



Universitat Autònoma de Barcelona

ADVERTIMENT. L'accés als continguts d'aquesta tesi queda condicionat a l'acceptació de les condicions d'ús establertes per la següent llicència Creative Commons:  http://cat.creativecommons.org/?page_id=184

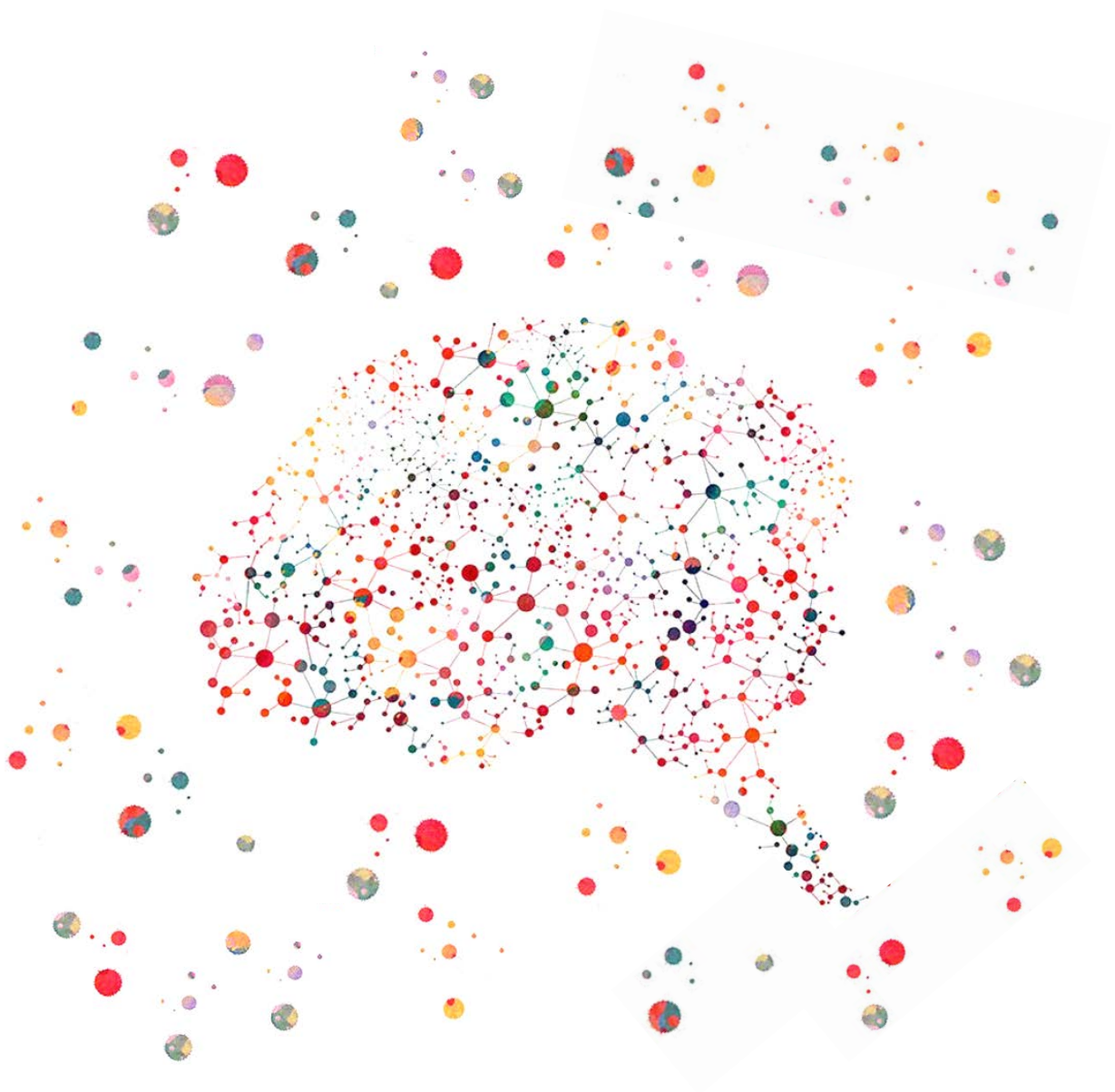
ADVERTENCIA. El acceso a los contenidos de esta tesis queda condicionado a la aceptación de las condiciones de uso establecidas por la siguiente licencia Creative Commons:  <http://es.creativecommons.org/blog/licencias/>

WARNING. The access to the contents of this doctoral thesis it is limited to the acceptance of the use conditions set by the following Creative Commons license:  <https://creativecommons.org/licenses/?lang=en>

TESI DOCTORAL 2020

Identificació de nous gens responsables de la síndrome d'Angelman-*like*

Cinthia Aguilera Román





Identificació de nous gens causants de la síndrome d'Angelman-*like*

TESI DOCTORAL
Cinthia Aguilera Román
Sabadell, Juny 2020

Identificació de nous gens responsables de la síndrome d'Angelman-*like*

Cinthia Aguilera Román

Memòria presentada per optar al títol de **doctora en Biologia Cel·lular**
per la Universitat Autònoma de Barcelona.

Tesi doctoral inscrita en el departament de Biologia Cel·lular, Fisiologia
i Immunologia, Facultat de Medicina, Universitat Autònoma de
Barcelona i realitzada a la Corporació Sanitària Parc Taulí.

Directores:

Anna Ruiz Nel·lo
Neus Baena Díez

Tutora:

Joaquima Navarro Ferreté

Sabadell, Juny 2020

La **Dra. Neus Baena Díez**, facultativa del laboratori de Genètica de la Corporació Sanitària Parc Taulí i la **Dra. Anna Ruiz Nel·lo**, facultativa del laboratori de Genètica de la Corporació Sanitària Parc Taulí,

CERTIFIQUEN:

Que Cinthia Aguilera Román ha realitzat sota la seva direcció el treball d'investigació que s'exposa en la memòria titulada **"Identificació de nous gens responsables de la síndrome d'Angelman-like"** per optar al títol de Doctora per la Universitat Autònoma de Barcelona.

Que aquest treball s'ha portat a terme en el laboratori de Genètica de la Corporació Sanitària Parc Taulí en el grup d'investigació Discapacitat intel·lectual associada a trastorns psiquiàtrics d'origen genètic i síndrome d'Angelman.

I per a que així consti, signen el present certificat.

Dra. Neus Baena Díez

Dra. Anna Ruiz Nel·lo

Sabadell, Juny 2020

La recerca portada a terme en aquesta tesi doctoral ha estat possible gràcies al finançament de l'Institut de Salut Carlos III (PI16/01411), l'associació espanyola de la síndrome d'Angelman (ASA) i l'Institut d'Investigació i Innovació Parc Taulí I3PT (CIR2016/025, CIR2018/021). Aquesta recerca també ha estat realitzada en col·laboració amb els grups de Recerca Translacional en Càncer en la Infància i l'Adolescència, Recerca Biomèdica en Urologia i Patologia Molecular Translacional de l'Institut de Recerca de la Vall d'Hebron (VHIR).



Als meus pares

A la meva germana

Agraïments

Aquesta tesi doctoral ha estat possible gràcies a la dedicació i l'ajuda que he rebut per part de les meves directores de tesi Anna Ruiz i Neus Baena. També vull donar les gràcies a la Miriam Guitart i Elisabeth Gabau, sense les quals aquest projecte no hagués estat possible i de les que he après molt de genètica, de l'esforç i la dedicació per la feina, i a la Núria Capdevila pel temps dedicat i l'esforç contactant amb les famílies.

Vull agrair a tots els tècnics de genètica i d'immunologia, Mari Carmen, Laura, Carme, Edu, Encarna, Mila i Anna. Ha estat un plaer treballar amb vosaltres aquests quatre anys, he après moltíssimes coses i, sobretot, m'heu fet sentir com a casa.

També agrair a l'Anna Santamaria, Miquel Segura, Stefan Hümmer i Marc Masanes del VHIR, els quals amb la seva col·laboració, també han fet possible aquesta tesi.

I no m'oblido de la meva família que sense el seu suport no hagués arribat tan lluny.

A tots moltes gràcies!

Índex

ABREVIATURES	2
1.INTRODUCCIÓ.....	9
1.1 Trastorns del neurodesenvolupament	9
1.1.1 Gens associats a trastorns del neurodesenvolupament	10
1.2 La síndrome d'Angelman (SA)	12
1.2.1 Aspectes clínics	12
1.2.2 Etiologia de la SA	15
1.2.2.1 Mecanismes moleculars	16
1.2.2.2 Gen <i>UBE3A</i>	17
1.2.3 Correlació genotip-fenotip	20
1.2.4 Guany de funció del gen <i>UBE3A</i>	22
1.2.5 Diagnòstic diferencial de la SA	22
1.3 Variació genètica i aplicació de les noves tecnologies de seqüenciació en els trastorns del neurodesenvolupament	24
1.3.1 Evolució de les tecnologies de seqüenciació	26
1.3.2 Aplicació de les noves tecnologies de seqüenciació en els trastorns del neurodesenvolupament	27
1.3.2.1 Heterogeneïtat genètica i variabilitat fenotípica en els trastorns del neurodesenvolupament	28
1.4 Classificació de les variants genètiques segons la guia ACMG	30
1.5 Impacte funcional de les variants genètiques	33
1.5.1 Mecanismes de patogenicitat	33
1.5.2 Mecanisme de qualitat post-transcripcional <i>Nonsense-Mediated mRNA Decay</i> (NMD)	34
1.5.3 Estudis funcionals de variants	36
1.6 Quinesina KIF1A	37
2. OBJECTIUS.....	44
3.MATERIAL I MÈTODES	46
3.1 Subjectes d'estudi	46
3.1.1 Selecció dels pacients amb fenotip <i>SA-like</i>	46

3.1.2	Descripció clínica de la cohort d'estudi	46
3.2	Mètodes	47
3.2.1	Separació de limfòcits a partir de sang perifèrica	47
3.2.2	Extracció d'ADN	47
3.2.3	Extracció d'ARN	48
3.2.4	Retrotranscripció	48
3.2.5	Seqüenciació de l'exoma i filtratge de les variants	48
3.2.5.1	Seqüenciació Sanger	52
3.2.6	Anàlisi de les CNVs exòniques a partir de les dades de l'exoma	54
3.2.6.1	<i>Multiplex Ligation-dependent Probe Amplification</i> (MLPA) customitzada	54
3.2.7	Quantificació de l'al·lel amb la variant amb el programa <i>Minor Variant Finder</i>	55
3.2.8	<i>Real Time-Polymerase Chain Reaction</i> (RT-PCR) amb sondes Taqman	56
3.2.9	Clonatge del domini motor de KIF1A <i>wild type</i> (WT) i mutat (R169T) plàsmid d'expressió pQE-80L	57
3.2.9.1	Subclonatge en el vector d'expressió pQE-80L	58
3.2.10	Producció de proteïna recombinant	59
3.2.10.1	Anàlisi de l'expressió de les proteïnes recombinants per <i>Western Blot</i>	61
3.2.10.2	Purificació de la proteïna recombinant	62
3.2.11	Assaig ATPasa del domini motor WT i R169T de KIF1A	63
3.2.12	Anàlisi de xarxes d'interacció	64
3.2.13	Re-avaluació clínica dels pacients	66
4.	RESULTATS	68
Capítol 1.	Manuscrits	69
1.1	Manuscrit I	69
Annex I.	Manuscrit I	91
1.2	Manuscrit II	93
Capítol 2.	Anàlisi d'altres models d'herència mendelians	99
2.1	Herència lligada al cromosoma X	99
2.2	Herència autosòmica recessiva	102
Capítol 3.	Avaluació de la càrrega genètica <i>de novo</i> en els pacients amb un fenotip SA-like	103
Capítol 4.	Anàlisis complementaris realitzats als pacients sense diagnòstic molecular	108

4.1	Detecció de delecions/duplicacions intragèniques en gens implicats en el neurodesenvolupament	108
4.2	Avaluació de variants en mosaic.....	110
Capítol 5.	Activació del mecanisme <i>Nonsense-mediated mRNA Decay</i> (NMD) per variants stop i <i>frameshift</i>	112
Capítol 6.	Impacte de la variant R169T sobre la funció de KIF1A	121
6.1	Activitat ATPasa del domini motor de KIF1A amb la variant R169T.....	121
Capítol 7.	Xarxa d'interacció dels nous gens identificats responsables de la SA-like, els gens del diagnòstic diferencial de la SA i el gen <i>UBE3A</i>	125
Annex II.	Resum de les variants identificades en els pacients SA-like.....	130
5. DISCUSSIÓ	133
5.1	Anàlisi de la patogenicitat de les variants identificades	136
5.2	Anàlisis complementaris en els pacients sense diagnòstic molecular.....	146
5.3	Xarxa d'interacció dels gens identificats com a responsables del fenotip SA-like i <i>UBE3A</i>	150
5.4	Re-avaluació clínica dels pacients amb diagnòstic molecular.....	151
6. CONCLUSIONS	155
7. BIBLIOGRAFIA	157
ANNEX III	176

Índex de figures

Figura 1. Etapes del desenvolupament del cervell.....	9
Figura 2. Expressió diferencial dels gens de la regió 15q11-q13	16
Figura 3. Criteris per a la classificació d'una variant com a patogènica o benigna seguint la guia ACMG.....	31
Figura 4. Model que mostra l'activació del mecanisme de control post-transcripcional <i>Nonsense-Mediated mRNA Decay</i>	35
Figura 5. Activació del mecanisme <i>Nonsense-Mediated mRNA Decay</i> (NMD) segons la localització del codó stop prematur.....	35
Figura 6. Estructura general d'una quinesina en forma dimèrica unida a la vesícula per la regió de la cua i amb el domini motor unit a microtúbuls.....	38
Figura 7. Estructura de la quinesina KIF1A.....	39
Figura 8. Transport intracel·lular de vesícules per quinesines en neurones.....	41
Figura 9. Procediment utilitzat per a la preparació de la llibreria de seqüenciació.....	49
Figura 10. Esquema que mostra dels tres passos bàsics en la seqüenciació amb la tecnologia Illumina.....	49
Figura 11. Algorisme de l'anàlisi bioinformàtic per processar de les dades obtingudes de la seqüenciació de l'exoma.....	50
Figura 12. Protocol portat a terme per a la prioritització de variants segons un model d'herència dominant <i>de novo</i>	51
Figura 13. Disseny d'una sonda de MLPA, on poden observar la nomenclatura dels diferents components de la sonda (MRC-Holland)	54
Figura 14. Esquema del clonatge realitzat per introduir el domini motor de KIF1A WT i R169T en el plàsmid d'expressió pQE-80L	59
Figura 15. Sistema de producció de proteïnes recombinants en E.Coli.....	60
Figura 16. Esquema del procés de purificació de les proteïnes recombinants.....	62
Figura 17. Disseny de la placa de 96 pous on s'ha portat a terme l'assaig ATPasa.....	63
Figura 18. Pedigrí del trio 11 on es mostra els resultats de la validació per seqüenciació Sanger de la variant c.1237G>A en el gen <i>LAS1L</i>	100
Figura 19. Pedigrí del trio 12 on es mostra els resultats de la validació per seqüenciació Sanger de la variant c.220C>A en el gen <i>COX7B</i>	101
Figura 20. Estructura de la proteïna CSTF2 on es troben indicades la variant identificada en la pacient 14 de la nostra cohort i la variant identificada per Grozdanov et al., 2020	106

Figura 21. Gràfic que mostra la normalització de la profunditat de cobertura dels exons 1-4 del gen <i>TM4SF20</i> del trio format pel pacient 15 i els seus progenitors.....	109
Figura 22. Quantificació de l'al·lel amb la variant c.1861C>T en el gen <i>SYNGAP1</i> en ADN genòmic i ADNc	112
Figura 23. Expressió del gen <i>SYNGAP1</i>	113
Figura 24. Quantificació de l'al·lel amb la variant c.1826delA en el gen <i>SATB2</i> en ADN genòmic i ADNc	114
Figura 25. Percentatge al qual es troba la variant c.1826delA en ADN genòmic i ADNc	115
Figura 26. Expressió del gen <i>SATB2</i>	115
Figura 27. Quantificació de l'al·lel amb la variant c.459_459delTGAG en el gen <i>HSF2</i> en ADN genòmic i ADNc	116
Figura 28. Percentatge al qual es troba la variant c.456_459delTGAG en ADN genòmic i ADNc....	117
Figura 29. Expressió del gen <i>HSF2</i>	117
Figura 30. Quantificació de l'al·lel amb la variant c.1165C>T en el gen <i>CHMP7</i> en ADN genòmic i ADNc	118
Figura 31. Expressió del gen <i>CHMP7</i>	118
Figura 32. Expressió de l'ARN i de la proteïna ASXL3 segons la base de dades <i>The Human Protein Atlas</i> (www.proteinatlas.org).....	119
Figura 33. Gel d'agarosa on s'han corregut el producte d'amplificació dels exons 11-12 del gen <i>ASXL3</i> en ADNc del pacient 10, mostra control i línies de neuroblastoma SKNBE(2) i CHLA-90, fent servir la parella de primers 1 i amb un nou primer <i>forward</i>	120
Figura 34. Estructura de la proteïna KIF1A on es troba indicada la variant p.Arg169Thr identificada en el pacient 1	121
Figura 35. Anàlisi d'expressió de la proteïna recombinant.....	122
Figura 36. Tinció amb <i>Naphtol blue</i> de les membranes corresponents als diferents passos de la purificació de la proteïna	123
Figura 37. Resultats de l'assaig ATPasa	123
Figura 38. Tinció de la membrana on s'han carregat diferents dilucions del domini motor WT i amb la variant R169T.....	124
Figura 39. Resultats obtinguts de l'anàlisi de la xarxa d'interacció dels nous gens identificats, els gens del diagnòstic diferencial de la SA i el gen <i>UBE3A</i> amb STRING	125
Figura 40. Xarxa d'interacció dels nous gens identificats, els gens del diagnòstic diferencial de la SA i el gen <i>UBE3A</i> obtinguda amb el programa GeneMANIA.....	127

Figura 41. Estructura de la proteïna SYNGAP1 on es troben indicats els diferents dominis de la proteïna.....	138
Figura 42. Estructura de la proteïna SATB2 on es troben indicats els diferents dominis de la proteïna.....	138
Figura 43. Estructura de la proteïna ASXL3 on es troben indicats els diferents dominis de la proteïna.....	139
Figura 44. Estructura de la subunitat del canal de potassi codificada per <i>KCNQ3</i> i com la variant R230C estableix la forma oberta del canal a voltatges negatius.....	140
Figura 45. Modelatge estructural de la variant R169T.....	142
Figura 46. Estructura del transportador GAT-1 codificat pel gen <i>SLC6A1</i>	143
Figura 47. Estructura de la proteïna codificada pel gen <i>TBL1XR1</i>	145
Figura 48. Estructura de la proteïna codificada pel gen <i>RMND1</i>	149
Figura 49. Ressonàncies magnètiques dels pacients 1 i 9.....	152
Figura 50. Algoritme diagnòstic proposat per als pacients amb un fenotip compatible amb la SA.....	153

