers, M. Kets, S. Vermeer, C. van Ravenswaaij, C. de Kovel, H. Brunner, D. Smeets, B. de Vries, E. Sijm, & M. Smeets. 2004. Screening for subtelomeric rearrangements in 210 patients with unexplained mental retardation using multiplex ligation dependent probe amplification (MLPA). *J. Med. Genet.* 41: 892-899.
- Lau, A., C. Brown, M. Pena-Herrera, S. Langlois, D. Kalousek y W. Robinson. 1997. Skewed X-chromosome inactivation is common in fetuses or newborns associated with confined placental mosaicism. *Am J Hum Genet.* 61: 1353-1361.

- Launay, G., & Simonson, T. 2008. Homology modelling of protein-protein complexes: a simple method and its possibilities and limitations. *BMC bioinformatics*, 9:427-433.
- Leal, A., B. Morera, G. Del Valle, D. Heuss, C. Kayser, M. Berghoff, R. Villegas, E. Hernández, M. Méndez, H.C. Hennies, B. Neundörfer, R. Barrantes, A. Reis & B. Rautenstrauss. 2001. A second locus for an axonal form of autosomal recessive Charcot-Marie-Tooth disease maps to chromosome 19q13.3. *Am. J. Hum. Genet.* 68: 269-274.
- Lee, D. & M. Cox. 2008. Lehninger principles of biochemistry. Wisconsin, Freeman.
- Leonard, H. & X. Wen. 2002. The epidemiology of mental retardation: Challenges and opportunities in the new millennium. *Ment. Retard. Dev. D. R.* 8: 117-134.
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin & 1000 Genome Project Data Processing Subgroup. 2009. The Sequence alignment/map (SAM) format and SAMtools. *Bioinformatics*, 25: 2078-2079.
- McEntee, W. J. & T.H. Crook. 1993. Glutamate: its role in learning, memory, and the aging brain. *Psychopharmacol.* 111: 391-401.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M & DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20:1297-303.
- McLaren, J. & S.E. Bryson. 1987. Review of recent epidemiological studies of mental retardation: prevalence, associated disorders, and etiology. *Am. J. Ment. Retard.* 92: 243-254.
- Meldrum, B. S. 1994. The role of glutamate in epilepsy and other CNS disorders. *Neurology.* 44: 14-23.
- Meléndez, M. 2004. Importancia de la genealogía aplicada a estudios genéticos en Costa Rica. *Rev. Biol. Trop.* 52: 423-450.
- Molinari, F., M. Rio, V. Meskenaitė, F. Encha-Razavi, J. Augé, D. Bacq, S. Briault, M. Vekemans, A. Munnich, T. Attié-Bitach, P. Sonderegger & L. Colleaux. 2002. Truncating Neurotrypsin Mutation in Autosomal Recessive Nonsyndromic Mental Retardation. *Science.* 298: 1779-1781.
- Montenegro, G., E. Powell, J. Huang, F. Speziani, Y.J.K Edwards, G. Beecham, W. Hulme, C. Siskind, J. Vance, M. Shy & S. Züchner. 2011. Exome sequencing allows for rapid gene identification in a Charcot-Marie-Tooth family. *Ann. Neurol.* 69: 464-470.
- Motazacker, M. M., B.R. Rost, T. Hucho, M. Garshasbi, K. Kahrizi, R. Ullmann, S. Abedini, S. Nich, S. Amini, C. Goswami, A. Tzschach, L. Jensen, D. Schmitz, H. Ropers, H. Najmabadi & A. Kuss. 2007. A defect in the ionotropic glutamate receptor 6 gene (GRIK2) is associated with autosomal recessive mental retardation. *Am. J. Hum. Genet.* 81: 792-798.
- Najmabadi, H., H. Hu, M. Garshasbi, T. Zemojtel, S. Abeddini, W. Chen et al. 2011. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature.* 478: 57-63.

- Niciu, M. J., Kelmendi, B., & Sanacora, G. 2012. Overview of glutamatergic neurotransmission in the nervous system. *Pharmacol Biochem Be.* 100: 656-664.
- Ng CP & Henikoff S. 2003. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Research* 31: 3812–3814.
- Noor, A., C. Windpassinger, M. Patel, B. Stachowiak, A. Mikhailov, M. Azam, M. Irfan, Z.K. Siddiqui, F. Nacem, A.D. Paterson, M. Lutfullah, J.B. Vicent & M. Ayub. 2008. CC2D2A, encoding a coiled-coil and C2 domain protein, causes autosomal-recessive mental retardation with retinitis pigmentosa. *Am. J. Hum. Genet.* 82: 1011-1018.
- Noor, A., C. Windpassinger, I. Vitcu, M. Orlic, M. Arshad, M. Khalid, M. Nasir, M. Ayub, B. Alman & J.B. Vicent.. 2009. Oligodontia is caused by mutation in LTBP3, the gene encoding latent TGF- β binding protein 3. *Am. J. Hum. Genet.* 84: 345-349.
- Peng, L. A., Schousboe, A., and Hertz, L. (1991). Utilization of α -ketoglutarate as a precursor for transmitter glutamate in cultured cerebellar granule cells. *Neurochem. Res.* 16: 29–34.
- Philippe, O., M. Rio, A. Carioux, J.M. Plaza, P. Guigue, F. Molinari, N. Boddaert, C. Bole-Feysot, P. Nitschke, A. Smahi, A. Munnich & L. Colleaux. 2009. Combination of linkage mapping and microarray expression analysis identifies NF-kB signaling defect as a cause of autosomal-recessive mental retardation. *Am. J. Hum. Genet.* 85: 903-908.
- Pieretti, M., F. Zhang, Y. Fu, S. Warreb, B. Oostra, C. Caskey & D. Nelson. 1991. Absence of expression of the FMR1 gene in fragile X syndrome. *Cell*, 66: 817-822.
- Plaitakis, A. & P. Shashidharan. 2000. Glutamate transport and metabolism in dopaminergic neurons of substantia nigra: implications for the pathogenesis of Parkinson's disease. *J. neurol.* 247: 25-35.
- Porreca, G. J., Zhang, K., Li, J. B., Xie, B., Austin, D., Vassallo, S. L., & Shendure, J. 2007. Multiplex amplification of large sets of human exons. *Nature methods*, 4: 931-936.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, MAR Ferreira, D. Bender, J. Maller, P. Sklar, PIW de Baker, MJ Daly & PC Sham. 2007. *Am J Hum Genet.* 81: 559-575.
- Purcell, S. PLINK v.1.07. 10 octubre 2009. Consultado en: <http://pngu.mgh.harvard.edu/purcell/plink/>
- Rauch, A., J. Hoyer, S. Guth, C. Zweier, C. Kraus, C. Becker, M. Zenker, U. Hu, C. Thiel & F. Ru. 2006. Diagnostic yield of various genetic approaches in patients with unexplained developmental delay or mental retardation. *Am. J. Med. Genet.* 140: 2063-2074.
- Ropers, H.H. & B.C.J. Hamel. 2005. X-linked mental retardation. *Nat. Genet.* 6: 46-57.
- Ropers, H. H. 2010. Genetics of early onset cognitive impairment. *Annu. Rev. Genomics Hum. Genet.* 11: 161-87.
- Rutledge, J., Andersen, J., Fink, C. W., Cook, J. & Strickland, A. 1985. Persistent hypertransaminasemia as the presenting finding of childhood muscle disease. *Clin. Pediatr.* 24: 500–503.