Índex de taules

Taula 1. Consens clínic pel diagnòstic de la SA (Williams et al., 2005).....	14
Taula 2. Resum de les síndromes incloses en el diagnòstic diferencial de la SA	23
Taula 3. Número de gens associats a un o més fenotips	29
Taula 4. Criteris combinats per a la classificació de variants genètiques	32
Taula 5. Primers dissenyats per a la validació de variants per seqüenciació Sanger	53
Taula 6. Sondes de MLPA dissenyades en les regions d'interès	55
Taula 7. Primers dissenyats per quantificar l'al·lel amb la variant en ADNc dels gens <i>CHMP7</i> , <i>SYNGAP1</i> , <i>SATB2</i> , <i>HSF2</i> i <i>ASXL3</i>	56
Taula 8. Sondes TaqMan utilitzades per analitzar l'expressió gènica	57
Taula 9. Primers dissenyats per seqüenciar l'insert clonat en el vector pQE-80L	59
Taula 10. Variants obtingudes del filtratge segons un model d'herència dominant <i>de novo</i> en el pacient 17.....	91
Taula 11. Variants resultants de l'anàlisi del model d'herència lligat al cromosoma X	99
Taula 12. Variants rellevants resultants de l'anàlisi del model d'herència autosòmic recessiu ...	102
Taula 13. Variants <i>de novo</i> que potencialment alteren la proteïna en els pacients amb un fenotip <i>SA-like</i>	103
Taula 14. CNVs predites <i>de novo</i> en gens implicats en el neurodesenvolupament.....	108
Taula 15. Variants candidates resultants de l'anàlisi de variants en mosaic <i>de novo</i>	110
Taula 16. Enriquiment funcional de la xarxa dels gens identificats en la nostra cohort, els del diagnòstic diferencial de la SA i el gen <i>UBE3A</i> obtingut amb el programa STRING.....	126
Taula 17. Enriquiment funcional de la xarxa dels gens identificats en la nostra cohort, els del diagnòstic diferencial de la SA i el gen <i>UBE3A</i> obtingut amb el programa GeneMANIA	128
Taula 18. Resultats obtinguts de l'anàlisi de les vies moleculars enriquides dels gens identificats en la nostra cohort, els del diagnòstic diferencial de la SA i el gen <i>UBE3A</i> amb la base de dades Reactome.....	129
Taula 19. Variants patogèniques/probablement patogèniques responsables del fenotip <i>SA-like</i>	130
Taula 20. Variants de significat incert candidates o modificadores del fenotip <i>SA-like</i>	131



Abreviatures

ABREVIATURES

aCGH	<i>array Comparative Genomic Hybridization</i>
ACMG	<i>American College of Medical Genetics and Genomics</i>
ADN	Àcid desoxiribonucleic
ADNc	ADN complementari
ADP	<i>Adenosine diphosphate</i>
AMPAR	<i>α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid</i>
ARN	Àcid ribonucleic
ARNm	ARN missatger
ARNt	ARN de transferència
AS-SRO	<i>Angelman syndrome-Smallest Region of deletion Overlap</i>
ATP	<i>Adenosine triphosphate</i>
ATPasa	Adenosintrifosfatasa
BAM	<i>Binari Alignment Map</i>
BDNF	<i>Brain-derived neurotrophic factor</i>
BWA aligner	<i>Burrow-Wheeler Aligner</i>
ClinVar	<i>Clinical Variation Database</i>
CNV	<i>Copy Number Variant</i>
DD2GP	<i>Development Disorder Genotype - Phenotype Database</i>
DDD	<i>Deciphering Developmental Disorders</i>
ddNTPs	<i>Dideoxynucleotides</i>
Decipher	<i>Database of genomic variation and Phenotype in Humans using Ensembl Resources</i>
DI	Discapacitat intel·lectual
DMSO	<i>Dimethyl Sulfoxide</i>
DO600	Densitat òptica a 600nm
Dup	Duplicació
E. Coli	<i>Escherichia Coli</i>
EEG	Electroencefalograma
ExAC	<i>Exome Aggregation Consortium</i>
FBS	<i>Fetal Bovine Serum</i>
FDR	<i>False Discovery Rate</i>
GABA	<i>γ-aminobutyric acid</i>
GATK	<i>Genome Analysis Toolkit</i>
GDP	<i>Guanine diphosphate</i>
gnomAD	<i>Genome Aggregation Database</i>

ABREVIATURES

GO	<i>Gene Ontology</i>
GTP	<i>Guanosine-5'- triphosphate</i>
hESC	<i>human Embryonic Stem Cells</i>
hiPSC	<i>human induced Pluripotent Stem Cells</i>
IC	<i>Imprinting Center</i>
INDEL	<i>Inserció-deleció</i>
IPTG	<i>Isopropyl β-D-1-thiogalactopyranoside</i>
KDa	<i>Kilodalton</i>
LPO	<i>Left Probe Oligonucleotid</i>
Mb	<i>Megabases</i>
MLPA	<i>Multiplex Ligation-dependent Probe Amplification</i>
MS-MLPA	<i>Methylation Sensitive- Multiplex Ligation-dependent Probe Amplification</i>
NFQ	<i>Non-fluorescent quencher</i>
<i>Ni-NTA</i>	<i>Nickel-nitrilotriacetic acid</i>
NMD	<i>Nonsense-mRNA mediated decay</i>
OMIM	<i>Online Mendelian Inheritance in Man</i>
Pb	<i>Parell de bases</i>
PBS	<i>Phosphate Buffered Saline</i>
PCR	<i>Polymerase Chain Reaction</i>
Pi	<i>Inorganic Phosphate</i>
PKA	<i>Protein Kinase A</i>
pLI	<i>Probability of Loss of Function Intolerance</i>
PTC	<i>Premature stop codon</i>
PVDF	<i>Polyvinylidene fluoride</i>
PWS-SRO	<i>Prader Willi syndrome-Smallest Region of deletion Overlap</i>
QD	<i>Quocient de dosi</i>
QI	<i>Quocient Intel·lectual</i>
qPCR	<i>Quantitative Polymerase Chain Reaction</i>
Rpm	<i>Revolucions per minut</i>
RPO	<i>Right Probe Oligonucleotid</i>
SA	<i>Síndrome d'Angelman</i>
<i>SA-like</i>	<i>Síndrome d'Angelman-like</i>
SNP	<i>Single Nucleotid Polymorphism</i>
SNV	<i>Single Nucleotid Variant</i>

SWI/SNF	<i>SWItch/Sucrose Non-Fermentable</i>
SysID	<i>Systems Biology Approaches to ID</i>
TD	Trastorn del desenvolupament
TDAH	Trastorn de dèficit d'atenció i hiperactivitat
TEA	Trastorn d'espectre autista
TND	Trastorns del neurodesenvolupament
UTR	<i>Untranslated Region</i>
VCF	<i>Variant Call Format</i>
WES	<i>Whole Exome Sequencing</i>
WGS	<i>Whole Genome Sequencing</i>
WT	<i>Wild Type</i>

Recopilatori de gens

<i>ADSL</i>	<i>Adenylossuccinate lyase</i>
<i>ARID1A</i>	<i>AT-rich interaction domain 1A</i>
<i>ARID1B</i>	<i>AT-rich interaction domain 1B</i>
<i>ASXL3</i>	<i>ASXL transcriptional regulator 3</i>
<i>ATAD3A</i>	<i>ATPase family AAA domain containing 3A</i>
<i>ATP10A</i>	<i>ATPase phospholipid transporting 10A</i>
<i>ATRX</i>	<i>ATRX chromatin remodeler</i>
<i>BAP1</i>	<i>BRCA1 associated protein 1</i>
<i>CDKL5</i>	<i>Cyclin dependent kinase like 5</i>
<i>CHMP7</i>	<i>Charged multivesicular body protein 7</i>
<i>COX7B</i>	<i>Cytochrome c oxidase subunit 7B</i>
<i>CREBBP</i>	<i>CREB binding protein</i>
<i>CSTF2</i>	<i>Cleavage Stimulation Factor Subunit 2</i>
<i>DEAF1</i>	<i>DEAF1 transcription factor</i>
<i>E6-AP</i>	<i>E6-associated protein</i>
<i>EEF1A2</i>	<i>Eukaryotic translation elongation factor 1 alpha 2</i>
<i>EHMT1</i>	<i>Euchromatic histone lysine methyltransferase 1</i>
<i>EIF2B2</i>	<i>Eukaryotic translation initiation factor 2B subunit beta</i>
<i>EIF4G1</i>	<i>Eukaryotic translation initiation factor 4 gamma 1</i>
<i>EP300</i>	<i>E1A binding protein p300</i>
<i>FOXG1</i>	<i>Forkhead box G1</i>

<i>GABRB2</i>	<i>Gamma-aminobutyric acid type A receptor beta2 subunit</i>
<i>GAPDH</i>	<i>Glyceraldehyde-3-Phosphate Dehydrogenase</i>
<i>GRIN1</i>	<i>Glutamate ionotropic receptor NMDA type subunit 1</i>
<i>GRIN2A</i>	<i>Glutamate ionotropic receptor NMDA type subunit 2A</i>
<i>GRIN2B</i>	<i>Glutamate ionotropic receptor NMDA type subunit 2B</i>
<i>HERC2</i>	<i>HECT and RLD domain containing E3 ubiquitin protein ligase 2</i>
<i>HSF2</i>	<i>Heat Shock factor 2</i>
<i>HSPG2</i>	<i>Heparan Sulfate Proteoglycan 2</i>
<i>HTT</i>	<i>Huntingtin</i>
<i>IMPG1</i>	<i>Interphotoreceptor matrix proteoglycan 1</i>
<i>IQSEC2</i>	<i>IQ motif and Sec7 domain ArfGEF 2</i>
<i>KANSL1</i>	<i>KAT8 regulatory NSL complex subunit 1</i>
<i>KAT5</i>	<i>Lysine acetyltransferase 5</i>
<i>KCNQ2</i>	<i>Potassium voltage-gated channel subfamily Q member 2</i>
<i>KCNQ3</i>	<i>Potassium voltage-gated channel subfamily Q member 3</i>
<i>KIF1A</i>	<i>Kinesin family member 1A</i>
<i>KMT2C</i>	<i>Lysine methyltransferase 2C</i>
<i>LAS1L</i>	<i>LAS1 like ribosome biogenesis factor</i>
<i>MAGEL2</i>	<i>MAGE family member L2</i>
<i>MAPK8IP2</i>	<i>Mitogen-activated protein kinase 8 interacting protein 2</i>
<i>MAPK8IP3</i>	<i>Mitogen-activated protein kinase 8 interacting protein 3</i>
<i>MBD5</i>	<i>Methyl-CpG binding domain protein 5</i>
<i>MECP2</i>	<i>Methyl-CpG binding protein 2</i>
<i>MEF2C</i>	<i>Myocyte enhancer factor 2C</i>
<i>METTL24</i>	<i>Methyltransferase like 24</i>
<i>MKNR3</i>	<i>Makorin ring finger protein 3</i>
<i>NDN</i>	<i>Necdin, MAGE family member</i>
<i>NLGN</i>	<i>Neuroigin</i>
<i>NLGN4X</i>	<i>Neuroigin 4 X-Linked</i>
<i>NR1I3</i>	<i>Nuclear receptor subfamily 1 group I member 3</i>
<i>NRXN</i>	<i>Neurexin 1</i>
<i>PI4KA</i>	<i>Phosphatidylinositol 4-kinase alpha</i>
<i>PLXNB2</i>	<i>Plexin B2</i>
<i>PWRN1</i>	<i>Prader-Willi region non-protein coding RNA 1</i>

ABREVIATURES

<i>PWRN2</i>	<i>Prader-Willi region non-protein coding RNA 2</i>
<i>RAB5B</i>	<i>RAB5B, member RAS oncogene family</i>
<i>RAI1</i>	<i>Retinoic acid induced 1</i>
<i>RHOBTB2</i>	<i>Rho related BTB domain containing 2</i>
<i>RMND1</i>	<i>Required for Meiotic Nuclear Division 1 homolog</i>
<i>SATB2</i>	<i>SATB homeobox 2</i>
<i>SCN2A</i>	<i>Sodium voltage-gated channel alpha subunit 2</i>
<i>SHANK2</i>	<i>SH3 and Multiple Ankyrin Repeat Domains 2</i>
<i>SHANK3</i>	<i>SH3 and multiple ankyrin repeat domains 3</i>
<i>SHROOM4</i>	<i>Shroom family member 4</i>
<i>SLC35A2</i>	<i>Solute carrier family 35 member A2</i>
<i>SLC6A1</i>	<i>Solute carrier family 6 member 1</i>
<i>SLC9A6</i>	<i>Solute carrier family 9 member A6</i>
<i>SMARCA1</i>	<i>SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 1</i>
<i>SMARCA2</i>	<i>SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2</i>
<i>SMARCA4</i>	<i>SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4</i>
<i>SMARCB1</i>	<i>SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1</i>
<i>SMARCE1</i>	<i>SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily e, member 1</i>
<i>SMCR8</i>	<i>SMCR8-C9orf72 complex subunit</i>
<i>SMG1</i>	<i>MG1 nonsense mediated mRNA decay associated PI3K related kinase</i>
<i>SNORD116</i>	<i>Small nucleolar RNA, C/D box 116 cluster</i>
<i>SNURF-SNRPN</i>	<i>Small nuclear ribonucleoprotein polypeptide N</i>
<i>SOX11</i>	<i>SRY-box transcription factor 11</i>
<i>SPEG</i>	<i>Striated Muscle Enriched protein kinase</i>
<i>SPTAN1</i>	<i>Spectrin alpha, non-erythrocytic 1</i>
<i>SPW</i>	<i>Síndrome de Prader Willi</i>
<i>STXBP1</i>	<i>Syntaxin binding protein 1</i>
<i>SYNGAP1</i>	<i>Synaptic Ras GTPase activating protein 1</i>
<i>TBL1XR1</i>	<i>Transducin beta like 1 X-linked receptor 1</i>

ABREVIATURES

<i>TCF4</i>	<i>Transcription factor 4</i>
<i>TM4SF20</i>	<i>Transmembrane 4 L six family member 20</i>
<i>TUBB2B</i>	<i>Tubulin beta 2B class IIb</i>
<i>UBE3A</i>	<i>Ubiquitin protein ligase E3A</i>
<i>UBE3A-ATS</i>	<i>UBE3A antisense transcript</i>
<i>UPF1</i>	<i>UPF1 RNA helicase and ATPase</i>
<i>UPF2</i>	<i>UPF2 regulator of nonsense mediated mRNA decay</i>
<i>UPF3B</i>	<i>UPF3B regulator of nonsense mediated mRNA decay</i>
<i>VAMP2</i>	<i>Vesicle associated membrane protein 2</i>
<i>ZEB2</i>	<i>Zinc finger E-box binding homeobox 2</i>
<i>ZNF238</i>	<i>Zinc Finger Protein 238</i>



Introducció

1.1 Trastorns del neurodesenvolupament (TND)

El neurodesenvolupament es defineix com el procés biològic que compren el desenvolupament i la maduració del sistema nerviós. El desenvolupament del cervell comença en la tercera setmana de gestació amb la formació del tub neural. A partir de la novena setmana fins al final de la gestació es produeixen processos de proliferació, migració i diferenciació cel·lular que esdevenen cap al final del període embrionari en la formació dels compartiments del sistema nerviós central, del sistema nerviós perifèric i en la formació dels circuits neuronals (1,2) (Figura 1). Qualsevol disrupció en aquesta complexa cadena d'esdeveniments, pot causar disfuncions en el desenvolupament del cervell i com a conseqüència resultar en trastorns del neurodesenvolupament (TND) (3,4).

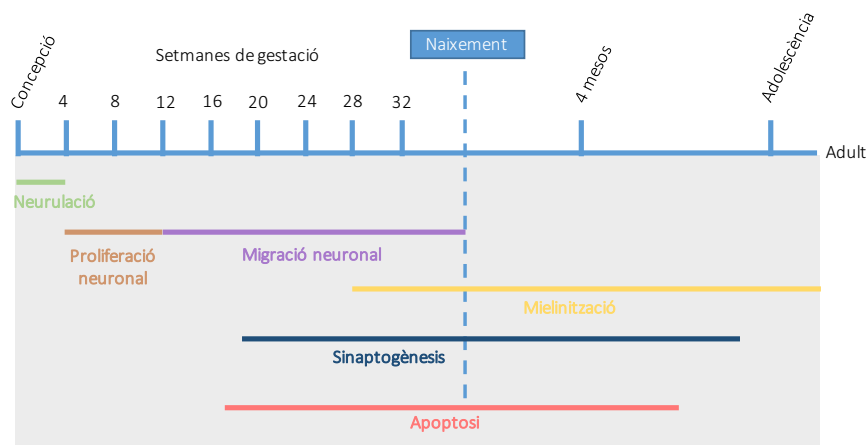


Figura 1 | Etapes del desenvolupament del cervell. Imatge modificada i adaptada de Tau GZ et al., 2009.

Els TND són un grup de trastorns complexos i heterogenis que es caracteritzen per presentar afectacions a nivell cognitiu, motor, de comunicació i de comportament adaptatiu que es manifesten en etapes primerenques del desenvolupament (5). L'espectre dels dèficits en el desenvolupament compren des de problemes d'aprenentatge i de les funcions executives a problemes globals d'interacció social o intel·ligència. Dins dels TND s'inclouen el Trastorn d'Espectre Autista (TEA), Discapacitat Intel·lectual (DI), Trastorn de Dèficit d'Atenció i Hiperactivitat (TDAH), i problemes de llenguatge (DSM-5) (4–7). En molts casos, hi ha comorbiditat de més d'un TND, per exemple, els individus amb TEA freqüentment presenten DI i els individus amb TDAH freqüentment presenten un trastorn específic de llenguatge (8). S'ha suggerit que es deu a què es troben afectades vies moleculars comunes en els TND (9).

La DI s'engloba dins dels TND i es caracteritza per dèficits a nivell intel·lectual i de funcionament adaptatiu, que inclou habilitats conceptuals, socials i pràctiques. És un dels TND més comuns, la seva prevalença en població general és de l'1-3%. S'inicia durant el període del desenvolupament i es pot manifestar fins a l'edat adulta. El terme retard global del desenvolupament s'utilitza en els individus menors de 5 anys que no es poden sotmetre als tests estandarditzats per valorar la gravetat clínica. La DI es subdivideix en quatre grups segons el quocient intel·lectual (QI): i) lleu (QI 50-69), ii) moderada (QI 35-49), greu (QI 20-34) i profunda (QI inferior a 20) (8,10,11).

Dins de la DI, la DI greu compren el 4% dels casos i té una prevalença del 0,3-0,5% (12). Els individus amb DI greu, a part de presentar un problema de funcionament intel·lectual greu, en molts casos també presenten problemes motors, un llenguatge expressiu i habilitats comunicatives greument afectades, utilitzant només paraules aïllades o frases curtes (8,12-14). L'etiologia es divideix en: i) causes genètiques, que constitueixen aproximadament el 50% dels casos, on s'inclouen els defectes monogènics, anomalies cromosòmiques i canvis en número de còpies (CNVs), ii) el 25% dels casos corresponen a factors ambientals, entre els que s'inclouen l'exposició a toxines, infeccions que afecten al sistema nerviós central, malalties que afecten a la mare durant l'embaràs i hipòxia/isquèmia i iii) un 25% dels casos presenta una etiologia desconeguda (12,15,16).