- Sali, A. y T.L. Blundell. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234: 779-815.
- Schwarz JM, Rödelsperger C, Schuelke M, Seelow D. 2010. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods.* 7: 575-576.
- Sherman, S. L., P.A. Jacobs, N.E. Morton, U. Froster-Iskenius, P.N. Howard-Peebles, K.B. Nielsen, M. Partington, G. Sutherland, G. Turner & M. Watson. 1985. Further segregation analysis of the fragile X syndrome with special reference to transmitting males. *Hum. Genet.* 69: 289-299.
- Srinivasan, R., Sailasuta, N., Hurd, R., Nelson, S., & Pelletier, D. 2005. Evidence of elevated glutamate in multiple sclerosis using magnetic resonance spectroscopy at 3 T. *Brain.* 128: 1016-1025.
- Teichberg, V. I., Cohen-Kashi-Malina, K., Cooper, I., & Zlotnik, A. 2009. Homeostasis of glutamate in brain fluids: an accelerated brain-to-blood efflux of excess glutamate is produced by blood glutamate scavenging and offers protection from neuropathologies. *Neuroscience.* 158: 301-308.
- Telatar, M., Z. Wang, S. Castellvi-Bel, L.Q. Tai, S. Sheikhavandi, J.R. Regueiro, O. Porras & R.A. Gatti. 1998. A model for ATM heterozygote identification in a large population: four founder-effect ATM mutations identify most of Costa Rican patients with Ataxia Telangiectasia. *Mol. Genet. Metab.* 64: 36-43.
- Thiele, H. & P. Nürnberg. 2004. HaploPainter: a tool for drawing pedigrees with complex haplotypes. *Bioinformatics.* 21:1730-1732.
- Senderek, J., C. Bergmann, C. Stendel, J. Kirfel, N. Verpoorten, P. De Jonghe, V. Timmerman, R. Chrast, M. Verheijen, G. Lemke, E. Battaloglu, Y. Parman, S. Erdem, E. Tan, H. Topaloglu, A. Hahn, W. Müller-Felber, N. Rizzuto, G.M. Fabrizi, M. Stuhmann, S. Rudnik-Schöneborn, S. Züchner, S.J. Michael, E. Buchheim, V. Straub, J. Klepper, K. Huehne, B. Rautenstrauss, R. Büttner, E. Nelis & K. Zerres. 2003. Mutations in a gene encoding a novel SH3/TPR domain protein cause autosomal recessive Charcot-Marie-Tooth Type 4C neuropathy. *Am. J. Hum. Genet.* 73: 1106-1119.
- Shank, R. P., Baldy, W. J., & Ash, C. W. 1989. Glutamine and 2-oxoglutarate as metabolic precursors of the transmitter pools of glutamate and GABA: correlation of regional uptake by rat brain synaptosomes. *Neurochem Res.* 14: 371-376.
- Sherman, K. E. 1991. Alanine aminotransferase in clinical practice. A review. *Arch. Intern. Med.* 151: 260-265.
- Strauss, R. S., Barlow, S. E., and Dietz, W. H. 2000. Prevalence of abnormal serum aminotransferase values in overweight and obese adolescents. *J. Pediatr.* 136: 727-733.
- Sulonen, A. M., Ellonen, P., Almusa, H., Lepistö, M., Eldfors, S., Hannula, S., ... & Saarela, J. 2011. Comparison of solution-based exome capture methods for next generation sequencing. *Genome biology,* 12: 1-17.

- Verpy, E., S. Masmoudi, I. Zwaenepoel, M. Leibovici, T.P. Hutchin, I. Del Castillo, S. Nouaille, S. Nouaille, S. Blanchard, S. Lainé, J.L. Popot, F. Moreno, R.F. Mueller & C. Petit. 2001. Mutations in a new gene encoding a protein of the hair bundle cause non-syndromic deafness at the DFNB16 locus. *Nature genetics*. 29: 345-349.
- Vissers, L. E. L. M., J. Ligt, C. Gilissen, I. Janssen, M. Stehouwer, P. Vries, P. Arts, N. Wieskamp, M. Rosario, B. van Bon, A. Hoischen, B. de Vries, H. Brunner & J. Veltman. 2010. A de novo paradigm for mental retardation. *Nat. Genet.* 42: 1109-1112.
- Wang, K., Li, M., & Hakonarson, H. 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic acids research*, 38: 164-164.
- Webber, C., J. Hehir-Kwa, D.Q. Nguyen, B. de Vries, J. Veltman & C. Ponting. 2009. Forging links between human mental retardation associated CNVs and mouse gene knockout models. *PLoS Genet.* 5: 1-12.
- Wechsler, D. 1997. Wechsler Adult Intelligence Scale—3rd Edition (WAIS-3®) San Antonio, Texas: Harcourt Assessment.
- Wechsler, D. 2003. Wechsler Intelligence Scale for Children—4th Edition (WISC-IV®). San Antonio, Texas: Harcourt Assessment.
- Yang, R. Z., G. Blaileanu, B.C. Hansen, A.R. Shuldiner & D.W. Gong. 2002. cDNA cloning, genomic structure, chromosomal mapping, and functional expression of a novel human alanine aminotransferase. *Genomics*, 79: 445-450.

9. ANEXOS

Escala del desarrollo integral del niño (EDIN)

CAJA COSTARRICENSE DE SEGURO SOCIAL

CENTRO: _____ NOMBRE: _____ Nº EXPEDIENTE: _____
 SERVICIO: _____ FECHA NACIMIENTO: _____
 No. CAMA: _____ SEXO: _____ FECHA: _____

NIÑOS MENORES DE UN AÑO

FASE	AREA						EDAD DEL NIÑO
	Motora gruesa	Motora fina	Cognoscitiva	Lenguaje	Socio-afectiva	Habitos de Salud y Nutrición	
	CONDUCTA						
MESES	Se pone de pie apoyándose	Mete y saca cubos de una taza	Devuelve la bola	Dice algunas palabras cortas (dos por lo menos)	Demuestra afecto	Bebe en taza derramando poco	
11-12							
10-11	Camina sostenido de una mano	Golpea un cubo contra otro	Retiene dos cubos en la mano al darle otro	Al escuchar la palabra adiós, hace adiós con la mano	Se enoja cuando tratan de quitarle un juguete	Mete la mano en el plato de comida y se la lleva a la boca	
9-10	Gatea	Coge una cuerda con dos dedos	Jala el pañal para alcanzar un objeto	Responde a órdenes simples acompañadas de gestos	Repite actividad si se celebra su actuación	Intenta usar cuchara y plato	
8-9	Permanece sentado sin apoyo	Toma un dulce con el pulgar y el índice	Quita el pañal y toma la bola escondida	Suspende actividad cuando se le dice no	Se observa en el espejo lo frota y lo chupa	Demuestra preferencia o rechazo por los alimentos	
7-8	Se para brevemente sostenido de manos	Toma gomita entre pulgar, índice y medio	Busca objetos perdidos frente a él	Usa gestos para comunicar sus deseos	Cambia de actitud cuando llegan extraños	Mastica los alimentos	
6-7	Se sienta sin apoyo por pocos segundos	Toma gomita con pulgar y otros dedos	Recupera chilindrín que se le cae de la mano	Combina sílabas ma-ma, ta-ta	Disfruta jugando al escondido	Se lleva a la boca sin ayuda una galleta, pan o tortilla	
5-6	Da vuelta de boca abajo a boca arriba	Coge el objeto que se le pone frente a él	Sigue con la mirada momentáneamente los objetos que caen	Emite sonidos mientras juega solo	Explora la cara de la madre con la vista y manos	Se alimenta con leche exclusivamente	
4-5	Se sienta con apoyo, cabeza firme	Intenta coger los objetos sin lograrlo	Observa un objeto antes de meterlo a la boca	Emite sonidos repetidamente "p", "m", "b", "g", "o"	Se ríe a carcajadas cuando juega con otra persona	Se alimenta con leche exclusivamente	
3-4	Se vuelve parcialmente	Mantiene las manos abiertas	Sigue con la mirada la bola que se desplaza en movimientos circulares	Emite sonidos vocales a,e,u	Mira sus manos, juega con ellas y se las lleva a la boca	Se alimenta con leche exclusivamente	
2-3	Al sentarlo mantiene la cabeza erecta unos segundos	Rasca los objetos que toca	Mueve la cabeza al sonar la campanilla	Emite sonidos guturales (ggggu)	Utiliza sonrisa social	Se alimenta con leche exclusivamente	
1-2	Sostiene la cabeza por breves segundos	Aprieta fuerte un objeto y luego lo deja caer	Sigue objetos con la mirada	Responde a los sonidos cambiando de actitud	Se calma al alzarlo y hablarle	Se alimenta con leche exclusivamente	
0-1	Boca abajo mantiene la cabeza de lado. Moro Marcha autom.	Mantiene manos fuertemente cerradas. Succión Prens. Palmar Prens. Palmar	Fija la mirada en el chilindrín	Busca sonido	Muestra rostro poco expresivo	Se alimenta con leche exclusivamente	

DESARROLLO PSICOMOTOR DEL NIÑO

Cód. 4-70-03-3035

NIÑOS DE 1 A 6 AÑOS

FASE	AREA						EDAD DEL NIÑO
	Motora gruesa	Motora fina	Cognoscitiva	Lenguaje	Socio-afectiva	Habitos de Salud y Nutrición	
	CONDUCTA						
5-6	Salta abriendo y cerrando las piernas	Apaña con una mano una bolsita de arroz	Puede contar de 7 a 10 objetos	Usa los verbos en pasado, presente y futuro	Participa en actividades de grupo	No se orina en la cama por las noches	
4-5	Salta hacia atrás por imitación	Toca con el pulgar los demás dedos de la mano	Dibuja una figura humana con 4 partes	Empieza verbos en pasado	Gusta de juegos competitivos	Puede vestirse y desvestirse solo	
3-4	Mantiene el equilibrio en un pie por 8 segundos sin ayuda	Construye puentes con 3 cubos	Dice si un objeto es blando o duro	Utiliza oraciones de 5 a 7 palabras	Dice su sexo	Se lava y seca la cara solo.	
2,5-3	Se mantiene de pie con los talones juntos	Ensarta cuentas en un cordón	Coloca un cubo encima de un objeto y debajo de él.	Usa algunos plurales	Dice su nombre	Usa solo el inodoro o letrina cuando lo necesita	
2-2,5	Se para en un solo pie con ayuda	Construye una torre de 4 a 6 cubos	Señala tres partes de su cuerpo	Construye frases	Comparte juegos y juguetes, reconoce su nombre cuando lo escucha	Avisa para defecar u orinar	
1,5-2	Se baja de un silla sin ayuda	Construye torre de 2 a 3 cubos por imitación	Señala una parte de su cuerpo cuando le pregunta	Sigue dos órdenes consecutivas	Hace berrinche cuando no se le da lo que quiere	Indica en forma verbal o no verbal que pañal está sucio.	
1-1,5	Da unos pasos solo.	Sostiene dos cubos pequeños en una mano	Recupera objetos escondidos bajo su pañal o taza	Dice 2 a 6 palabras	Juega solo	Trata de comer por si solo.	