1.1.1 Gens associats amb els trastorns del neurodesenvolupament

Els gens que estan implicats en els TND porten a terme bàsicament tres funcions que són essencials en el desenvolupament del cervell: remodelació de la cromatina, regulació de la transcripció i la sinapsi.

Remodelació de la cromatina

La remodelació de la cromatina està implicada en la formació de connexions a nivell neuronal durant la neurogènesi i la diferenciació neuronal (17). Els remodeladors de la cromatina intervien regulant l'estructura de la cromatina i l'expressió gènica, a través de marques epigenètiques o modificacions post-transcripcionals d'histones com la fosforilació, metilació, acetilació i ubiquitinació. Aquestes modificacions post-transcripcionals d'histones fan que la cromatina sigui o no accessible als factors de transcripció. Estan regulades per diferents proteïnes que es classifiquen en quatre grups: i) les d'escriptura (*writers*) implicades en l'addició de grups metil, acetil, ubiquïtines o fosfat a les histones, ii) les implicades en l'eliminació de les modificacions post-transcripcionals (*erasers*), iii) les lectores (*readers*), que exerceixen la seva

funció depenent de les modificacions o marques en les histones i iv) els complexos remodeladors dependents d'ATP. Un dels complexos més importants implicats en la remodelació de la cromatina i associat a TND és el complex dependent d'ATP BAF (SWI/SNF) (17). Aquest complex intervé en el desenvolupament neuronal i en la formació d'estructures del cervell en etapa embrionària com el prosencèfal o l'epiteli olfatiu (18,19). En les cèl·lules progenitores neuronals, npBAF, actua interaccionant amb el factor de transcripció repressor REST i el seu co-repressor, inhibint gens neuronals i mantenint la identitat dels progenitors neuronals (20). Variants en algunes de les subunitats del complex SWI/SNF estan associades a la síndrome de Coffin-Siris (*SMARCB1*, *SMARCA4*, *SMARCE1*, *ARID1A*, *ARID1B* i *SOX11*) i a la síndrome Nicolaidis-Baraitser (*SMARCA2*) (21,22). Altres gens modificadors de la cromatina que estan implicats en TND són *MECP2*, *EHMT1*, *CREBBP* i *EP300* que estan associats a la síndrome de Rett, síndrome de Kleefstra i a la síndrome de Rubinstein-Taybi, respectivament (23,24).

Factors de transcripció

En el desenvolupament del cervell és crucial la regulació espacial i temporal de l'expressió gènica per controlar els processos de proliferació, diferenciació i migració neuronal, així com la formació i eliminació de sinapsis. La regulació de l'expressió gènica controla la presència i la producció de productes gènics que permeten a la cèl·lula respondre a senyals intra- i extracel·lulars, definir la identitat cel·lular durant el desenvolupament i coordinar l'activitat cel·lular. El control de la transcripció es porta a terme principalment, per factors de transcripció i modificadors de la cromatina. Els factors de transcripció s'uneixen a seqüències específiques en l'ADN i permeten la transcripció selectiva de gens per l'ARN polimerasa II (25). S'han reportat diversos factors de transcripció implicats en TND, un exemple és el factor de transcripció 4 (*TCF4*), que s'ha descrit que té llocs d'unió en gens neuronals implicats en el neurodesenvolupament. L'haploinsuficiència de *TCF4* és la causa de la síndrome de Pitt-Hopkins, variants comunes en la població general que provoquen canvis subtils en l'expressió de *TCF4* s'han associat a esquizofrènia (26,27). Un altre exemple, és el gen *SATB2* que codifica per una proteïna de la matriu nuclear que s'associa a regions riques en AT i regula l'expressió gènica en el nucli, variants en aquest gen causen la síndrome associada a *SATB2* (SAS) (28).

Sinapsis

Les neurones són cèl·lules polaritzades les quals a través d'impulsos nerviosos es comuniquen amb altres neurones a través de la sinapsi, mitjançant neurotransmissors. Els neurotransmissors són sintetitzats per les neurones presinàptiques i emmagatzemats en vesícules sinàptiques.

Aquestes vesícules es transfereixen a la zona activa on es fusionen amb la membrana presinàptica provocant l'alliberació dels neurotransmissors per exocitosi a la fenedura sinàptica. Els neurotransmissors alliberats s'uneixen als seus receptors o a canals iònics activats per la unió de lligand en la membrana postsinàptica, desencadenant una transducció del senyal químic en un flux d'ions a través de la membrana (29).

En la diferenciació neuronal i en la formació dels circuits neuronals és essencial la correcta projecció dels axons neuronals, formació i arborització de les dendrites, el manteniment i l'eliminació de les sinapsis (30). La modificació dels circuits neuronals a nivell estructural i funcional, com a conseqüència de l'activitat neuronal s'anomena plasticitat sinàptica. Durant les primeres etapes del desenvolupament en la infància, la plasticitat sinàptica és essencial per a la maduració dels circuits neuronals (31). Els gens que codifiquen proteïnes que intervenen en els processos anteriorment esmentats s'han associat a TND. Un exemple són les neuroxines (*NRXN*) que són proteïnes d'adhesió sinàptica que es troben localitzades en la membrana presinàptica i s'uneixen a les neuroligines (*NLGN*) situades en la membrana postsinàptica. Variants tant en *NRXN* com en *NLGN* han estat associades a TEA. Un altre exemple, són els receptors d'àcid γ -aminobutíric (GABA) com *GABRB3* o de glutamat com *GRIN2A* o *GRIN2B* que intervenen en les sinapsis inhibidores i excitatòries, respectivament, i han estat associats a encefalopaties epilèptiques i TEA (32,33).

1.2 Síndrome d'Angelman (SA)

La síndrome d'Angelman (SA) és un TND que es caracteritza per una DI greu amb absència de llenguatge, característiques craneofacials dismòrfiques distintives, problemes neurològics com l'atàxia i/o tremolor de les extremitats i epilèpsia associada a un patró específic en l'electroencefalograma (EEG) que persisteix en l'edat adulta. Entre els trets del fenotip conductual es troben l'aparença feliç, hiperactivitat i dèficit d'atenció. Altres problemes clínics inclouen problemes d'alimentació i gastrointestinals, trastorn del cicle son-vigília, escoliosi i hipersensibilitat a la calor (34–36). La prevalença de la SA és d'aproximadament 1/15.000 naixements (37,38). L'edat mitja del diagnòstic es troba per sobre dels 18 mesos a causa de la inespecificitat clínica que presenten els individus durant el primer any de vida (36).

1.2.1 Aspectes clínics de la SA

El consens clínic establert per Williams et al., (36) l'any 2005 i que actualment segueix sent vigent pel diagnòstic de la SA considera trets clínics consistents els presents en el 100% dels individus.

Aquests trets inclouen el retard del desenvolupament greu, trastorn del moviment o l'equilibri (atàxia), afectació greu del llenguatge i fenotip conductual característic amb qualsevol combinació d'atacs de riure, aparença feliç, hiperexcitabilitat, estereotípies (principalment aleteig de mans), hiperactivitat i tendència a la dispersió. Els trets clínics freqüents es consideren aquells presents en més del 80% dels casos, com són la microcefàlia, les crisis epilèptiques i EEG patològic amb un patró característic. Els trets associats estan presents en el 20-80% dels casos. Entre ells es troben: occipuci pla, solc occipital, protrusió lingual, trastorns de succió/deglució, problemes d'alimentació, prognatisme, boca gran amb espai interdental augmentat, baveig freqüent, moviments masticatoris, estrabisme, hipopigmentació de pell i ulls, hiperreflèxia de les extremitats inferiors, braços aixecats i semiflexionats durant la marxa, trastorn del cycle son-vigília, atracció i fascinació per l'aigua, trastorns d'alimentació, obesitat en adults, escoliosi i restrenyiment (36,39) (Taula 1).

Trets clínics consistents i freqüents de la SA:

Capacitat cognitiva i llenguatge

Les habilitats cognitives es troben greument afectades i l'edat mental dels pacients SA correspon a la d'un individu de 3 a 17 mesos. El llenguatge també està greument afectat, essent el llenguatge expressiu més afectat que el receptiu, ja que són capaços de comprendre i comunicar-se amb sistemes de comunicació alternatius com gestos i senyalant objectes o pictogrames (40,41).

Fenotip conductual

El fenotip conductual es caracteritza per una aparença feliç amb atacs de riure, dèficit d'atenció, hiperactivitat, hiperexcitabilitat, estereotípies, atracció/fascinació per l'aigua i exploració dels objectes amb la boca. Els individus amb SA mostren una sensibilitat disminuïda al dolor i una hipersensibilitat a la calor (42).

Desenvolupament motor

Els problemes motors més comuns en individus amb SA inclouen espasticitat, atàxia al caminar i tremolors. La majoria dels individus tenen un to muscular normal, tot i que, en un 25% dels casos hi ha una hipotonia persistent i un 30% dels casos desenvolupa hipertonia. L'edat de sedestació es situa en els 12 mesos, mentre que l'adquisició de la marxa es situa en els 3 anys d'edat. Un 10% dels individus no caminen i els que ho aconsegueixen presenten en molts casos una marxa atàxica i inestable. En alguns casos, mostren tremolors de les extremitats que s'incrementen en l'edat adulta (43-45).

Epilèpsia i EEG

Les crisis epilèptiques s'inicien abans dels tres anys de vida, els principals tipus d'epilèpsia que presenten els pacients amb SA són absències, generalitzada tònica-clònica i atònica-mioclònica (46,47). El 80% dels pacients presenten un EEG alterat, s'han descrit tres tipus de patró que són característics de la SA, el patró més comú és el delta, però també hi ha individus que presenten el patró theta i el de puntes en àrees posteriors (48).

Taula 1 | Consens clínic pel diagnòstic de la SA (Williams et al., 2005) (35).

<p>Característiques clíniques consistents (presentes en el 100% dels individus amb SA)</p> <ul style="list-style-type: none"> -Retard del desenvolupament greu -Trastorn del moviment: marxa atàxica, tremolor de les extremitats o inestabilitat psicomotriu -Fenotip conductual: combinació d'atacs de riure, aparença feliç, estereotípies (principalment aleteig de mans), comportament hipermotoric, hiperexcitabilitat i tendència a la dispersió -Afectació greu del llenguatge que inclou l'absència de llenguatge o l'ús mínim de paraules. Habilitats comunicatives no verbals i llenguatge receptiu superiors al llenguatge expressiu
<p>Característiques clíniques freqüents (presentes en més del 80% dels individus amb SA)</p> <ul style="list-style-type: none"> -Estantament en el creixement del perímetre cefàlic, resultant en microcefàlia postnatal -Epilèpsia. L'aparició de les crisis comicials es dona abans dels 3 anys de vida i millora en l'edat adulta -Anomalies en l'EEG amb un patró característic
<p>Característiques clíniques associades (presentes en el 20- 80% dels individus amb SA)</p> <ul style="list-style-type: none"> -Occipuci pla -Solc occipital -Protrusió lingual -Trastorns de succió/deglució -Problemes d'alimentació/hipotonia troncal -Prognatisme -Boca ampla i espai interdental augmentat -Baveig freqüent -Moviments masticatoris -Estrabisme -Hipopigmentació de la pell, cabell i ulls comparat amb altres membres de la família (present en individus amb deleció de la regió cromosòmica 15q11-q13) -Hiperreflèxia de les extremitats inferiors -Braços aixecats i semiflexionats durant la marxa -Hipersensibilitat a la calor -Trastorns del cicle son-vigília i disminució de la necessitat de dormir -Atracció/fascinació per l'aigua -Trastorns de l'alimentació -Obesitat (en adults) -Escoliosi -Restrenyiment

1.2.2 Etiologia de la SA

La causa de la SA és la pèrdua d'expressió en les neurones de la proteïna ubiquitina lligasa E6-AP (proteïna associada E6) codificada per la còpia materna del gen *UBE3A*, localitzat a la regió cromosòmica 15q11-q13. La regió cromosòmica 15q11-q13 compren 6Mb i inclou un conjunt de gens que mostren una expressió diferencial depenent de l'origen parental. Aquesta regió està flanquejada per repeticions en baix número de còpies que originen delecions i/o duplicacions com a conseqüència d'una recombinació homòloga desigual. Es coneixen tres punts de trencament principals (o *breakpoints*, BP): BP1, BP2 i BP3. També s'han descrit punts més distals a BP3 denominats BP4, BP5 i BP6.

La regió 15q11-q13 es pot subdividir en (Figura 2):

-La regió més centromèrica que està situada entre els punts de trencament BP1-BP2 i conté els gens *NIPA1*, *NIPA2*, *CYFIP1* i *TUBGCP5*, que tenen una expressió biparental.

-La regió entre els punts de trencament BP2 i BP3 compren 4 Mb i es troba sotmesa a impronta genòmica. Els gens *MKRN3*, *MAGEL2*, *NDN*, *PWRN1*, *C15orf2*, *SNURF-SNRPN*, i diversos gens que codifiquen per *small nucleolar RNAs* (snoRNA) tenen una expressió paterna. El gen *UBE3A* s'expressa de manera al·lel específica en neurones, mentre que en l'al·lel patern s'expressa el transcrit antisentit *UBE3A-ATS* que impedeix l'expressió del gen *UBE3A*. En canvi, en l'al·lel matern no hi ha expressió d'*UBE3A-ATS*, permetent l'expressió del gen *UBE3A*. Els gens *PWNR2*, *ATP10A*, *GABRB3*, *GABRB5*, *GABRG3*, *OCA2* i *HERC2* tenen una expressió biparental.

L'expressió gènica de la regió entre els punts de trencament BP2 i BP3 es troba regulada pel centre de impronta (IC) que es troba subdividit en dues regions crítiques, la de Prader Willi (PWS-SRO) i la de SA (AS-SRO). PW-SRO inclou la regió promotora i l'exó 1 del gen *SNURF-SNRPN* i promou l'expressió dels gens del cromosoma patern mentre que AS-SRO actua reprimint el centre PWS-SRO en el cromosoma matern (49,50) (Figura 2). La pèrdua de l'expressió dels gens de l'al·lel patern (en particular el snoRNA, *SNORD116*) produeix la síndrome de Prader-Willi (SPW).

El gen *UBE3A* té una expressió específica de l'al·lel matern en neurones madures, principalment en les regions de l'hipocamp, còrtex, tàlem, bulb olfatori i en el cerebel, mentre que en la resta de teixits l'expressió és bial·lèlica (51). En neurones, l'al·lel patern del gen *UBE3A* està silenciada per la síntesi del transcrit antisentit *UBE3A-ATS* que s'inicia en el gen *SNURF-SNRPN*. L'expressió tant del transcrit *UBE3A-ATS* com del transcrit *UBE3A* en la cadena oposada en l'al·lel patern provoca la col·lisió de les maquinàries de transcripció a nivell de l'intró 4. Com a conseqüència es produeix una dissociació de les polimerases, acabant la transcripció prematura dels dos transcrits, l'antisentit i el que codifica la proteïna E6-AP (52).

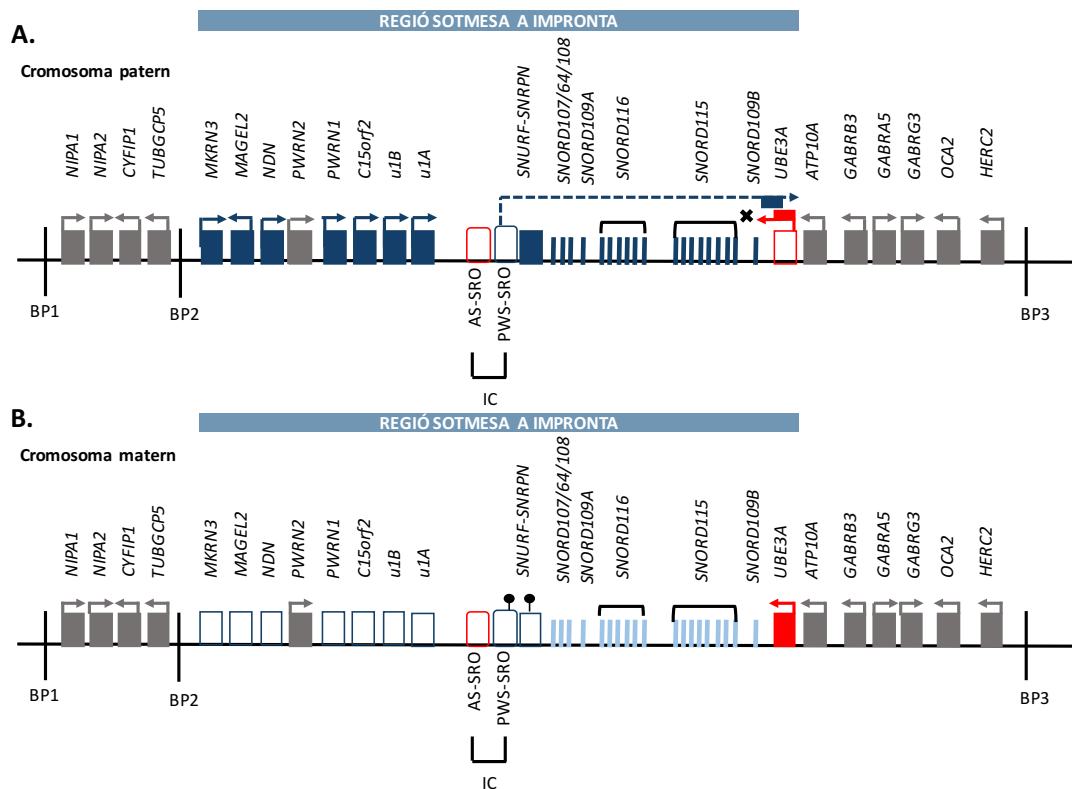


Figura 2 | Expressió diferencial dels gens de la regió 15q11-q13. **A)** Gens expressats en el cromosoma patern, **B)** Gens expressats en el cromosoma matern. Els rectangles en blau mostren gens d'expressió paterna, en vermell d'expressió materna i els grisos d'expressió bial·lèlica. Les fletxes indiquen la orientació de la transcripció. Els rectangles blancs emmarcats en blau o vermell, indiquen els gens que no s'expressen en el cromosoma patern o matern, respectivament. Les línies en color blau són snoRNAs i en color negre els punts de trencament (BP). El centre de impronta (IC) està representat en vermell el AS-SRO i en blau el PWS-SRO. Els cercles en negre indiquen la metilació de l'ADN. El transcrit *antisense* UBE3A-ATS s'inicia *upstream* del gen *SNURF-SNRPN* i impedeix la transcripció del gen *UBE3A* en el cromosoma patern.

1.2.2.1 Mecanismes moleculars.

La SA està causada pels següents mecanismes moleculars (53,54):

- La deleció de la regió 15q11-q13 en el cromosoma d'origen matern és la causa més freqüent de la SA (70-75% dels casos). Dins de les deleccions, la deleció de tipus I compren el 40% dels casos i inclou els punts de trencament BP1 a BP3, d'altra banda, la deleció de tipus II compren el 50% dels casos de deleció i inclou del punt de trencament BP2 a BP3. Deleccions que incloguin punts de trencament més llunyans (BP4, BP5 i BP6) es consideren deleccions atípiques (55). El risc de recurrència és inferior a l'1% si el cariotip de la mare és normal.
- Variants puntuals patogèniques/probablement patogèniques en el gen *UBE3A*. Constitueixen la segona causa més freqüent (10-15% dels casos) i poden ser *de novo* o familiars. En la revisió portada terme per Sadikovic et al., 2014 un 71,7% són *de novo* i un 28,9% heretades (56). Les variants puntuals identificades es localitzen principalment en els exons 3 i 10 del gen *UBE3A*

(NM_130838.1) considerant-se *hotspots* (56,57). El risc de recurrència és del 50% quan la mare és portadora. En molt baixa freqüència, s'han reportat delecions intragèniques o totals del gen *UBE3A* (58).

- La disomia uniparental paterna del cromosoma 15 és la causa del 2-5% dels casos. Majoritàriament són isodisomies degudes a una no disjunció materna, seguit d'una duplicació postzigòtica del cromosoma 15 patern (59,60). El risc de recurrència és inferior a l'1% si el cariotip dels progenitors és normal.
- El defecte en el centre de la impronta succeeix un 1-2% dels casos amb SA. És la causa menys freqüent, està produïda per errors en l'establiment de la impronta genòmica (85%) o delecions en el centre de la impronta (15%). En tots dos casos el resultat és un cromosoma matern amb un epigenotip patern. Els defectes en l'establiment de la impronta són esporàdics i el risc de recurrència és inferior a l'1%, mentre que les delecions en el centre de la impronta majoritàriament són familiars amb un risc de recurrència del 50% (61).

En el 10% dels individus amb diagnòstic clínic de SA no es confirma el mecanisme molecular descrit, és a dir, no presenten un patró de metilació anormal de la regió 15q11-q13 ni variants patogèniques en el gen *UBE3A*. Aquests pacients es defineixen com a síndrome d'Angelman-like. S'han descrit alguns pacients que presentarien alteracions genètiques en gens responsables de síndromes que són considerades diagnòstic diferencial de la SA ja que presenten característiques clíniques solapants (62).

1.2.2.2 Gen *UBE3A*

El gen *UBE3A* codifica per una E3 ubiquitina lligasa E6-AP (*E6-Associated Protein*) que catalitza la transferència d'ubiquitines als residus de lisina de les proteïnes subjectes de degradació pel sistema ubiquitina proteosoma. La família de proteïnes E3 lligases es caracteritza perquè totes presenten el domini C-terminal HECT (*Homologous to E6-AP C-Terminus*), el qual conté el lloc actiu ubiquitina lligasa (63). En absència d'*UBE3A*, no es poden ubiquitinar ni degradar les proteïnes diana per la via ubiquitina-proteosoma, la qual està implicada en la transducció de senyals, progressió del cicle cel·lular, reorganització del citoesquelet, reparació de l'ADN i regulació de la transcripció (39,56).

Funcions del gen *UBE3A*

Ubiquitinació de proteïnes

S'han identificat diferents substrats d'UBE3A com p53, p27, p18, pbl/ECT2, α -sinucleïnes, l'efexina-V, SK2 i BMAL1. Aquests substrats s'acumulen en les neurones en absència d'UBE3A i intervien en diverses funcions implicades en el neurodesenvolupament, entre les quals es troben: el creixement i migració neuronal, control de sinapsis excitatòries, densitat de les espines dendrítiques, aprenentatge, memòria, plasticitat sinàptica i regulació del ritme circadià (62,67–69).

Regulació del proteosoma

La funció d'UBE3A no es limita a la ubiquitinació de proteïnes per a què siguin degradades per la via-ubiquitina proteosoma, sinó que també controlaria l'activitat proteolítica del proteosoma, a través de la seva associació amb PSMD4, que és una subunitat reguladora del proteosoma. PSMD4 interviendria en el recanvi del proteosoma per autofàgia quan es troba ubiquitinat. S'ha proposat que sota determinades condicions, UBE3A regularia el recanvi per autofàgia del proteosoma a través de la interacció i ubiquitinació de PSMD4 (64,65). A més, s'han identificat 13 subunitats del proteosoma i proteïnes que interaccionen amb el proteosoma que són substrats d'UBE3A, entre elles Rpn10, Rng10 i DDI1. La ubiquitinació de DDI1 no induiria la seva degradació sinó que tindria una funció reguladora (66).

Complex HUN

En estudis de proteòmica s'ha observat que HERC2, UBE3A i NEURL4 formen el complex HUN, el qual s'ha suggerit que estaria implicat en la via de senyalització de mTORC1 i en la reorganització del citoesquelet d'actina, afectant a la plasticitat sinàptica i a la morfologia de les espines dendrítiques en neurones (70,71).

Regulació de la transcripció per UBE3A

De manera independent a la seva funció ubiquitina lligasa, UBE3A pot actuar com un coactivador de receptors nuclears d'hormones o regulant negativament l'expressió de gens induïts per l'hormona estradiol (72). Un dels gens induïts per l'hormona estradiol i que està regulat negativament per UBE3A és el gen *ARC*, que està implicat en el control dels nivells basals de receptors AMPAR (*α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid*). També s'ha observat que UBE3A podria regular els nivells de la proteïna *ARC* a través de la seva ubiquitinació directa (73,74). UBE3A també està implicat en la transcripció de receptors de glucocorticoides, quan

s'altera la via de senyalització d'aquests receptors es produeixen afectacions de l'eix hipotàlem-hipòfisi-suprarenal, resultant en una major susceptibilitat a l'estrès i a l'ansietat (75). En relació amb l'activitat transcripcional d'UBE3A, aquest interacciona amb la proteïna MCM6 que forma part del complex MCM, necessari per a la iniciació i l'elongació en la replicació de l'ADN (76). Aquest fet suggereix que UBE3A podria ser rellevant en la replicació, reparació i transcripció de l'ADN (70).

Models de ratolí SA

Els models de ratolí han contribuït a la comprensió de la patogènesi molecular de la SA i a la identificació de noves dianes terapèutiques pel seu tractament. El model de SA més utilitzat és *knockout* del gen $UBE3A^{m-/p+}$, generat a través de la deleció de l'exó 2 de l'al·lel matern del gen $UBE3A$. Els ratolins $UBE3A^{m-/p+}$ mostren una reducció del pes del cervell, atàxia, alteracions motores i un patró anormal en l'EEG. A més, també mostren defectes en l'aprenentatge, deteriorament de la memòria, reducció de la densitat d'espines dendrítiques i dèficits en la potenciació a llarg terme (*long-term potentiation*, LTP) en l'hipocamp (77,78). Els ratolins $UBE3A^{m-/p+}$ mostren una activitat reduïda de la proteïna calci-calmodulina quinasa II (CAMKII), a causa d'un augment en la seva fosforilació, que té un impacte en la inducció LTP i com a conseqüència en la memòria. El restabliment de l'activitat de CAMKII produeix una millora en el comportament i en l'aprenentatge (79,80). L'anàlisi de xarxes d'interacció evidencia que CAMKII i UBE3A es relacionarien a través de les proteïnes ASPP2 i HIF1AN i confluirien en una mateixa via molecular (70).

Els ratolins $UBE3A^{m-/p+}$ també presenten un augment en la senyalització de la neuroregulina-ErbB4 produint una plasticitat sinàptica anòmala i un deteriorament de la memòria. El dèficit en LTP pot ser rescatat per inhibidors d'ErbB introduïts directament en l'hipocamp de ratolins $UBE3A^{m-/p+}$ (81). Estudis de neurones de l'hipocamp de ratolins $UBE3A^{m-/p+}$ també han mostrat defectes en la guia dels axons neuronals, a causa d'anomalies en l'establiment d'adhesions focals (82).

El restabliment de la còpia materna en models de ratolí on es pot induir la expressió del gen $UBE3A$ ($Ube3a^{Stop/p+}; Cre^+$) mostren que hi ha finestres en el neurodesenvolupament on es poden rescatar algunes de les característiques clíniques. Els dèficits a nivell motor es rescaten en ratolins adolescents, mentre que l'ansietat, els comportaments repetitius i l'epilèpsia només es rescata durant les etapes primerenques del desenvolupament. La plasticitat sinàptica en l'hipocamp es pot rescatar a qualsevol edat. D'altra banda, s'ha descrit que la deleció de la còpia materna en ratolins induïbles ($Cre^{ERT}; Ube3a^{mlox/p+}$) a les 3 setmanes o a les 12 setmanes de vida, no té efecte

a nivell de comportament ni provoca una sensibilitat augmentada a l'epilèpsia. Es demostra que el gen *UBE3A* té un paper essencial en les etapes primerenques del desenvolupament del cervell, però més limitat en el cervell adult (83,84).

En la recerca de noves dianes terapèutiques, recentment, s'ha descrit que els ratolins *UBE3A*^{m-/p+} presenten un augment en els nivells de PTPA que és un activador de la proteïna fosfatasa 2 (PP2A). En el sistema nerviós PP2A és essencial en el creixement neuronal, diferenciació, assemblatge del citoesquelet, en la morfologia de les espines dendrítiques i en la plasticitat sinàptica. En ratolins *UBE3A*^{m-/p+}, PTPA no pot ser ubiquitinat ni degradat per la via ubiquitina proteosoma, augmentant el seus nivells i com a conseqüència produint un increment en l'activitat de PP2A (85,86). La inhibició de PP2A amb el fàrmac LB-100 rescata els dèficits en la morfologia de les espines dendrítiques i la transmissió de la sinapsi excitatòria, així com els dèficits en el comportament (87).

Interaccions d'*UBE3A* amb altres gens del diagnòstic diferencial de la SA

Algunes de les síndromes incloses en el diagnòstic diferencial de la SA estan causades per variants patogèniques en gens que interactuen o estan relacionats amb *UBE3A*. El gen *MECP2* causant de la síndrome de Rett (MIM # 312750) interacciona físicament amb *UBE3A* i actuen regulant a nivell transcripcional gens comuns, fet que explicaria perquè la SA i la síndrome de Rett presenten característiques clíniques solapants (88). D'altra banda, el gen *HERC2* responsable de la síndrome de deficiència de *HERC2* (MIM # 615516) interacciona físicament amb *UBE3A*, ja que *HERC2* també és una E3 ubiquitina lligasa que participa en la via de degradació de proteïnes ubiquitina-proteosoma. A més, *HERC2* i *UBE3A* formen part del complex proteic HUN, juntament amb *NEURL4* (70,89). El gen *TCF4* causant de la síndrome de Pitt-Hopkins (MIM # 610954) també s'ha relacionat amb *UBE3A*. En la línia cel·lular de neuroblastoma SH-SY5Y, el *knockdown* de *TCF4* provoca una disminució de l'expressió del gen *UBE3A*, suggerint que el gen *TCF4* actua regulant l'expressió d'*UBE3A* (90). Recentment, s'ha reportat que els gens *UBE3A* i *MEF2C* (síndrome *MEF2C* (MIM # 613443)) interaccionen físicament en estudis realitzats en *Drosophila Melanogaster*. S'ha suggerit que *UBE3A* regularia els nivells de *MEF2C* a través de la seva ubiquitinació i la subseqüent degradació en el proteosoma (91).

1.2.3 Correlació genotip-fenotip

Els pacients amb la SA presenten un major o menor grau de gravetat, depenent del mecanisme molecular. El fenotip més greu el presenten els pacients amb deleció respecte a la resta de classes

moleculars. S'ha suggerit que hi ha diferències fenotípiques entre els individus amb la deleció de tipus I (BP1-BP3) respecte als pacients amb la deleció de tipus II (BP2-BP3). Els individus amb deleció de tipus I presentarien major afectació del llenguatge, trets autistes i una epilèpsia més greu i de pitjor pronòstic que els pacients amb deleció de tipus II. No obstant, no hi ha diferències en la capacitat cognitiva i les habilitats adaptatives (92–95).

Els pacients amb deleció respecte als altres mecanismes moleculars presenten una major prevalença de TEA, absència de llenguatge o ús de paraules aïllades, major taxa d'epilèpsia, un patró en l'EEG més greu i a nivell cognitiu no progressen més dels 12-14 mesos (92,96–100). En la revisió portada a terme per Moncla et al., 1999 (101) en la qual comparen 20 pacients amb deleció amb 20 pacients sense deleció (disomia uniparental paterna, defecte d'impronta i variants patogèniques en el gen *UBE3A*), observen que els pacients sense deleció presenten microcefàlia en un 35%, respecte al 90% dels pacients amb deleció. Els pacients sense deleció són capaços de caminar de manera independent i un 70% presenten epilèpsia comparat amb el 90% dels pacients amb deleció. Els pacients sense deleció podien adquirir de 4 a 10 paraules i en quatre casos eren capaços d'utilitzar més de 20 paraules.

El fenotip dels pacients amb variants puntuals patogèniques i delecions intragèniques en *UBE3A* s'ha suggerit que és molt similar (41). Aquests individus comparats amb els pacients amb deleció presenten una puntuació millor a nivell cognitiu, en les habilitats motores i de llenguatge receptiu i expressiu. La prevalença de convulsions i microcefàlia és similar a la dels pacients amb deleció. El retard del desenvolupament dels individus amb variants puntuals patogèniques i delecions intragèniques en *UBE3A* és comparable amb el present en la disomia uniparental i el defecte de la impronta (97).

Finalment, els pacients amb disomia uniparental paterna i defecte de la impronta tenen un millor creixement físic, menor taxa d'anomalies del moviment i una menor prevalença d'epilèpsia, comparat amb els pacients amb deleció (101,102).

S'han descrit pacients amb SA amb un defecte de la impronta en mosaic, aquests individus en general presenten un fenotip més lleu o atípic que recorda a la SPW. En aquests pacients s'ha suggerit que com més alt és el percentatge de cèl·lules normals (1-40%) més lleu és el fenotip. Respecte a les diferents classes moleculars que conformen el SA clàssic, els pacients amb un defecte de la impronta en mosaic presenten: i) retard del desenvolupament, que només en un 20% dels casos és greu, ii) poden adquirir entre 100-1000 paraules i utilitzar frases, iii) l'atàxia

està present en un 33-41% dels individus, iv) l'adquisició de la marxa es situa als 21 mesos de mitjana, v) un 19% presenta microcefàlia, vi) l'EEG típic es troba present en un 50% dels individus, vii) entre el 5-20% presenta epilèpsia i viii) un 73% presenta problemes de la son. Pel que fa al fenotip conductual, presenten una major taxa d'ansietat (95%) i hiperactivitat (59%). Finalment, un dels trets característics dels SA en mosaic és la obesitat (64-76%) com a conseqüència d'una hiperfàgia, solapant-se amb la SPW i portant a què en alguns casos es diagnostiquin clínicament com SPW (103–106).

1.2.4 Guany de funció del gen *UBE3A*

La pèrdua de funció del gen *UBE3A* és responsable de la SA, mentre que les duplicacions de la regió 15q11-q13 en el cromosoma matern (dup15q11-q13), la presència d'un cromosoma extra isodivertent 15q11-q13 (idic(15) matern), la duplicació del gen *UBE3A* i la variant puntual T485A identificada en el gen *UBE3A*, produeixen un guany de funció, causant altres TND (107–111).

La síndrome dup15q11-q13 es caracteritza per una hipotonia, DI, retard del desenvolupament, afectació del llenguatge, epilèpsia i TEA. Els pacients amb idic(15) matern es troben més greument afectats que els portadors de la dup15q11-q13 (108). Estudis realitzats en cervells post-mortem d'individus que presentaven la dup15q11-q13 mostren un augment de l'expressió de transcrits del gen *UBE3A* i de la proteïna codificada per aquest gen (109), recolzant que el guany de funció del gen *UBE3A* és el principal mecanisme de patogenicitat. També s'ha descrit la duplicació completa del gen *UBE3A* en una pacient que presentava retard global del desenvolupament i TEA. La duplicació havia estat heretada per via materna i en la família hi havia casos d'individus amb dificultat d'aprenentatge, ansietat i depressió (110).

Finalment, la seqüenciació del genoma complet (WGS) en un pacient que presentava TEA va identificar la variant T485A en el gen *UBE3A*. La fosforilació de la Thr485 per la proteïna quinasa A (PKA) provoca una disminució de l'activitat d'UBE3A, inhibint la seva pròpia ubiquitinació, així com la dels seus substrats. La variant T485A impedeix la fosforilació d'UBE3A per part de la PKA, provocant un increment de l'activitat d'UBE3A que té com a conseqüència un augment en la densitat de les espines dendrítiques *in vivo* i canvis a nivell estructural en el cervell, els quals contribuirien a la patogènesi de l'autisme (111).

1.2.5 Diagnòstic diferencial de la SA

En el 10% dels individus amb diagnòstic clínic de SA, no s'identifiquen les alteracions moleculars reportades per la SA, aquests individus es classifiquen com SA-like. S'han descrit diferents

síndromes que poden estar causades per CNVs i/o per variants puntuals, que presenten característiques clíniques solapants amb la SA i que en alguns casos són les responsables del fenotip SA-like (Taula 2).

Taula 2| Resum de les síndromes incloses en el diagnòstic diferencial de la SA. Taula adaptada de l'article Tan et al., 2014 (112,113).

Síndrome	Gen	Característiques solapants amb la SA
Phelan-McDermid (MIM # 606232)	<i>SHANK3</i>	Retard del desenvolupament moderat-greu Absència de llenguatge o ús mínim de paraules Hipotonia Problemes d'alimentació Moviments masticatoris Hiperactivitat i dèficit d'atenció
Haploinsuficiència de <i>MBD5</i> (MIM # 156200)	<i>MBD5</i>	DI greu Afectació greu del llenguatge Epilèpsia Dèficit d'atenció
Koolen-de Vries (MIM # 610443)	<i>KANSL1</i>	Retard del desenvolupament lleu-moderat Hipotonia Epilèpsia Problemes d'alimentació Aparença felç
Pitt-Hopkins (MIM # 610954)	<i>TCF4</i>	DI greu Absència de llenguatge Marxa atàxica Aparença felç Hipotonia Boca gran i espai interdental augmentat
Christianson (MIM # 300243)	<i>SLC6A9</i>	DI greu-profunda Absència de llenguatge Microcefàlia postnatal Epilèpsia Marxa atàxica Hipotonia Atacs de riure freqüents
Mowat-Wilson (MIM # 235730)	<i>ZEB2</i>	DI moderada-greu Ús mínim de llenguatge expressiu Microcefàlia Hipotonia Epilèpsia Caminar de base ample amb els braços aixecats Moviments masticatoris
Kleefstra (MIM # 610253)	<i>EHMT1</i>	DI moderada-greu Afectació del llenguatge amb ús mínim de paraules Hipotonia Epilèpsia Problemes de la son Prognatisme
Deficiència de <i>HERC2</i> (MIM # 615516)	<i>HERC2</i>	Retard del desenvolupament moderat Afectació del llenguatge Caminar de base ample amb els braços aixecats Dèficit d'atenció Hipotonia

Taula 2 | Continuació.