DESARROLLO PSICOMOTOR DEL NIÑO

9.2 Fórmula de consentimiento informado



UNIVERSIDAD DE COSTA RICA
VICERRECTORÍA DE INVESTIGACIÓN
COMITÉ ÉTICO CIENTÍFICO
 Teléfonos:(506) 2207-5006 Telefax: (506) 2224-9367

Escuela de Biología
 Sección de Genética y Biotecnología

FÓRMULA DE CONSENTIMIENTO INFORMADO (Para ser sujeto de investigación)

Búsqueda de genes responsables de discapacidad cognitiva con herencia autosómica recesiva

Código de proyecto: 111-

Nombre del Investigador Principal: Dr. Alejandro Leal Esquivel

Nombre de la Investigadora Asociada: Dra. Gabriela Chavarría Soley

Nombre del participante: _____ Código: _____

A. PROPÓSITO DEL PROYECTO

El objetivo del proyecto es buscar alguna mutación (cambio genético) en pacientes que presentan retardo mental con herencia autosómica recesiva, o sea afectados con una discapacidad cognitiva que afecte tanto a hombres como a mujeres que son hijos de padres sanos, sin que haya otros parientes en la familia que padezcan la enfermedad. La idea es tratar de encontrar la causa genética de estas enfermedades.

Se requiere analizar a todos los miembros de la familia porque si se encuentra un cambio genético que podría ser responsable de la enfermedad, hay que comparar a los afectados con los sanos de la familia, para ver si la presencia de la mutación concuerda con las características clínicas de los miembros de la familia.

No se puede garantizar que se encontrarán las mutaciones responsables, ni tampoco se puede asegurar que en un futuro los pacientes van a recibir un tratamiento que les cure la enfermedad luego de realizado este diagnóstico. No obstante, esta investigación podría dar las bases para desarrollar futuras terapias y permitirá el diagnóstico de personas que no están enfermas pero tienen un gen mutado (portadores). El investigador principal de este proyecto es el Dr. Alejandro Leal Esquivel, genetista, y la profesora asociada es la Dra. Gabriela Chavarría Soley, genetista.

B. ¿QUÉ SE HARÁ?

Si acepta participar en este estudio, se le realizará lo siguiente:

- Se le hará una entrevista con datos generales del paciente, de la historia familiar, y de datos clínicos. Esta entrevista será de carácter confidencial y la información suministrada se utilizará sólo para los fines de este estudio.
- Se hará un análisis clínico de los pacientes y de los familiares.

epyrature

- c) Se tomará una muestra de sangre de 10 mL (cantidad parecida a la que se toma para un análisis de sangre en un laboratorio clínico), de sangre periférica (a extraerse del brazo).

C. RIESGOS

La participación en este estudio puede significar molestia por lo siguiente: el tomar la muestra de sangre puede causar ardor por la punción, y podría causar un pequeño moretón.

D. BENEFICIOS:

Como resultado de su participación en este estudio, el beneficio que obtendrá será saber la mutación causante de la discapacidad cognitiva que sufre. Esta información podría ser utilizada para detectar portadores en el resto de la familia, si esto es deseado. También, identificando la mutación causante, podrían hacerse, en un futuro, sugerencias terapéuticas que beneficien al paciente o a otros pacientes. No obstante, no se obtendrá un beneficio directo de tratamiento a raíz del resultado del diagnóstico. Tampoco se puede asegurar que se localice el cambio genético que está causando la enfermedad.

- E. Antes de dar su autorización para este estudio usted debe haber hablado con el Dr. Alejandro Leal o con alguno de los investigadores sobre este estudio y ellos deben haber contestado satisfactoriamente todas sus preguntas. Si quisiera más información más adelante, puedo obtenerla llamando a investigador principal al teléfono 2207-4043 (de 8 a.m. a 12 m. o de 1 a 5 p.m.). Cualquier consulta adicional puede comunicarse a la Vicerrectoría de Investigación de la Universidad de Costa Rica a los teléfonos 2207-4201 ó 2207-5839, de lunes a viernes de 8 a.m. a 5 p.m.
- F. Recibirá una copia de esta fórmula firmada para su uso personal.
- G. Su participación en este estudio es voluntaria. Tiene derecho de negarse a participar o a discontinuar su participación en cualquier momento, sin que esta decisión afecte la calidad de la atención médica (o de otra índole) que requiera.
- H. Su participación en este estudio es confidencial, los resultados podrían aparecer en una publicación científica o ser divulgados en una reunión científica pero de una manera anónima.
- I. No perderá ningún derecho legal por firmar este documento.
- J. ¿Autoriza que en caso de necesidad, la información y las muestras obtenidas para esta investigación podrían transferirse a otros investigadores bajo el Acuerdo de Transferencia de Material Biológico (MTA), siempre que se respete totalmente su confidencialidad? (Independientemente de su respuesta, Ud. podrá participar en este estudio).

Sí _____ No _____



- K. ¿Autoriza el uso del material genético para otros estudios relacionados con la enfermedad, siempre que se proteja su confidencialidad? (Independientemente de su respuesta, Ud. podrá participar en este estudio).

Sí _____ No _____

CONSENTIMIENTO

He leído o se me ha leído, toda la información descrita en esta fórmula, antes de firmarla. Se me ha brindado la oportunidad de hacer preguntas y éstas han sido contestadas en forma adecuada. Por lo tanto, accedo a participar como sujeto de investigación en este estudio

Nombre, cédula y firma del sujeto (niños mayores de 12 años y adultos) fecha

Nombre, cédula y firma del testigo fecha

Nombre, cédula y firma del Investigador que solicita el consentimiento fecha

Nombre, cédula y firma del padre/madre/representante legal (menores de edad) fecha



9.3 Datos clínicos que conforman el expediente de los pacientes.

Perspectiva clínica

Nombre

Edad

Embarazo (i.e. tipo de embarazo, complicaciones)

Desarrollo del niño

-Edad de cuando se sentó por primera vez:

-Edad de la primera vez que caminó:

-Edad de la primera vez que habló:

Defectos orgánicos (i.e. espina bífida, defectos en el corazón)