Deficiència d'adenilsuccinasa (MIM # 103050)	<i>ADSL</i>	Retard del desenvolupament greu Afectació del llenguatge Microcefàlia Epilèpsia
Rett (MIM # 312750)	<i>MECP2</i>	DI greu Absència de llenguatge Microcefàlia postnatal Hipotonia Epilèpsia Problemes de la son Atacs de riure inapropiats
Encefalopatia epilèptica infantil primerenca, 2 (MIM # 300672)	<i>CDKL5</i>	DI greu Absència de llenguatge Hipotonia Problemes de la son
Haploinsuficiència de <i>FOXP1</i> (MIM # 613454)	<i>FOXP1</i>	DI greu Absència de llenguatge Microcefàlia postnatal Epilèpsia Hipotonia Problemes de la son
Duplicació <i>MECP2</i> (MIM # 300260)	<i>MECP2</i>	DI greu Ús mínim de paraules Hipotonia Epilèpsia
Haploinsuficiència de <i>MEF2C</i> (MIM # 613443)	<i>MEF2C</i>	DI greu Absència de llenguatge Epilèpsia Caminar de base ample Hipotonia
Alfa-talassèmia-DI lligada al cromosoma X (MIM # 301040)	<i>ATRX</i>	DI greu-profunda Afectació del llenguatge Microcefàlia Hipotonia Baveig Reflux gastroesofàgic
Encefalopatia epilèptica infantil primerenca, 4 (MIM # 612164)	<i>STXBP1</i>	Retard del desenvolupament greu Absència de llenguatge Hipotonia Atàxia
DI associada a WAC (MIM # 616708)	<i>WAC</i>	Retard del desenvolupament lleu a greu Afectació del llenguatge Problemes de la son Epilèpsia Hipotonia Problemes d'alimentació Boca ampla

1.3 Variació genètica i aplicació de les noves tecnologies de seqüenciació en els trastorns del neurodesenvolupament

L'any 1990 es va iniciar el Projecte Genoma Humà amb el propòsit de conèixer la seqüència de l'ADN del genoma humà. L'objectiu d'aquest projecte va ser determinar la posició dels 3×10^9 parells de bases (pb) que conformen l'ADN i identificar els 25.000 gens que estan presents. El

projecte va finalitzar l'any 2003 i va permetre establir la seqüència del genoma humà de referència, així com promoure el desenvolupament de noves tecnologies de seqüenciació (114). La variació genètica contribueix a les diferències fenotípiques interindividuais, entre poblacions i espècies. S'estima que la variació genètica entre dos humans és de l'1%, això implica que 1 de cada 1000 pb serà diferent entre dos individus (115). Les variacions genètiques es poden dividir en dos grups: les variacions que afecten a la seqüència de nucleòtids i les variants estructurals. Dins de les variacions en la seqüència s'inclouen els canvis d'un únic nucleòtid (SNVs) i les delecions o insercions (INDELS). Les variants estructurals consisteixen en reorganitzacions de material genètic i es classifiquen en dos grups depenent de si afecten o no a la dosi gènica: i) les CNVs que consisteixen en delecions, duplicacions, insercions i triplicacions i ii) les variacions equilibrades que no afecten a la dosi i on s'inclouen les translocacions equilibrades i les inversions (116,117).

Amb l'objectiu d'estudiar quina és la variació genètica en el genoma humà, s'han desenvolupat diferents projectes. L'any 2002 es va iniciar el projecte *HapMap* amb l'objectiu de mapar la diversitat d'haplotips en el genoma humà, a través de la caracterització de canvis d'un únic nucleòtid polimòrfic (SNPs) comuns (freqüència al·lèlica superior al 5%) en diferents poblacions humanes. Es va finalitzar el projecte amb la identificació d'un total de 8 milions de SNPs comuns en el genoma humà. L'any 2003 es va iniciar el projecte ENCODE (*Encyclopedia of DNA elements*) amb l'objectiu d'interpretar les dades obtingudes del projecte genoma humà. En la primera fase es va analitzar l'1% del genoma corresponent a les regions funcionals o que codifiquen per proteïnes. En la segona fase es va analitzar el 99% del genoma restant, incloent la quantificació dels diferents tipus d'ARN, es van mapar les regions codificants, les modificacions d'histones i els llocs d'unió de factors de transcripció i de metilació de l'ADN. En la darrera fase que va culminar l'any 2016, es van revisar els resultats obtinguts prèviament amb l'objectiu d'aplicar-los a estudis genòmics de malalties (118).

La incorporació de les noves tecnologies de seqüenciació (*Next Generation Sequencing*, NGS) ha permès caracteritzar la variació en el genoma humà a gran escala. El *1000 Genomes Project* es va iniciar l'any 2007 amb la finalitat d'estudiar la distribució de variants genètiques comunes i rares en el genoma humà. En aquest projecte s'han identificat un total de 88 milions de variants, d'aquestes 84,7 milions són SNPs, 3,6 milions INDELS i 60.000 variants estructurals (119). Finalment, convé ressaltar el projecte *100.000 Genomes* que es va iniciar a Gran Bretanya l'any 2013 amb l'objectiu d'implementar la medicina genòmica en el sistema nacional de salut (NHS).

Dins d'aquest projecte es distingeixen dos programes: i) un per identificar la causa genètica de malalties rares pediàtriques i ii) un programa de càncer per determinar l'etiologia d'alguns tipus concrets de càncer, el pronòstic i avaluar la resposta a potencials tractaments. Els resultats preliminars indiquen que s'ha pogut establir un diagnòstic en un 25-50% dels individus amb malalties rares pediàtriques (120,121).

1.3.1 Evolució de les tecnologies de seqüenciació

La tecnologia anomenada de primera generació consisteix en l'automatització portada a terme per *Applied Biosystem Instrument* (ABI) del mètode de seqüenciació descrit l'any 1977 per Sanger et al., (122). Posteriorment, a partir del projecte genoma humà, es van desenvolupar noves tecnologies de seqüenciació anomenades tecnologies de segona generació (NGS) o tecnologies de seqüenciació massiva en paral·lel (MPS). Aquestes tecnologies permeten la seqüenciació de molts fragments d'ADN a la vegada, reduint el temps i el cost de la seqüenciació (123). A partir de l'any 2010, es van començar a desenvolupar les tecnologies de tercera generació, que a diferència de les de segona generació permeten la seqüenciació d'una única molècula d'ADN sense necessitat d'amplificació. Això es tradueix en la generació de *long-reads* tenint avantatges respecte a les tecnologies de segona generació: i) millor assemblatge de genomes, ii) és més fàcil mapar els fragments seqüenciats en el genoma de referència, iii) permet diferenciar entre isoformes i iv) avenços en la identificació de variants estructurals. Tenint en compte aquestes avantatges, s'anticipa que en els pròxims anys s'implementaran les tecnologies de tercera generació en el camp del diagnòstic genètic (124,125).

En relació amb les aproximacions diagnòstiques que es poden realitzar a través de la tecnologia de NGS es diferencien: i) els panells de gens, que corresponen a la seqüenciació d'un conjunt de gens determinat associats a una patologia concreta, ii) la seqüenciació de l'exoma complet (WES), que inclou totes les regions codificants (1-2% del genoma) i iii) la seqüenciació del genoma complet (WGS). Les principals avantatges del panell de gens respecte al WES i al WGS és que permet obtenir una cobertura més elevada dels gens estudiats, minimitza la identificació de variants de significat incert (VUS), redueix la probabilitat d'identificar troballes secundaries, així com disminuir el cost i el temps de resposta. Per contra, la utilització del panell de gens implica la necessitat de realitzar actualitzacions periòdiques perquè es van descobrint nous gens associats a malalties i el re-anàlisi de les dades té un ús limitat en els casos sense diagnòstic molecular (126,127).

El WES a diferència dels panells de gens, inclou totes les regions codificants del genoma, permetent analitzar gens que encara no han estat associats a una malaltia determinada i per tant, la identificació de noves variants puntuals. L'estratègia del re-anàlisi de les dades introduint canvis en la *pipeline* bioinformàtica al llarg del temps en els pacients sense diagnòstic és molt eficaç i incrementa el número de diagnòstics (128,129). Entre els inconvenients es troben que la cobertura de les regions analitzades és inferior que en els panells de gens, el volum de dades a analitzar és superior, existint una major probabilitat de detectar troballes secundàries i augmentant la identificació de VUS (130–132). Tant els panells de gens com el WES tenen limitacions a l'hora de detectar variants en regions repetitives, variants estructurals, INDELS i variants en el genoma mitocondrial (127,133).

En darrer lloc, el WGS permet seqüenciar el 98% del genoma humà amb una cobertura més uniforme i millors paràmetres de qualitat de seqüenciació respecte el WES. El WGS és la tecnologia de seqüenciació més eficient per detectar tant CNVs com altres variants estructurals, ja que cobreix tots els punts de trencament. Permet la detecció de variants en ARNs no codificants, expansions de *short tandem repeats* i variants en regions codificants que no es troben ben cobertes en el WES. Així mateix, es poden identificar variants en regions intròniques profundes (*deep intronic*) i en regions reguladores. Entre les desavantatges respecte al WES es troben: i) el cost és més elevat, ii) es necessita una gran capacitat d'emmagatzematge de les dades i iii) l'anàlisi de les dades és més complex (134–137).

1.3.2 Aplicació de les noves tecnologies de seqüenciació en els trastorns del neurodesenvolupament

La implementació de la tecnologia de NGS en el diagnòstic dels TND caracteritzats per una gran heterogeneïtat genètica i variabilitat fenotípica ha permès un increment en la taxa de diagnòstic, així com el descobriment de nous gens implicats en aquests trastorns. A dia d'avui, les guies clíniques consideren el WES com a primer test diagnòstic per als individus amb TND, essent especialment útil en formes atípiques de síndromes ja conegudes, malalties rares, trastorns genètics que no presenten trets clínics distintius i en la identificació de nous gens (138,139).

El principal projecte que ha demostrat la utilitat diagnòstica de l'aplicació del WES en cohorts amb TND ha estat *Deciphering Developmental Disorders* (DDD). S'ha portat a terme a Gran Bretanya, on han reclutat unes 14.000 famílies amb trastorns del desenvolupament (TD) greus i defectes congènits. Aquest projecte va ser capdavanter en la implementació de WES en trios en

casos esporàdics aportant una taxa elevada de diagnòstics i el descobriment d'un gran nombre de nous gens implicats en TND.

En una primera fase, l'aplicació del WES en 1.133 trios va permetre identificar SNVs responsables del fenotip present en el 27% dels individus (140). El re-anàlisi de les dades de WES introduint millores en la *pipeline* bioinformàtica va resultar en un 13% de diagnòstics addicionals, que donen lloc a una taxa total de diagnòstic del 40%. Dins del 40% de diagnòstics, el 80% estan causats per variants *de novo* (141). Posteriorment, s'han realitzat diferents anàlisis de les dades de WES de la cohort d'estudi i s'han incorporat dades de nous pacients en col·laboració amb altres grups d'Europa i Estats Units, fet que ha suposat el descobriment de nous gens i ha permès establir nous diagnòstics.

L'anàlisi de l'enriquiment de variants *de novo*, estima que aproximadament un 42% de la cohort d'estudi (4.293 individus) és portadora de variants patogèniques *de novo* en regions codificants, la meitat de les quals són variants de pèrdua de funció. D'aquests resultats s'ha relacionat l'edat paterna avançada amb un augment en la prevalença dels TD (142). D'altra banda, la seqüenciació dirigida d'elements reguladors en individus amb TD greus sense diagnòstic per WES, ha permès detectar en un 1-3% dels individus variants patogèniques *de novo* en elements reguladors que es troben actius en el cervell fetal (143).

En canvi, l'anàlisi de la contribució de variants recessives als TD indica que només el 3,6% dels casos són atribuïbles a variants recessives en població europea. Per contra, en individus amb ancestres del Pakistan, és del 31% degut a l'alta freqüència de consanguinitat en la població paquistanesa. En més de la meitat dels casos, els gens identificats com a responsables de TD, ja havien estat descrits prèviament (144).

Cal destacar que l'anàlisi de les dades de WES en TND greus ha demostrat que variants genètiques comunes que influeixen en trets del neurodesenvolupament (com són intel·ligència, autisme, circumferència del cap i el pes al néixer) en la població general afecten també la presentació i gravetat dels TND greus, causats o no, per variants monogèniques. (6).

1.3.2.1 Heterogeneïtat genètica i variabilitat fenotípica en els trastorns del neurodesenvolupament

La gran variabilitat fenotípica i comorbiditat en els TND, suposa un gran repte pel seu diagnòstic clínic i molecular ja que molts símptomes o característiques clíniques no són úniques per a un

trastorn o una síndrome concreta sinó que són comuns a més d'una (5). Variants en un mateix gen poden donar lloc a diferents fenotips (Taula 3), així com diferents gens poden estar associats a un mateix fenotip (heterogeneïtat genètica).

Taula 3 | Número de gens associats a un o més fenotips. Dades extretes d'OMIM (www.omim.org).

Base de dades OMIM (Actualitzada Juny 2020)	
Gens associats a 1 fenotip	2971
Gens associats a 2 fenotips	791
Gens associats a 3 fenotips	278
Gens associats a 4 o més fenotips	241

El fet de què un mateix gen pugui estar implicat en diferents fenotips depèn de com afecten les variants genètiques a la funció de la proteïna (p.e guany de funció versus pèrdua de funció). Un exemple, és el gen *SMARCE1*, variants *missense* s'associen a la síndrome de Coffin-Siris (MIM # 616938), mentre que variants de pèrdua de funció (*nonsense* i *frameshift*) són responsables de meningiomes espinals i craneals i no presenten les característiques clíniques de Coffin-Siris (145,146). Un altra bona mostra és el gen *CREBBP*, variants de pèrdua de funció causen la síndrome de Rubinstein-Taybi (MIM # 180849), mentre que variants *missense* en els dos últims exons (30 i 31) s'associen amb la síndrome Menke-Hennekam (MIM # 618332) (147).

En relació a l'heterogeneïtat genètica, l'aplicació de WES ha permès identificar nous gens responsables de pacients amb característiques clíniques solapants amb síndromes ja conegudes, anomenades síndromes-*like*. Un exemple és la síndrome de Kleefstra, causada per la pèrdua de funció del gen *EHMT1*. L'estudi tant per WES en trios com la seqüenciació dirigida de gens implicats en la remodelació de la cromatina de pacients que presenten les característiques clíniques de la síndrome però sense variants en el gen *EHMT1*, ha permès la identificació dels gens *KMT2C*, *SMARCB1*, *MBD5* i *NR1I3* com a responsables del fenotip Kleefstra-*like* (148,149).

De forma similar, l'aplicació del WES en pacients Rett-*like* ha identificat nous gens responsables (*SLC35A2*, *STXBP1*, *ZNF238*, *EEF1A2*, *EIF2B2*, *SHROOM4*, *SCN2A*, *KCNQ2*, *MEF2C*, *SYNGAP1*, *KIF1A* i *GRINI*), així com nous gens candidats (*RHOBTB2*, *SMARCA1*, *GABBR2*, *EIF4G1* i *HTT*) (150–152). En la síndrome Smith-Magenis-*like* també s'han detectat variants en nous gens (*IQSEC2*, *DEAF1*, *NDN*, *KAT5*, *BAP1* i *MAPK8IP3*) com a responsables d'aquest fenotip (153).

Com ja s'ha mencionant anteriorment, la variabilitat en la presentació i expressió de les característiques clíniques d'un individu, depèn també de la seva càrrega genètica individual. Existeixen diferents mecanismes d'interacció o modificació genètica en el que una variant genètica responsable d'un fenotip determinat pot interaccionar amb altres variants rares i freqüents localitzades en altres gens o bé en el mateix gen, actuant com a modificadores. Per exemple, en la deleció 16p11.2 s'ha observat que la càrrega de variants genètiques rares modula la variabilitat clínica intra i inter-familiar (154). Els factors ambientals també s'ha observat que tenen un paper important en l'expressió de determinats trets fenotípics (155).

1.4 Classificació de les variants genètiques segons la guia ACMG

L'aplicació de les noves tecnologies de seqüenciació en els laboratoris de diagnòstic ha fet que s'incrementi de forma exponencial la detecció de noves variants de seqüència en gens associats a trastorns genètics. La interpretació clínica d'aquestes variants és fonamental per un diagnòstic genètic correcte. Per aquest motiu, l'*American College of Medical Genetics and Genomics (ACMG)* juntament amb l'*Association for Molecular Pathology (AMP) and the College of American Pathologists (CAP)* han publicat una guia on es revisen els criteris per a la interpretació de variants de seqüència i es descriuen recomanacions per a la seva classificació. Els criteris utilitzats es basen en diferents tipus d'evidències com són: dades poblacionals, dades computacionals, dades funcionals i dades de segregació, entre d'altres (156).

Les variants de seqüència es poden classificar en 5 categories (157), on la variant és assignada com:

- Patogènica**. Contribueix de manera directa al desenvolupament de la malaltia. En el cas de les malalties amb una herència autosòmica recessiva o lligada al cromosoma X, s'ha de considerar que només una variant patogènica pot no ser suficient per causar la malaltia.
- Probablement patogènica**. Existeix una alta probabilitat (>90%) de què sigui la responsable de la malaltia. Es necessiten evidències addicionals que demostrin la patogenicitat de la variant.
- Significat incert**. No hi ha informació suficient que permeti discernir si es tracta d'una variant patogènica o benigna.
- Probablement benigna**. Existeix una alta probabilitat (>90%) de què sigui una variant benigna i no responsable de la malaltia. Es necessiten evidències addicionals que demostrin que és una variant benigna.
- Benigna**. La variant no és la responsable de la malaltia.

Per classificar una variant dins d'una de les 5 categories esmentades anteriorment, s'ha de tenir en compte la combinació i el pes dels diferents criteris que permeten classificar una variant com a patogènica o benigna. Entre els criteris molt sòlids i sòlids de patogenicitat es troben: i) la pèrdua de funció (p.e *nonsense*, *frameshift* i de *splicing*) en un gen haploinsuficient, ii) un canvi d'aminoàcid ja descrit anteriorment com a causant de malaltia, iii) l'herència *de novo*, iv) estudis funcionals *in vivo* o *in vitro* que demostren l'efecte deleteri de la variant sobre la proteïna i v) una prevalença de la variant superior en individus afectats comparat amb individus control. Els criteris sòlids que permeten classificar una variant com a benigna inclouen: i) una freqüència superior al 5% en població control, ii) una freqüència al·lèlica superior a l'esperada per un trastorn concret, iii) estudis funcionals *in vivo* i *in vitro* que demostren que la variant no té un efecte deleteri sobre la funció de la proteïna i iv) no hi ha segregació de la variant en altres individus de la família afectats. En la Figura 3 i la Taula 4, es troben resumits els diferents criteris i la combinació per a la classificació de variants.

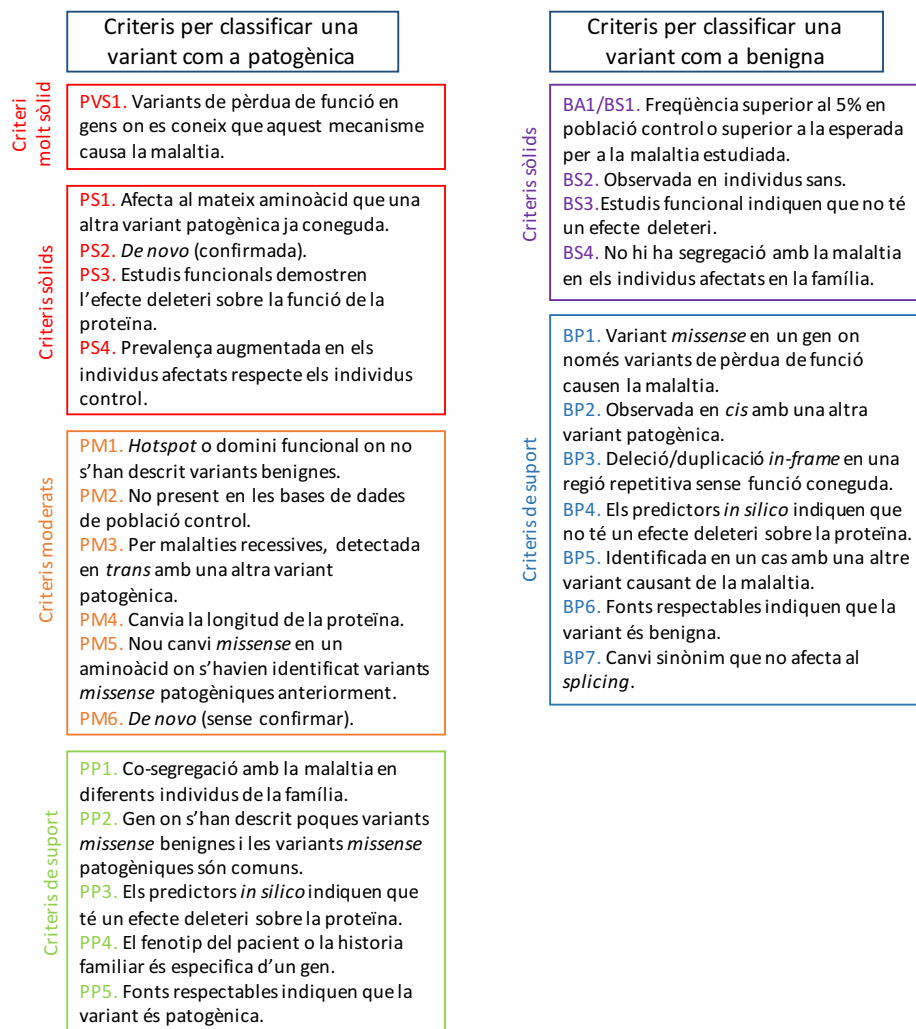


Figura 3 | Críteris per a la classificació d'una variant com a patogènica o benigna seguint la guia ACMG.

Taula 4 | Criteris combinats per a la classificació de variants genètiques. Taula adaptada de Richards et al., 2015 (156).

Classificació	Combinació de criteris
Patogènica	(i) 1 molt sòlid (PSV1) i <ul style="list-style-type: none"> a) ≥ 1 sòlids (PS1-PS4) o b) ≥ 2 moderats (PM1-PM6) o c) 1 moderat (PM1-PM6) i 1 de suport (PP1-PP5) o d) ≥ 2 de suport (PP1-PP5) (ii) ≥ 2 sòlids (PS1-PS4) o (iii) 1 sòlid (PS1-PS4) i <ul style="list-style-type: none"> a) ≥ 3 moderats (PM1-PM6) o b) 2 moderats (PM1-PM6) i ≥ 2 de suport (PP1-PP5) o c) 1 moderat (PM1-PM6) i ≥ 4 de suport (PP1-PP5)
Probablement patogènica	(i) 1 molt sòlid (PSV1) i 1 moderat (PM1-PM6) o (ii) 1 sòlid (PS1-PS4) i 1-2 moderats (PM1-PM6) o (iii) 1 sòlid (PS1-PS4) i ≥ 2 de suport (PP1-PP5) o (iv) ≥ 3 moderats (PM1-PM6) o (v) 2 moderats (PM1-PM6) i ≥ 2 de suport (PP1-PP5) o (vi) 1 moderat (PM1-PM6) i ≥ 4 de suport (PP1-PP5)
Benigna	(i) 1 sòlid BA1 o (ii) ≥ 2 sòlids (BS1-BS4)
Probablement benigna	(i) 1 sòlid (BS1-BS4) i 1 de suport (BP1-BP7) o (ii) ≥ 2 suport (BP1-BP7)
De significat incert	(i) No compleix els criteris anteriors (ii) Els criteris de classificació benignes i patogènics són contradictoris

Degut a la complexitat en la classificació de les variants seguint aquests criteris, s'han desenvolupat eines, com InterVar (158) i Varsome (159), que seguint les recomanacions de la guia ACMG, faciliten la classificació de les variants, ja que obtenen de forma automàtica la informació necessària de bases de dades, programes de predicció *in silico*, etc. No obstant, l'usuari ha de revisar els diferents criteris que compleix la variant i modificar o afegir els que siguin necessaris, com l'herència o la segregació familiar, així com si s'han realitzat estudis funcionals de variants.

1.5 Impacte funcional de les variants genètiques

L'impacte funcional d'una variant es pot predir tenint en compte el tipus de variant: les variants truncants (*nonsense*, *frameshift* i en alguns casos de *splicing*) produeixen amb una alta freqüència una pèrdua de funció, degut a què el mecanisme *Nonsense-mediated mRNA decay* (NMD) degrada els transcrits amb codons de stop prematurs. En el cas de les variants *missense* és molt difícil predir l'efecte sobre la funció de la proteïna, per aquest motiu s'han desenvolupat programes bioinformàtics que ajuden a predir el efecte potencial sobre la proteïna i així poder classificar-les. No obstant, els models funcionals *in vitro* i *in vivo* són els que realment mostren les conseqüències sobre la funció de la proteïna i permeten determinar si una variant és patogènica.

1.5.1 Mecanismes de patogenicitat

Existeixen principalment tres mecanismes de patogenicitat: pèrdua de funció, guany de funció i dominant negatiu.

Mecanisme de pèrdua de funció

El mecanisme de pèrdua de funció està produït principalment per variants que trunquen la proteïna, com és el cas de les variants *nonsense*, *frameshift* i de *splicing* (160). No obstant, les variants *missense* també poden produir una pèrdua de funció, ja què poden afectar negativament a l'activitat, el plegament o l'estabilitat de la proteïna, (161). Dins del mecanisme de pèrdua de funció es poden distingir dos tipus de variants:

- Les variants **amorfes** donen lloc a una pèrdua de funció completa, no hi ha activitat de la proteïna codificada per aquell al·lel amb la variant.
- Les variants **hipomorfes** produeixen una pèrdua parcial de la funció de la proteïna com a conseqüència d'una disminució en la transcripció o una proteïna amb una funció disminuïda.

S'han descrit tres classes de gens: i) els que toleren canvis que trunquen la proteïna, ii) gens recessius que toleren canvis truncants en heterozigosi i iii) gens haploinsuficients. Un indicador de què un gen és intolerant a canvis de pèrdua de funció és el pLI, que és un *score* de probabilitat que valora el número de variants truncants observades i les esperades en un determinat gen. Com més proper és el pLI a 1 més intolerant és el gen a variants de pèrdua de funció. També hi ha un indicador (*z-score*) que avalua les variants *missense* esperades i les observades per un determinat gen, permetent valorar si un gen és intolerant a variants *missense* (160).

Guany de funció

El mecanisme de guany de funció pot estar causat per: i) un increment en la producció d'una proteïna a causa d'una duplicació/amplificació, un control traduccional incorrecte o un augment en l'estabilitat de l'ARN/proteïna o ii) un increment en l'activitat de la proteïna (162,163).

Es poden diferenciar dos tipus de variants que actuen per un mecanisme de guany de funció:

-Les variants **hipermorfes** donen lloc a una major activitat pròpia de la proteïna.

-Les variants **neomorfes** produeixen una proteïna amb una funció diferent a la de la proteïna *wild type* (WT).

Dominant negatiu

Aquest mecanisme també conegut com variants antimorfes, consisteix en la producció d'una proteïna que interfereix en l'activitat de la proteïna WT. Quan les variants antimorfes es troben en heterozigosi, redueixen l'activitat de la proteïna codificada per l'al·lel WT. Un exemple clar són els complexos multimèrics de proteïnes, on diferents proporcions de proteïnes mutades i WT, alteren la funció de tot el complex (162).

1.5.2 Mecanisme de qualitat post-transcripcional *Nonsense-Mediated*

mRNA Decay (NMD)

Nonsense-mediated mRNA Decay (NMD) és un mecanisme de vigilància de l'ARNm que es troba en les cèl·lules eucariotes i que evita la traducció de proteïnes truncades. La seva funció principal és fer que el procés de producció de proteïnes sigui eficient per a les cèl·lules i evitar la producció de proteïnes que puguin tenir un potencial efecte deleteri, siguin inestables o que no es puguin plegar correctament. Aquest mecanisme actua en l'espai de temps entre la transcripció i la traducció (164,165). El principal model que explica com els transcrits amb codons de stop prematurs són reconeguts i activen el mecanisme NMD es basa en els complexos EJC (*Exon-Junction Complex*). Aquests complexos estan implicats en el *splicing* del pre-ARNm. A mesura que el ribosoma va avançant desplaça els complexos EJC, però quan hi ha un codó stop prematur el ribosoma s'atura. Com a conseqüència, els complexos EJC que no són desplaçats pel ribosoma activen el mecanisme NMD reclutant el factor Upf1 que s'uneix als factors Upf2 i Upf3. Upf1 és fosforilat per la proteïna quinasa SMG1, desencadenant el reclutament d'altres factors i quinases que acaben eliminant el casquet 5' i deadenilant l'extrem 3' del ARNm, provocant l'exposició dels seus extrems a la degradació per exonucleases (Figura 4) (166–169).

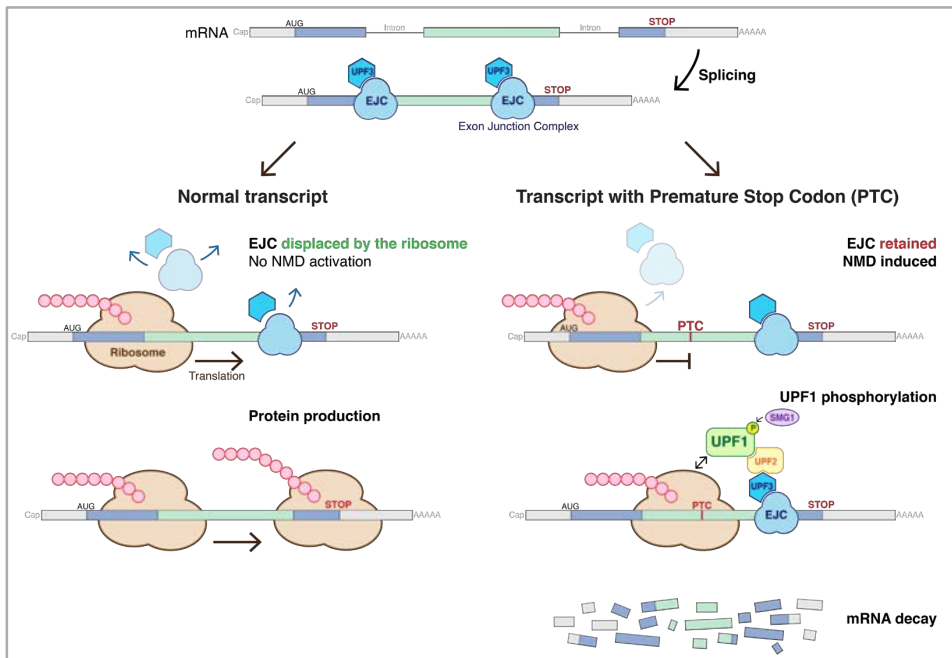


Figura 4 | Model que mostra l'activació del mecanisme de control post-transcripcional *Nonsense-Mediated mRNA Decay*. Imatge obtinguda de Marius Walter - Own work. PTC, *Premature stop codon*.

Les variants *nonsense* i *frameshift* que introdueixen codons de stop prematurs que es troben a >50-55 nucleòtids *upstream* de la unió exó-exó situada més en l'extrem 3', desencadenen el mecanisme NMD. Quan el codó *stop* prematur es troba situat en l'últim exó o a <50-55 nucleòtids de l'última unió exó-exó, escaparien del mecanisme NMD (170) (Figura 5). Els gens que no tenen introns són immunes al mecanisme NMD (171).

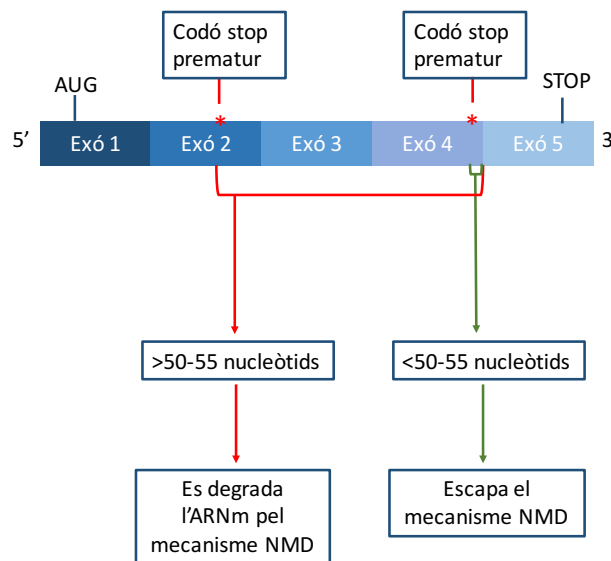


Figura 5 | Activació del mecanisme *Nonsense-Mediated mRNA Decay* (NMD) segons la localització del codó stop prematur. Els exons estan representats com rectangles en blau i els asteriscs en vermell indiquen codons de stop prematurs.

Hi ha algunes excepcions que provoquen la inhibició del mecanisme NMD com són els codons stop prematurs localitzats en exons de més de 400 nucleòtids o a <150 nucleòtids del codó d'inici (172,173). Aquestes excepcions explicarien la variabilitat en l'activació del mecanisme NMD observada entre diferents variants i gens en humans (173,174).

L'eficiència del mecanisme NMD és variable, essent un factor important que contribueix a la diversitat fenotípica observada entre diferents individus. Entre els factors que poden influenciar el mecanisme NMD es troben SNPs en gens de la maquinaria de NMD, competició amb altres mecanismes en els que intervenen les mateixes proteïnes i factors ambientals com l'estrès cel·lular, la sobreexpressió d'oncogens o la hipòxia que inhibeixen l'activitat d'aquest mecanisme (175–177).

Alguns dels gens que estan implicats en el mecanisme NMD han estat associats a TND. En concret, variants patogèniques en el gen *UPF3B* i delecions en heterozigosi que inclouen el gen *UPF2*, que són dos gens que formen part del complex EJC, s'han associat a DI. Els individus amb variants de pèrdua de funció en el gen *UPF3B* també presenten autisme, esquizofrènia i/o dèficit d'atenció i hiperactivitat (178–180).

1.5.3 Estudis funcionals de variants

Un dels mètodes més robusts per avaluar la patogenicitat d'una variant són els estudis funcionals. Aquests estudis es poden realitzar tant *in vivo* en models animals, generant la variant que presenta el pacient i comprovar quins són els efectes fenotípics, així com *in vitro*, utilitzant línies cel·lulars. El model *in vivo* més utilitzat és el ratolí, degut a què comparteix entre un 95-98% del genoma amb els humans, es reproduïxen ràpidament, són cost-efectius i permeten manipular el genoma de forma precisa. Per contra, no es poden avaluar certs aspectes fenotípics, com per exemple, l'afectació del llenguatge (181). D'altra banda, dins dels models *in vitro*, la reprogramació de cèl·lules mare embrionàries humanes (hESC) o de cèl·lules mare pluripotents induïdes (hiPSC), ha permès la generació i l'estudi de neurones (182). Els models *in vitro* també presenten inconvenients atès que hi ha molta heterogeneïtat entre els clons de iPSC i al tractar-se de cultius en 2D falten connexions amb altres neurones per donar lloc a estructures més complexes. Considerant aquestes limitacions, s'han desenvolupat protocols per a la generació d'organoides de cervell, que són cultius cel·lulars en 3D que donen lloc a estructures cerebrals i permeten estudiar el desenvolupament del cervell *in vitro* (183).

Les variants *missense* són particularment difícils d'interpretar, ja que poden produir un ampli espectre d'efectes. L'alteració de la seqüència d'aminoàcids d'una proteïna pot modificar la seva estructura afectant al plegament, l'estabilitat, la solubilitat i/o la unió a cofactors i a altres proteïnes i com a conseqüència afectar la seva activitat (184). A nivell fenotípic poden donar lloc a diferents presentacions clíniques amb una expressivitat variable i penetrància incompleta (185,186).

Els estudis funcionals *in vivo* i *in vitro* són un criteri sòlid per a la classificació d'una variant com a patogènica o benigna. Considerant que en molts casos l'abordatge d'aquests estudis és difícil s'han desenvolupat diferents assajos que permeten generalitzar i aplicar-se a diferents gens. Entre ells es troben assajos basats en el rescat d'una via cel·lular, en l'activació transcripcional de gens diana i localització cel·lular (187). Assajos que prediuen l'impacte de variants sobre l'ARNm (en el seu processament o en l'estabilitat) poden ser de molta utilitat quan es valoren variants que afecten el splicing, en regions no codificants, regions no traduïdes (5' i 3'UTR) o variants *deep intronic* (156).

Els programes de predicció *in silico* s'han desenvolupat com una eina per valorar el potencial efecte deleteri de variants genètiques sobre la funció de la proteïna (188). Hi ha dos tipus de programes: i) els de primera generació que es basen en algorismes que valoren la conservació, homologia de seqüència i estructura de la proteïna i ii) els de segona generació basats en *machine-learning*, que parteixen d'un conjunt de dades que serveixen per entrenar el programa (*training dataset*) i incorporen la informació de les noves dades que es van introduint (187). Segons la guia ACMG es recomana utilitzar diferents programes de predicció *in silico* i no utilitzar aquestes prediccions com a una única evidència per classificar una variant com a patogènica, probablement patogènica o benigna (156).

1.6 Quinesina KIF1A

Les quinesines són una família de proteïnes motores dependents d'ATP que intervien en el transport intracel·lular de molècules i orgànuls a través dels microtúbuls. Aquestes proteïnes contribueixen al transport, l'organització i el manteniment del citoesquelet, així com en la segregació del material genètic durant la mitosi i la meiosi cel·lular (189,190). Es caracteritzen per presentar un domini motor globular que està implicat en la hidròlisi de l'ATP i en la unió a microtúbuls, una regió filamentosa en forma de tija (*stalk region*) que utilitzen per dimeritzar i una regió on s'uneixen les vesícules que han de transportar anomenada cua (191–193) (Figura 6).