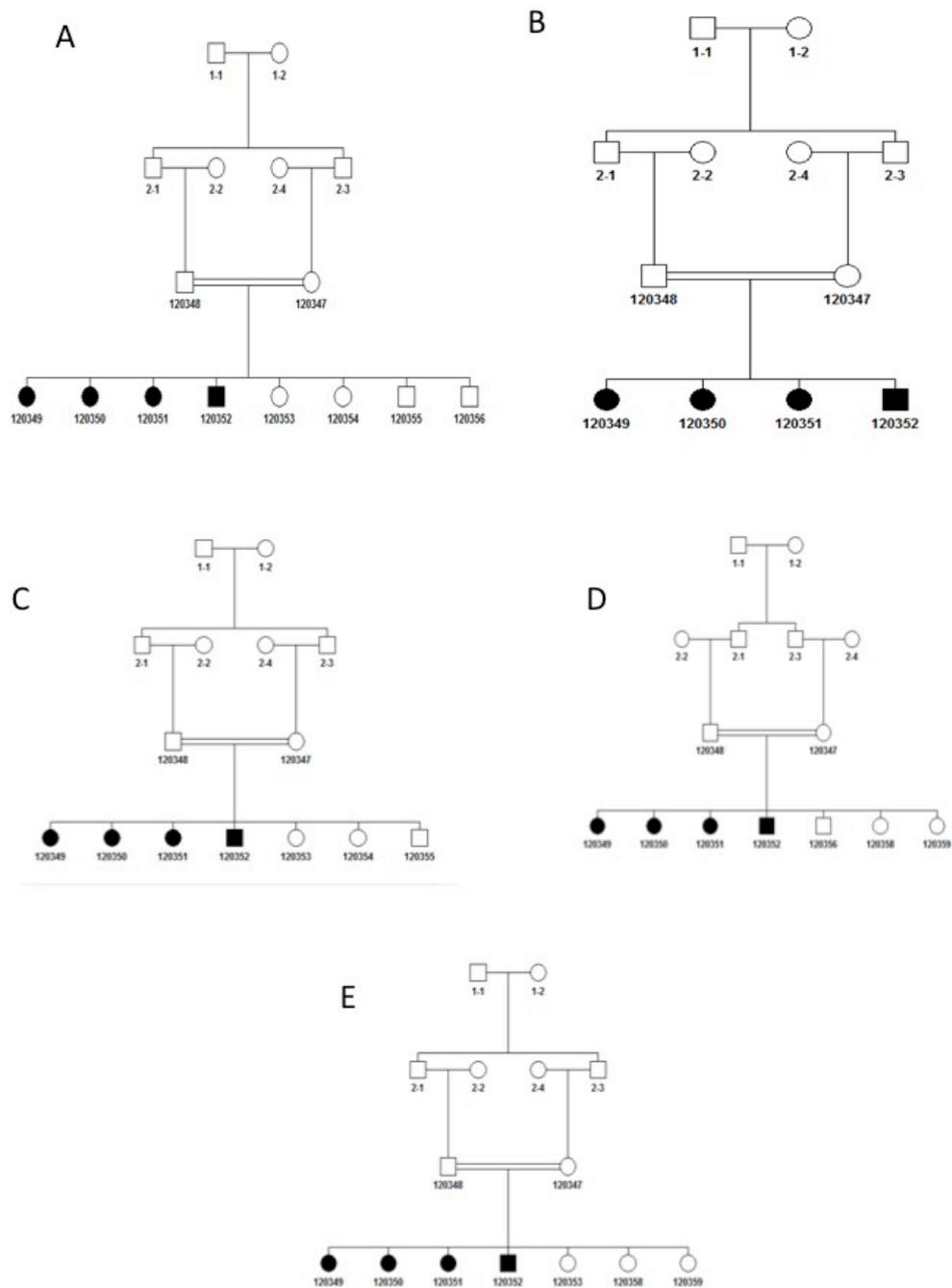
Habilidades (i.e. caminar, hablar, jugar con otros niños, comportamientos agresivos)

Exámenes realizados (MIT, EEG, IQ)

Complicaciones u hospitalizaciones

Tipo de escuela a la que asiste:

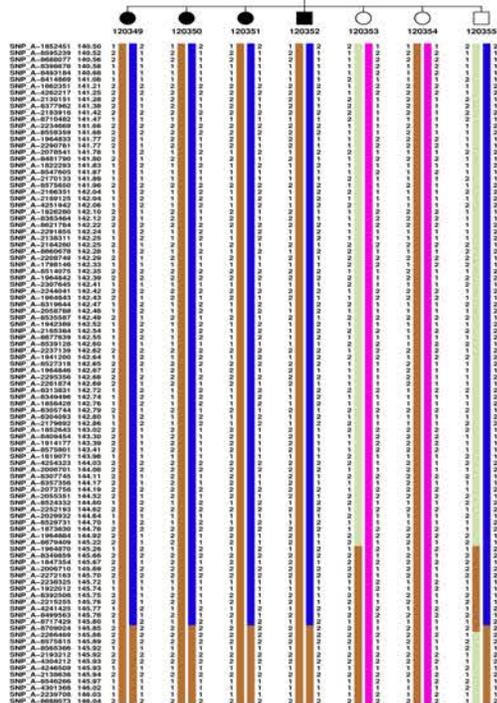
9.4 Genealogía de la familia PZ utilizados para el análisis de ligamiento.



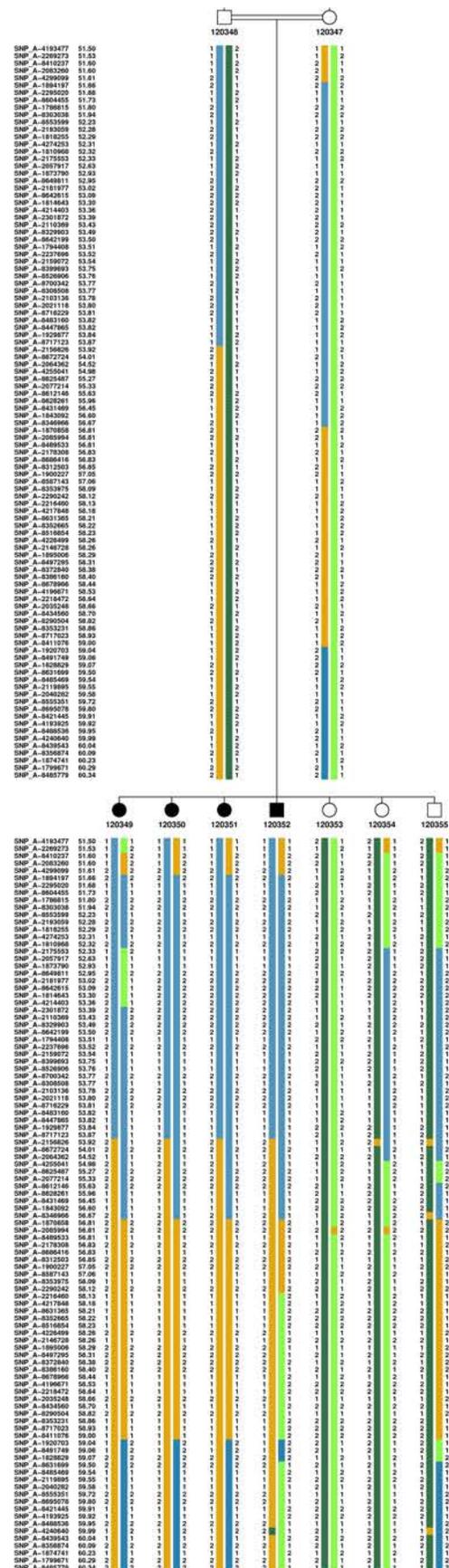
9.5 Haplotipos de la familia PZ: para poder realizar el análisis de ligamiento se tuvo que partir la familia y aquí se presentan los haplotipos de las regiones que presentaron los puntajes LOD más altos.