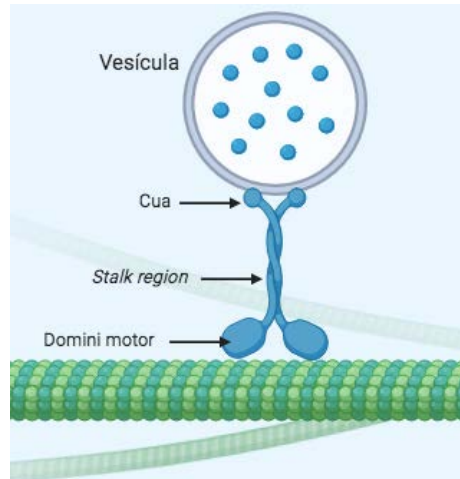


Figura 6 | Estructura general d'una quinesina en forma dimèrica unida a la vesícula per la regió de la cua i amb el domini motor unit a microtúbuls.

Les quinesines són proteïnes altament conservades entre les diferents espècies, existeixen diferents tipus dins d'un mateix organisme, suggerint que han divergit durant l'evolució per portar a terme diferents funcions cel·lulars. Es poden classificar en 14 famílies diferents d'acord amb els resultats d'anàlisi filogenètics, cadascuna de les famílies té característiques específiques i porten a terme funcions diferents (194). Segons on es localitza el domini motor les quinesines es poden agrupar en: i) en les N-quinesines es troba localitzat en l'extrem N-terminal, ii) en les M-quinesines en el centre de la proteïna i iii) les C-quinesines el tenen en l'extrem C-terminal. La localització del domini motor determina la direccionalitat del moviment. De les 14 famílies, només la família 13 està formada per M-quinesines i la família 14 per C-quinesines, la resta de famílies són N-quinesines (195).

La quinesa estudiada en aquesta tesi doctoral és la N-quinesina KIF1A, que pertany a la família 3 de les quinesines.

Estructura de KIF1A

L'estructura de KIF1A es pot dividir en: el domini motor, la regió *neck-coil*, el domini CC1-FHA-CC2 i el domini de plecstrina.

El domini motor es troba localitzat en l'extrem N-terminal i està format per 7 hèlixs α i 10 fulles β . Entre la fulla β_3 i la hèlix α_2 es situa una regió conservada anomenada *P-loop* que constitueix la butxaca d'unió a l'ATP. Les regions *switch I* (*loop* L9) i *switch II* (*loops* L11, L12 i L13) estan implicades en els canvis conformacionals que tenen lloc en la hidròlisi de l'ATP i l'alliberament del fòsfat. Les regions que intervenen en la unió dels microtúbuls estan formades pel *loop* L8 i pels *loops* L11, L12 (en KIF1A anomenat *K-loop*) (Figura 7A,B)(195,196).

La regió *neck-linker* uneix el domini motor amb el primer domini *coiled-coil* (anomenat *Neck-Coil*, NC). El domini *neck-coil* és molt flexible, experimentant diversos canvis conformacionals en diferents estats d'unió a ATP o ADP i està implicat en la dimerització i la motilitat de KIF1A (197). Els dominis CC1-FHA-CC2 tenen funcions reguladores sobre la pròpia activitat de KIF1A. D'una banda, CC1 regula la motilitat de KIF1A través d'interaccions intramoleculares amb la regió *neck-coil*. CC1 i FHA (*Forkhead-Associated*) controlen l'activació i dimerització de KIF1A. D'altra banda, els dominis FHA i CC2 inhibeixen l'activitat de KIF1A impedit la seva unió als microtúbuls (198,199). Finalment, en l'extrem C-terminal de la proteïna es localitza el domini homòleg de plecstrina (PH) que intervé en la unió de les vesícules, interaccionant de manera específica amb el fosfatidilinositol 4,5-bifosfat (PIP₂) present en la membrana de les vesícules (200) (Figura 7).

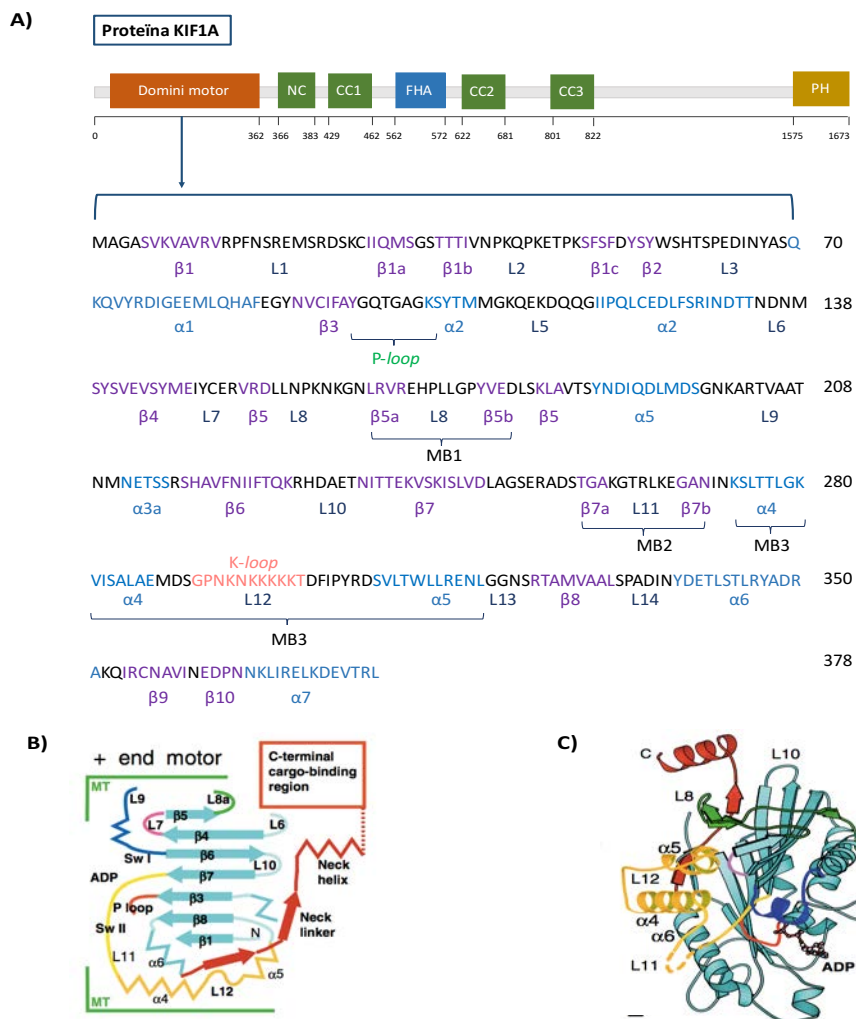


Figura 7 | Estructura de la quinesina KIF1A. **A)** Dominis de la proteïna KIF1A i estructura del domini motor on es troben indicat les hèlixs α , fulles β , els *loops* i les regions d'unió als microtúbuls. **B)** Visió esquemàtica de la organització de la proteïna KIF1A. **C)** Estructura tridimensional de la proteïna KIF1A en la conformació amb ADP unit. CC, *Coiled Coil*; FHA, *forkhead-associated*; PH, *pleckstrin homology domain*; MB1, *Microtubule binding region 1*; MB2, *Microtubule binding region 2*; MB3, *Microtubule binding region 3*; MT, *microtubule*; SW I, *Switch I*; SW II, *Switch II*. Les imatges B i C estan extretes de l'article de Amos and Hirose, 2007.

Regulació de l'activitat de KIF1A

En absència de vesícules que transportar la quinesina KIF1A roman inactiva en el citoplasma evitant la despesa innecessària d'ATP. S'han proposat dos models que expliquen com s'inhibeix l'activitat del domini motor de KIF1A. El primer model explicaria la inhibició de KIF1A en estat monomèric, on es produiria el plegament de la regió de la cua que interaccionaria amb el domini motor inhibint-lo. Es basa en estudis realitzats en la quinesina KIF5, en els últims anys, aquest model ha rebut un suport experimental creixent de models de quinesines implicades en funcions diverses com la polaritat epitelial, el transport intraflagel·lar i la mitosi (201–203). L'autoinhibició permet una regulació espacial i temporal precisa dels motors i es pot alliberar mitjançant la unió de les vesícules, fosforilació o altres mecanismes. En relació al segon model, inicialment, s'havia proposat que el mecanisme d'autoinhibició era alliberat per la unió de les vesícules a la regió de la cua, fet que induïa la dimerització i activació de KIF1A (204). Posteriorment, s'ha demostrat que la dimerització no és suficient per a l'activació i que KIF1A existeix en forma dimèrica però inactiva en el citoplasma. El mecanisme pel qual la forma dimèrica romandria inactiva s'ha suggerit que seria a través d'interaccions intramoleculares entre regions de la proteïna que no pertanyen al domini motor. La unió de les vesícules a la regió de la cua permetria alliberar el mecanisme d'autoinhibició i activar la proteïna (198).

Funcions de KIF1A

KIF1A és una quinesina específica de cervell que transporta vesícules pre-sinàptiques al llarg dels microtúbuls dels axons neuronals (195) (Figura 8). En concret, a través d'anàlisis de immunocitoquímica i immunoprecipitació, s'ha observat que les vesícules que transporta KIF1A corresponen a precursors sinàptics que contenen sinaptofisina, sinaptotagmina i Rab-3A (205). Estudis de la motilitat han mostrat que es mou cap a l'extrem positiu dels microtúbuls (moviment anterògrad) i que és una de les proteïnes motores més ràpides (1.2 m/s) (205,206).

Respecte a les funcions de KIF1A en el neurodesenvolupament, s'ha descrit que intervé principalment en la plasticitat sinàptica i en la supervivència, diferenciació i migració neuronal.

La plasticitat sinàptica comporta modificacions en l'estructura i la composició de les espines dendrítiques com a conseqüència de canvis en l'activitat neuronal. En neurones de l'hipocamp, KIF1A transporta de manera directa i específica vesícules que contenen sinaptotagmina-IV a les espines dendrítiques durant el procés de plasticitat sinàptica i impedeix la fusió no regulada d'aquestes vesícules (207). Estudis en ratolins també han demostrat que KIF1A té un paper fonamental en la sinaptogènesi mediada per BDNF (*Brain-derived Neurotrophic factor*) en

l'hipocamp i en la neuroplasticitat induïda per alts nivells d'estímuls sensorials, motors, socials i cognitius (208).

En relació amb la migració neuronal, KIF1A intervé en el desplaçament dels nuclis de les cèl·lules progenitores de la glia radials (RGCs) de l'àrea apical a la basal de la zona ventricular (localitzada en el tub neural) durant la neurogènesi. Les cèl·lules RGCs donen lloc a la majoria de les neurones i cèl·lules de la glia en el còrtex cerebral i presenten un moviment dels seus nuclis de la regió apical a la basal de manera dependent a l'estadi del cicle cel·lular en el que es troben (migració nuclear intercinètica, INM). Recolzant el paper de KIF1A en la migració neuronal es va identificar la variant *missense* R18W en un pacient que presentava paquigiria frontal, reducció de la matèria blanca i un engruïment del cos callós (209). Estudis d'aquesta variant en rates han confirmat que causa defectes en la migració nuclear intercinètica basal i en la migració neuronal, explicant el fenotip present en el pacient (210).

Finalment, KIF1A està implicat en la supervivència i diferenciació neuronal a través del transport de factors neurotròfics i la regulació de l'autofàgia. En les neurones sensorials localitzades en el sistema nerviós perifèric, KIF1A transporta el receptor de neurotrofines nTrkA (*tropomyosin receptor kinase A*), el qual intervé en la via de senyalització NGF/TrkA/PI3K que és essencial per una funció neuronal normal així com en els processos de diferenciació i supervivència neuronal (211). Així mateix, KIF1A determina la localització on s'han de formar els autofagosomes a través del transport de la proteïna d'autofàgia integral de membrana ATG-9. El procés d'autofàgia durant el desenvolupament del cervell és crucial per eliminar i crear noves estructures, i en adults per a la transmissió sinàptica i el reciclatge de vesícules (205).

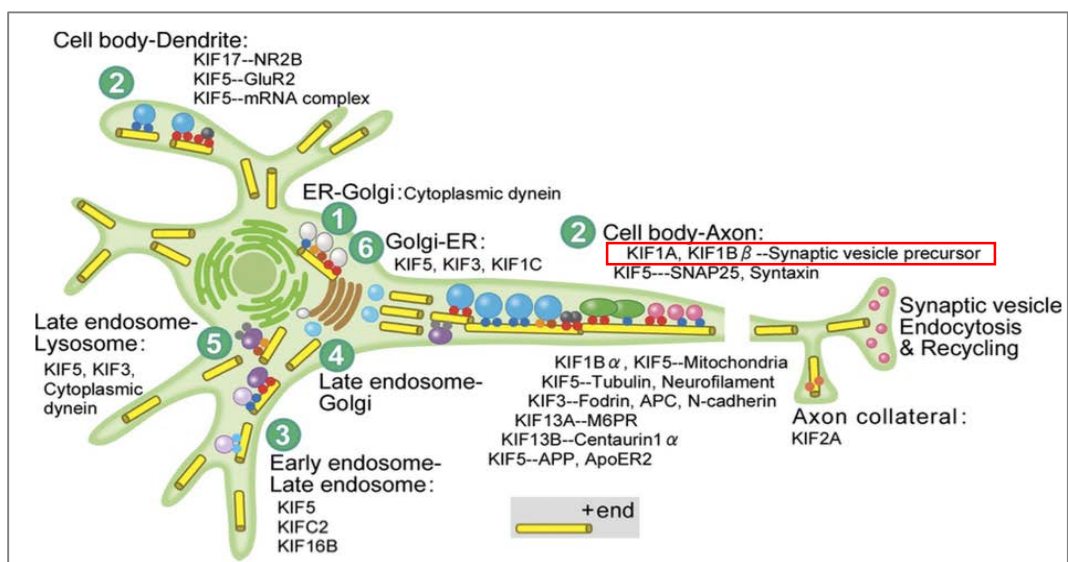


Figura 8 | Transport intracel·lular de vesícules per quinesines en neurones. Imatge extreta de Nobutaka Hirokawa and Yasuko Noda, 2008.

Fenotips associats a KIF1A

Variants patogèniques en el gen *KIF1A* s'han associat amb quatre trastorns diferents: i) neuropatia sensitiva hereditària tipus II (HSAN II, MIM # 614213) que segueix una herència autosòmica recessiva, ii) paraplegia espàstica amb una herència autosòmica recessiva (SPG30, MIM # 610357), iii) paraplegia espàstica amb una herència autosòmica dominant (SPG30, MIM # 610357) i iv) retard mental autosòmic dominant tipus 9 (inclòs en OMIM com NESCAV MIM # 260565). Recentment, també s'ha associat a la síndrome de Rett (152).

Els fenotips de paraplegia espàstica i retard mental autosòmic dominant tipus 9 estan causats principalment per variants *missense* en el domini motor. El fet de què variants localitzades en el domini motor resultin en diferents fenotips s'explicaria perquè afecten de manera diferent a la motilitat de la proteïna i en conseqüència a la seva capacitat de transport. Variants *de novo* en heterozigosi que actuen per un mecanisme dominant negatiu i que produeixen afectacions greus de la motilitat s'associen al fenotip de DI (212,213). Per contra, s'ha suggerit que variants que produeixen afectacions més lleus de la motilitat s'associen al fenotip de paraplegia espàstica sense DI. No obstant, un estudi recent, ha demostrat que variants associades al fenotip de paraplegia espàstica confereixen un guany de funció de KIF1A, que resulta en una sobreactivació del domini motor (214).



Objectius

Objectius

La recerca portada a terme en aquesta tesi doctoral té com a principal objectiu la identificació de nous gens responsables del fenotip SA utilitzant la seqüenciació de l'exoma en trios pacient-progenitors.

S'han establert els següents objectius específics:

1. Seqüenciació de l'exoma en trios pacient-progenitors i anàlisi dels models d'herència mendelians.
2. Anàlisi de la càrrega genètica de variants puntuals *de novo* en els pacients amb un fenotip SA.
3. Identificació de variants en número de còpies i variants en mosaic en els trios en els quals no s'ha identificat la causa del fenotip SA.
 - Variants en mosaic *de novo* en el pacient.
 - Variants en mosaic en els progenitors i present en el pacient en heterozigosi.
4. Anàlisi de l'activació del mecanisme *Nonsense-mediated mRNA Decay* per variants *nonsense* i *frameshift*.
5. Determinar l'impacte funcional de la variant p.Arg169Thr sobre l'activitat de la quinesina KIF1A.
6. Construir una xarxa d'interacció, fent servir eines bioinformàtiques, integrant els gens identificats en la nostra cohort responsables del fenotip *SA-like*, els gens del diagnòstic diferencial de la SA i el gen *UBE3A*.



Material i mètodes

3.1 Subjectes d'estudi

3.1.1 Selecció dels pacients amb fenotip SA-like

Es van seleccionar 17 pacients (9 nens i 8 nenes) de la cohort de pacients de la Corporació Sanitària Parc Taulí (CSPT) que complien les característiques clíniques de la SA descrites per Williams et al., 2005 (36) i que no tenien un diagnòstic molecular.

Un diagnòstic molecular de la SA negatiu consisteix en:

- Un patró de metilació normal del locus *SNURF-SNRPN*.
- No presentar variants puntuals patogèniques/probablement patogèniques ni delecions/duplicacions exòniques en el gen *UBE3A*.

A més a més, es va realitzar un array d'hibridació genòmica comparada (aCGH, ISCA, 60Kb, Agilent Technologies) per comprovar que els pacients no presentaven cap deleción/duplicación patogènica responsable de la seva clínica.

3.1.2 Descripció clínica de la cohort d'estudi

Els pacients inclosos a l'estudi complien les característiques clíniques consistents de la SA descrites per Williams et al., 2005 (36). En la nostra cohort:

- 100% presenta retard del desenvolupament greu i afectació del llenguatge.
- 82,35% (14/17) presenta absència de llenguatge o adquisició de 5-10 paraules, mentre que un 17,64% (3/17) ha adquirit més de 20 paraules.
- 94,1% (16/17) mostra un llenguatge receptiu i comunicació no verbal més preservats que el llenguatge expressiu.
- 58,8% (10/17) presenta atàxia i inestabilitat al caminar. Tot i que l'atàxia és una característica clínica consistent, la revisió portada a terme per Bell L et al., 2019 mostra que depenent de l'etiologia, el percentatge varia del 72,7% al 100% (215).
- 94,1% (16/17) presenta una aparença feliç.
- 76,4% (13/17) presenta estereotípies.

Pel que fa a les característiques clíniques freqüents de la SA, aquestes es troben presents en més del 80% dels individus amb SA. En la nostra cohort:

- 47% (8/17) presenta microcefàlia.
- 76,47% (13/17) presenta epilèpsia.
- 76,9% (10/13) presenta anomalies en l'EEG.

Les característiques clíniques associades es troben entre el 20-80% dels individus amb SA. En la nostra cohort:

- 35,29% (6/17) presenta problemes d'alimentació.
- 52,9% (8/17) presenta hipotonia.
- 17,64% (3/17) presenta la boca gran i l'espai interdental augmentat.
- 41,1% (7/17) mostra problemes de la son.

Tots els pacients presenten trets autistes. A més, tots són casos esporàdics, no hi ha consanguinitat ni troballes rellevants en la història familiar.