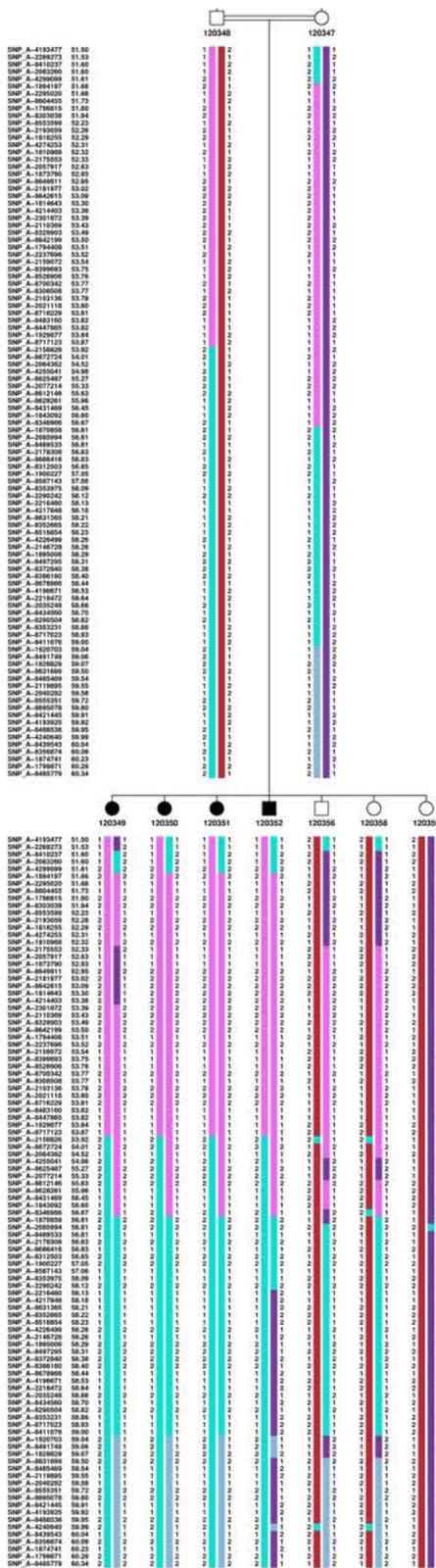
A) Generalología C, Cromosoma 2

SNP_A_1802451	140.50
SNP_A_8992279	140.52
SNP_A_8396676	140.56
SNP_A_8491194	140.56
SNP_A_8414898	141.08
SNP_A_1802241	141.21
SNP_A_8202177	141.25
SNP_A_2130151	141.26
SNP_A_8077862	141.26
SNP_A_2162818	141.42
SNP_A_8715482	141.47
SNP_A_2234669	141.66
SNP_A_8033838	141.66
SNP_A_2297641	141.77
SNP_A_2292761	141.78
SNP_A_8481790	141.80
SNP_A_8841700	141.87
SNP_A_2178133	141.88
SNP_A_2166311	142.04
SNP_A_2168129	142.04
SNP_A_4351842	142.08
SNP_A_1828246	142.15
SNP_A_8621794	142.22
SNP_A_2291876	142.24
SNP_A_2136311	142.25
SNP_A_8098787	142.25
SNP_A_2200749	142.29
SNP_A_1798146	142.31
SNP_A_8514075	142.35
SNP_A_1968462	142.36
SNP_A_2207645	142.41
SNP_A_2244641	142.42
SNP_A_8318644	142.47
SNP_A_2204776	142.48
SNP_A_8032687	142.49
SNP_A_1842388	142.52
SNP_A_8677839	142.55
SNP_A_1841206	142.64
SNP_A_2227139	142.62
SNP_A_1941206	142.64
SNP_A_8027718	142.65
SNP_A_1964846	142.67
SNP_A_2281874	142.69
SNP_A_8511871	142.72
SNP_A_8348498	142.74
SNP_A_1858428	142.76
SNP_A_8350492	142.80
SNP_A_2179692	142.81
SNP_A_1852643	143.02
SNP_A_8499454	143.20
SNP_A_1974177	143.39
SNP_A_8071861	143.41
SNP_A_8254323	144.03
SNP_A_2206791	144.08
SNP_A_8207765	144.11
SNP_A_8237766	144.17
SNP_A_2055351	144.52
SNP_A_2202193	144.64
SNP_A_8029731	144.70
SNP_A_1873636	144.76
SNP_A_1967849	145.22
SNP_A_1968476	145.29
SNP_A_8348958	145.66
SNP_A_1847354	145.67
SNP_A_2292718	145.70
SNP_A_2272163	145.70
SNP_A_1922012	145.74
SNP_A_8392769	145.75
SNP_A_2212058	145.76
SNP_A_8241420	145.77
SNP_A_8777429	145.80
SNP_A_8792024	145.82
SNP_A_2284649	145.88
SNP_A_8071815	145.89
SNP_A_8264712	145.93
SNP_A_2192212	145.92
SNP_A_8246608	145.93
SNP_A_2138626	145.94
SNP_A_8848626	145.97
SNP_A_4301366	146.02
SNP_A_2232769	146.03
SNP_A_8886973	146.04

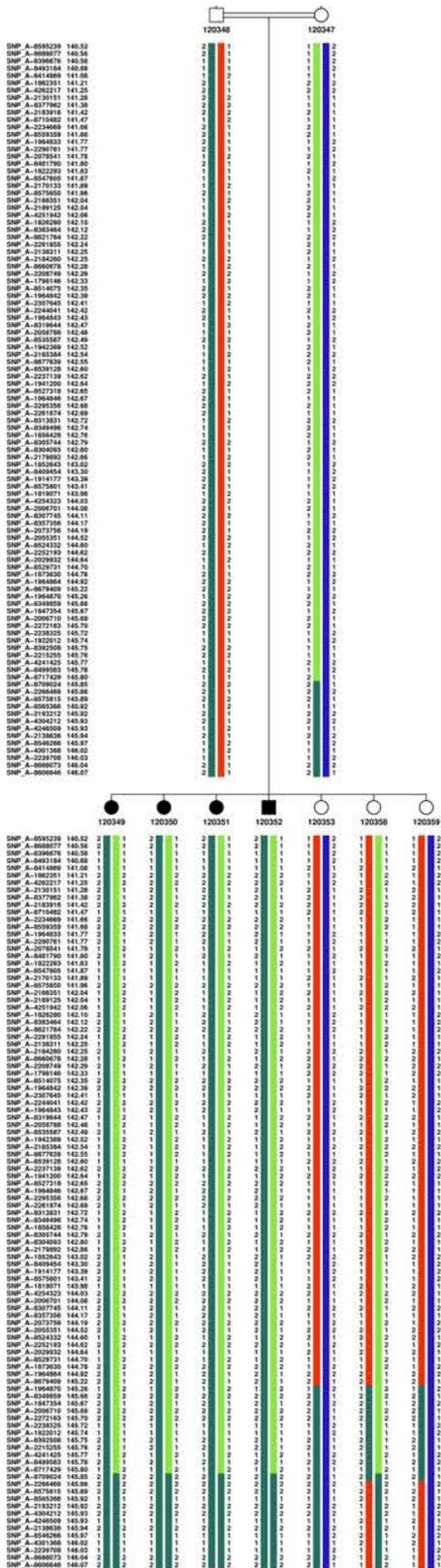


B) Genealogía C, Cromosoma16

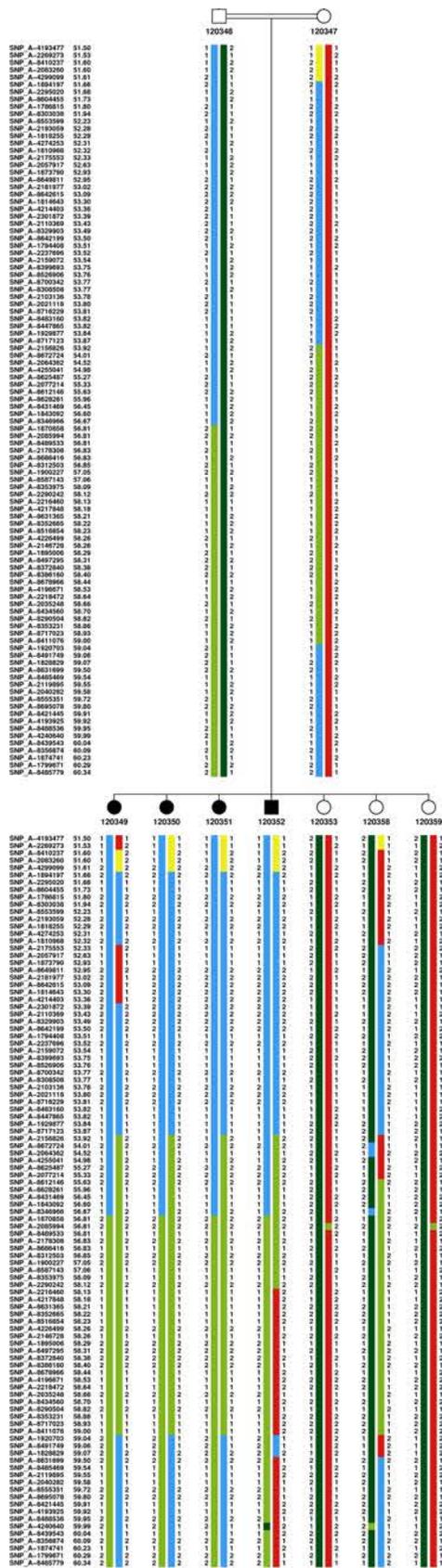




E) Genealogía E, cromosoma 2

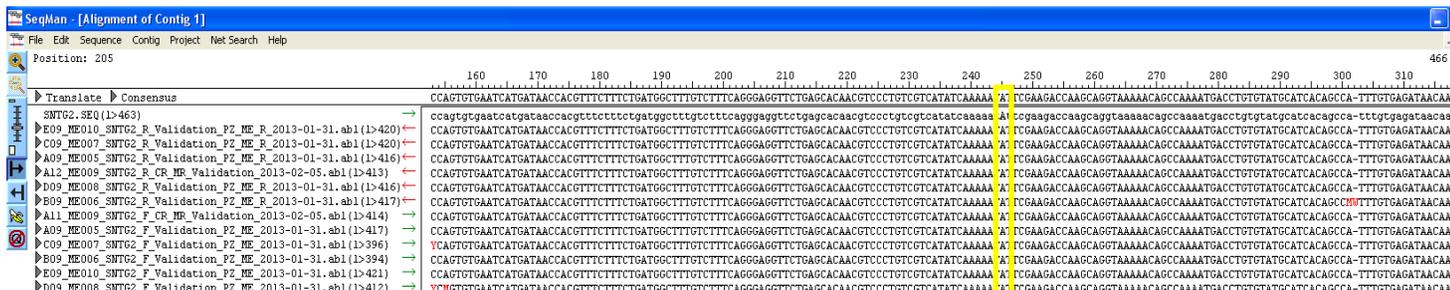


F) Genealogía E, cromosoma 16

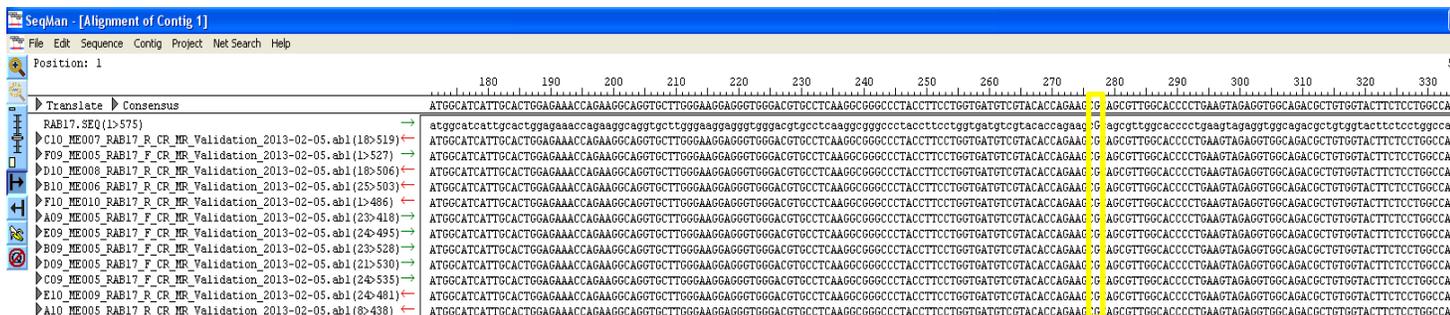


9.6 Análisis de secuencias para los cambios genéticos en los genes SNTG2, RAB17 y ILKAP en todos los miembros de la familia ME.

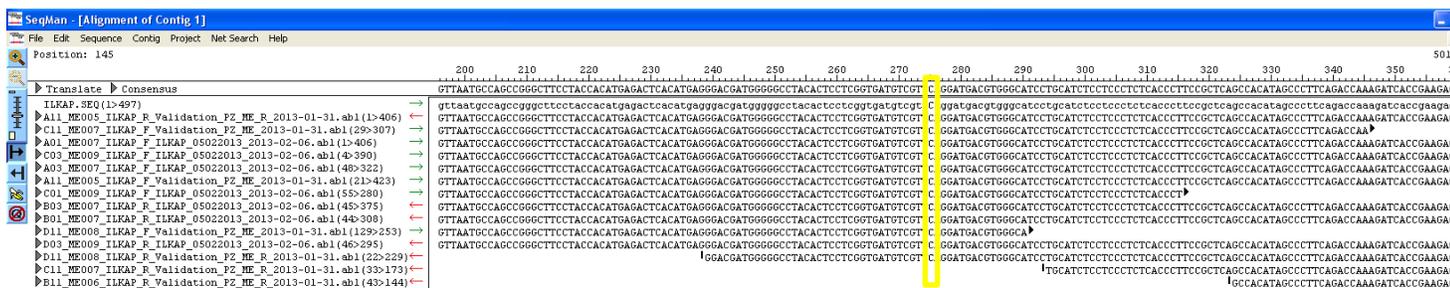
A. SNTG2



B. RAB17

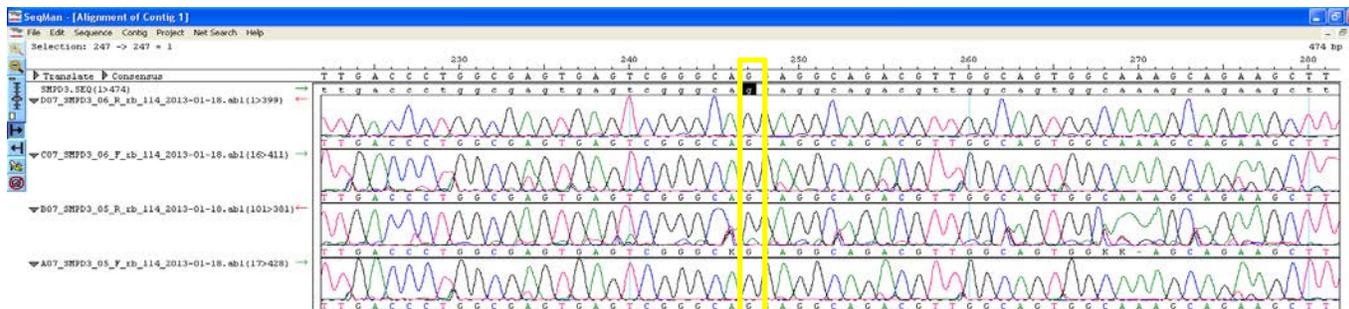


C. ILKAP

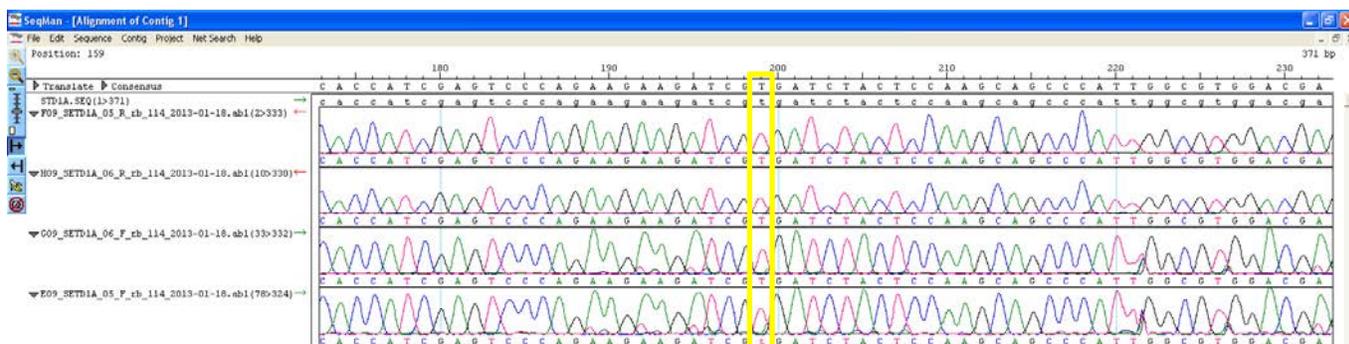


9.7 Análisis de secuencias para los cambios genéticos , que no fueron validados, en los genes SMPD3 y SETD1A en los cuatro individuos afectados y TAOK2 y KIAA0554 en todos los miembros de la familia PZ.

A. SMPD3



B. SETD1A



C. TAOK2