3.2 Mètodes

3.2.1 Separació de limfòcits a partir de sang perifèrica

La separació dels limfòcits es porta a terme a partir de 9 mL de sang, que es centrifuguen (3.000 rpm, 5 min) i es recull la capa blanca de cèl·lules, la qual es resuspèn en 15 mL de PBS (*Phosphate Buffered Saline*, Sigma-Aldrich). A continuació, el PBS amb la capa blanca de cèl·lules, es passa a un tub que conté 9 mL de *Lymphoprep* (Palex Medical), que és un medi que permet separar les cèl·lules mononuclears per un gradient de densitat. Es centrifuga a baixes revolucions (2000 rpm) sense fre durant 20 min per a què es pugui establir el gradient i s'agafa la capa de cèl·lules blanques, corresponents als limfòcits i monòcits. Els granulòcits i els eritròcits romanents, es sedimenten al final del tub, ja què tenen una major densitat que les cèl·lules mononuclears. Es fan dos rentats amb 10 mL de PBS i el pellet es resuspèn amb 1 mL de medi de conservació de limfòcits (10 mL DMSO (SULAB), 2 mL Sèrum fetal boví (Biological Industries), 80 mL RPMI (Biological Industries)), es congelen a -80 °C i al dia següent es transfereixen a nitrogen líquid a -196 °C.

3.2.2 Extracció d'ADN

L'extracció d'ADN s'ha realitzat per precipitació salina (reactius Gentra, Qiagen) a partir de sang perifèrica en els 17 pacients i els seus progenitors. Partint d'un volum de 6 mL de sang, primer es porta a terme una lisi dels hematies amb 18 mL *RBC lysis solution*. Es centrifuga (3.500 rpm, 5 min a 4 °C), s'eliminen els hematies lisats i el pellet de cèl·lules blanques es resuspèn amb 6 mL del reactiu *Cell lysis solution* i es deixa tota la nit a T^a amb. Al dia següent es precipiten les proteïnes amb 4 mL del reactiu *Protein Precipitation Solution*. Es centrifuga a 3.500 rpm 5 min a 4 °C i tot seguit, es recull el sobrenedant i es porta a terme la precipitació de l'ADN amb 6 mL d'isopropanol. En últim lloc, es fa un rentat amb etanol al 70%, es deixa assecar l'ADN precipitat

i es resuspèn amb 500 µL de *DNA Hydration Solution*. La quantificació de l'ADN i la determinació de les pureses es realitzen amb l'equip Nanodrop (ND-1000, Thermofisher Scientific).

3.2.3 Extracció d'ARN

L'extracció d'ARN dels limfòcits prèviament congelats en nitrogen líquid a -196 °C, s'ha portat a terme fent servir el kit comercial *Biostic Blood Total Isolation Kit sample* (MO Bio Laboratories) seguint les instruccions del fabricant. Aquesta tecnologia utilitza membranes de silica per a la purificació de l'ARN. En primer lloc, es realitza una lisi cel·lular utilitzant una solució que conté β-mercaptoetanol, que és un agent reductor que desnatura les proteïnes reduint els ponts disulfur, protegint l'ARN de què es pugui degradar per acció de les RNases. A continuació, s'uneixen els àcids nucleics a la columna de silica i es realitza un tractament amb DNaseI (Sigma Aldrich-Roche) per eliminar l'ADN genòmic. Posteriorment, es porten a terme diferents rentats de la columna i en darrer lloc, l'ARN unit a la columna s'elueix en 30 µL d'H₂O *RNase-free*. La quantificació i determinació de les pureses es realitza amb l'equip Nanodrop.

3.2.4 Retrotranscripció

La retrotranscripció és un procés en el que a través d'una retrotranscriptasa s'obté la cadena d'ADN complementaria (ADNc) a partir d'ARN. La retrotranscripció s'ha realitzat amb el kit comercial *PrimeScript RT Reagent Kit* (Takara Bio). Per a la reacció de transcripció s'utilitzen 500 ng d'ARN, 2 µL 5x *Primer Script Buffer*, 0,5 µL *Prime Script RT Enzyme Mix*, 0,5 µL *Oligo dT primers* (50 µM) i 0,5 µL *Random 6mers* (100 µM). Les condicions del termociclador són les següents: 37 °C 15 min i 85 °C 5 s.

3.2.5 Seqüenciació de l'exoma i filtratge de les variants

La seqüenciació de l'exoma complet (WES) s'ha realitzat en trios en 16 dels pacients i els seus progenitors, en un dels casos (pacient 17) el trio està format per la mare i la germana, ja que el pare era èxitus. En el cas de la pacient 5, només s'ha realitzat el WES de la mostra de la pacient i no dels progenitors. Les llibreries s'han generat amb el kit *Sure Select Human All Exon V5+ UTR* (Agilent Technologies) i s'han seqüenciat amb la plataforma Illumina HiSeq2000 que produeix *paired end reads* 2x100 nucleòtids. Aquests dos procediments s'han portat a terme en el Centre Nacional d'Anàlisi Genòmica (CNAG-CRG).

La tecnologia emprada permet la seqüenciació de >99% dels exons del genoma i les regions no traduïdes (UTR) 5' i 3', requerint-se aproximadament 10 µg d'ADN per individu. L'ADN es fragmenta per sonicació i se li uneixen seqüències adaptadores i codis de barres moleculars que permeten la multiplexació de la mostra per crear la llibreria de seqüenciació. A continuació, es porta a terme una captura de les regions d'interès i un enriquiment d'aquestes (Figura 9).

Preparació de la llibreria

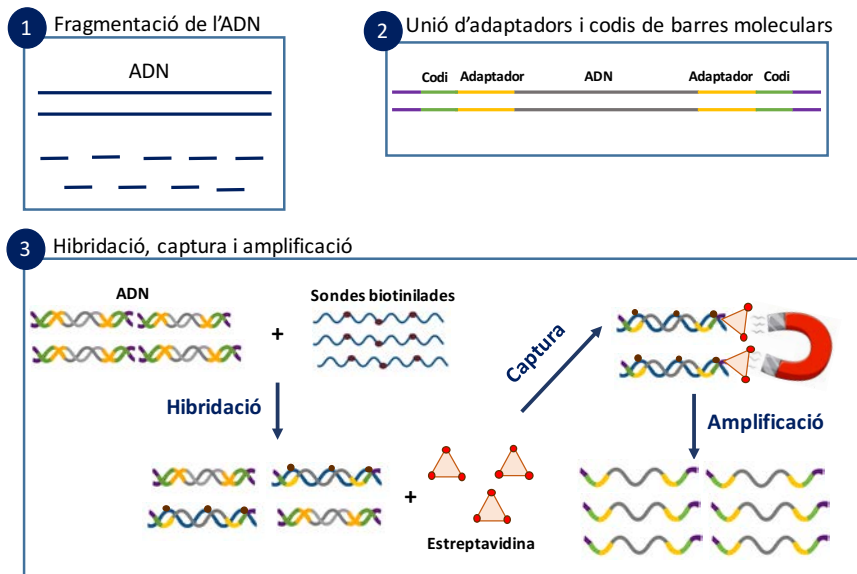


Figura 9 | Procediment utilitzat per a la preparació de la llibreria de seqüenciació.

La seqüenciació s'ha dut a terme en la plataforma Hiseq2000 (Illumina) on els fragments enriquits es sotmeten a una fase d'amplificació clonal per generar els clústers i seqüenciació (Figura 10).

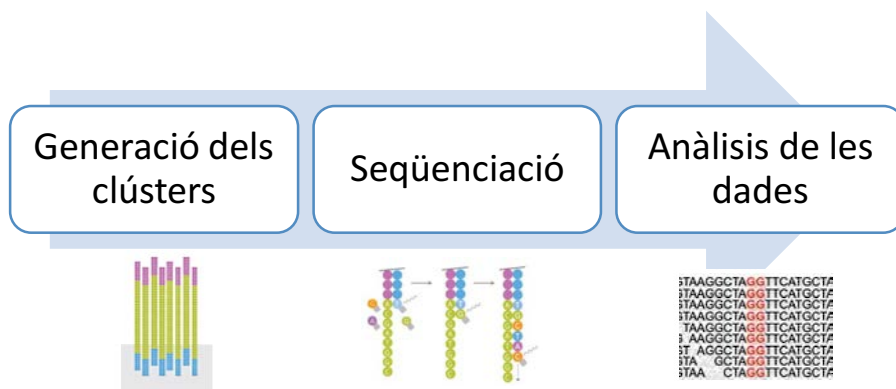


Figura 10 | Esquema que mostra dels tres passos bàsics en la seqüenciació amb la tecnologia Illumina.

Les dades crues de seqüenciació dels 16 trios i la pacient 5 s'han mapat contra el genoma humà de referència (hg19) utilitzant el *software* d'alineament *BWA aligner* (216). Posteriorment, les dades s'han processat amb el *software* GATK (217) per eliminar duplicats de PCR i portar a terme una re-calibració del *score* de qualitat per base. La crida de variants o *variant calling* s'ha realitzat amb l'eina GATK *Haplotype Caller* i les variants resultants s'han anotat amb el *software* ANNOVAR (218) (Figura 11).

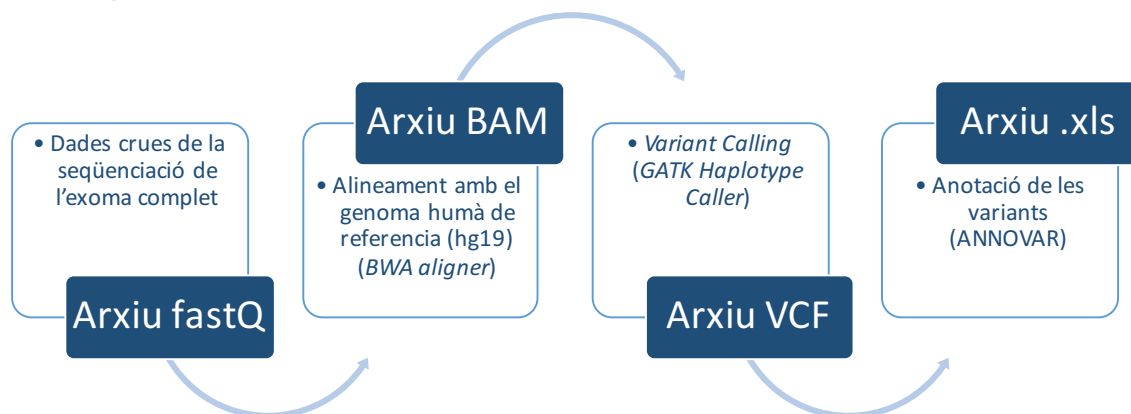


Figura 11 | Algorisme de l'anàlisi bioinformàtic per processar de les dades obtingudes de la seqüenciació de l'exoma.

Model d'herència dominant *de novo*

Inicialment, totes les variants del WES s'han filtrat per un model d'herència dominant *de novo* seguint el protocol que es descriu a continuació (Figura 12). Només s'han considerat les variants amb un genotip associat i amb una qualitat de genotip ≥ 20 en tots els membres del trio. Variants que presenten una profunditat de lectura < 5 en els progenitors i < 10 en els pacients s'han descartat. Només s'han tingut en compte les variants que es troben en heterozigosi en els pacients i els progenitors són homozigots per l'al·lel de referència. Finalment, les variants candidates *de novo* s'han filtrat considerant que l'al·lel alternatiu havia d'estar present en més del 25% dels *reads*.

Les variants *de novo* s'han anotat fent servir el *software* ANNOVAR que permet l'anotació funcional de les variants detectades de la seqüenciació i permet estimar l'impacte de variants *missense* fent servir diferents predictors *in silico* (SIFT, Mutation Assessor, Provean i Mutation Taster, entre d'altres). Les variants que afecten el *splicing* s'han avaluat utilitzant el programa *Human Splicing Finder* (219). Les variants anotades s'han prioritzat segons la freqüència al·lèlica en les bases de dades d'ExAC i gnomAD ($< 1/1000$) i l'impacte sobre la proteïna (*nonsense*, *frameshift*, *splicing* i *missense*). Així mateix, s'ha avaluat si les variants identificades es troben en les bases de dades ClinVar (220) i Decipher (221). La classificació de les variants s'ha portat a terme amb el suport de l'eina Varsome (159) que segueix la guia de l'ACMG (156).

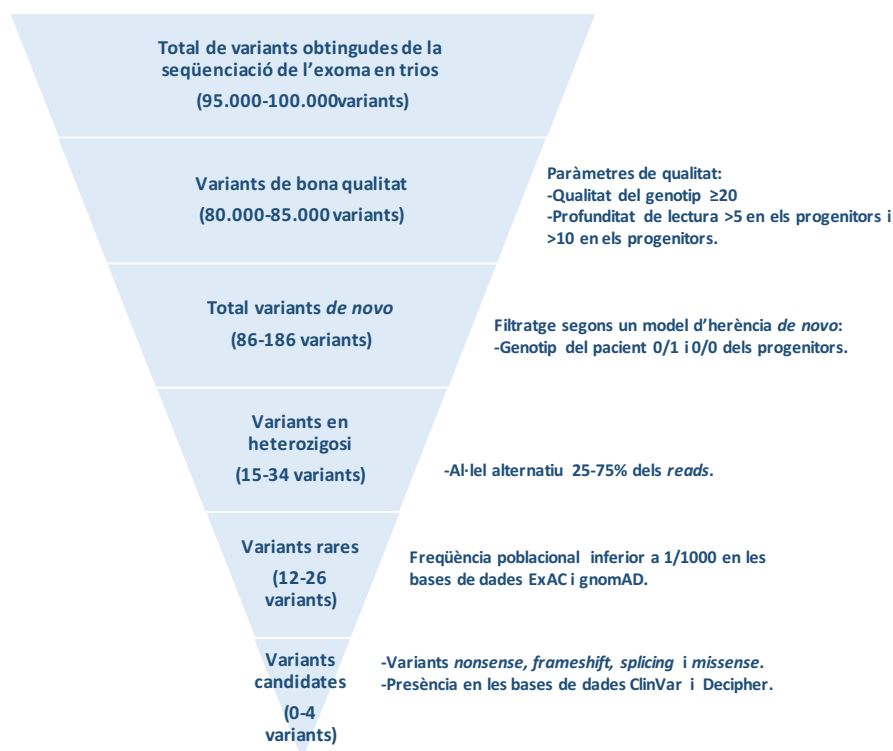


Figura 12 | Protocol portat a terme per a la prioritització de variants segons un model d'herència dominant *de novo*.

Anàlisi d'altres models d'herència mendelians i mosaics

En els pacients en els quals no s'han identificat variants patogèniques/probablement patogèniques en l'anàlisi del model d'herència dominant *de novo*, s'han analitzat altres models d'herència mendelians.

D'acord amb un model d'herència [autosòmic recessiu \(heterozigot compost\)](#):

- Només s'han considerat les posicions amb un mínim de 10 *reads* en els tres membres del trio.
- Les variants s'han filtrat per una freqüència al·lèlica $<1/200$ en les bases de dades ExAC i gnomAD.
- S'han prioritzat les variants amb un impacte sobre la proteïna (*nonsense*, *frameshift*, *splicing* i *missense*).
- Els progenitors havien de ser cadascun portador d'una variant (genotip 0/1) i homozigot per l'al·lel de referència de la segona variant (genotip 0/0) i a l'inrevés.

Seguint un model d'herència [lligat al cromosoma X](#):

- Només s'han considerat les posicions amb un mínim de 10 *reads* en els tres membres del trio.
- Les variants s'han filtrat per una freqüència al·lèlica $<1/1000$ en les bases de dades d'ExAC i gnomAD.
- S'han prioritzat les variants amb un impacte sobre la proteïna (*nonsense*, *frameshift*, *splicing* i *missense*).

En els pacients en els quals no s'han identificat variants patogèniques/probablement patogèniques en els models d'herència mendelians autosòmic recessiu i lligat al cromosoma X, també s'han analitzat les variants en mosaic. D'una banda, s'han estudiat les variants en [mosaic de novo](#) filtrant per:

- Un genotip 0/1 (heterozigot) en el pacient i 0/0 en els progenitors.
- Un mínim de 10 *reads* de cobertura en els tres membres del trio.
- Una freqüència al·lèlica entre 10-30%.
- Una freqüència inferior a <1/1000 en la bases de dades d'ExAC i gnomAD.

D'altra banda, s'han analitzat variants en [mosaic en els progenitors i que els pacients fossin portadors heterozigots de la variant](#). En aquest últim cas, s'ha filtrat per:

- Variants amb una cobertura mínima de 10 *reads* en els tres membres del trio.
- S'han considerat aquelles variants en les que el genotip del pacient és 0/1 i la proporció de l'al·lel alternatiu es troba entre 30-70%.
- Només un dels progenitors havia de ser portador de la variant i la proporció de l'al·lel alternatiu havia d'estar entre 5-30%.
- Les variants havien de tenir una freqüència <1/1000 en les bases de dades d'ExAC i gnomAD.

En els diferents anàlisis (autosòmic recessiu, lligat al cromosoma X i mosaic), s'ha comprovat si les variants identificades estan presents en les bases de dades ClinVar i Decipher. A més, s'ha realitzat una cerca de la informació funcional i del/s fenotip/s associat/s als gens identificats en les bases de dades d'OMIM (222), DD2GP (*Development Disorder Genotype – Phenotype, Decipher*) (223), SysID (224) i Pubmed (225).

Les variants candidates s'han validat per seqüenciació Sanger (amb el procediment que es descriu a continuació en l'apartat 3.2.5.1) en els progenitors i en els pacients per confirmar la presència de la variant i el patró d'herència.

3.2.5.1 Seqüenciació Sanger

La seqüenciació Sanger consisteix en la incorporació de dideoxinucleòtids (ddNTPs) marcats per una ADN polimerasa durant la replicació *in vitro* de l'ADN, fent que acabi la replicació. Els diferents fragments replicats es separen en una electroforesi capil·lar i els ddNTPs marcats s'exciten amb un làser, obtenint-se un electroferograma amb la seqüència de la regió d'estudi, on cada nucleòtid (A, T, G i C) té assignat un color. En la Taula 5 es troben els *primers* dissenyats per amplificar les regions on es localitzen les variants d'interès.

Taula 5 | Primers dissenyats per a la validació de variants per seqüenciació Sanger. En color blau fosc està marcada la seqüència dels primers M13 *forward* i *reverse*.

<p>Exó 10 gen <i>LAS1L</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTGGACTTGAATGACGTCCTG 3' T=58 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCGTTTGCCTGGTCTTCTTAG 3' T=60 °C</p>	<p>Exó 4 gen <i>KCNQ3</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTGAATAATCCCACCTTACCATG 3' T=56 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCCTTTCGTTATCATCAAGG 3' T=56 °C</p>
<p>Exó 11 gen <i>TBL1XR1</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTTAGTCAACAATGGTTCAAAGC 3' T=58 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCATTGTATAGACTGGGTAGC 3' T=58 °C</p>	<p>Exó 6 gen <i>KIF1A</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTGTCTTATGTTGACCCCTAC 3' T=60 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCACAGGCTAAACGACCTC 3' T=60 °C</p>
<p>Exó 50 gen <i>SPTAN1</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTGAACCTACCAAGTGCTCTGAG 3' T=60 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCCTTATCCAGGCTGCCTC 3' T=62 °C</p>	<p>Exó 3 gen <i>VAMP2</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTCTAGGCATGTTCTGTGTGC 3' T=60 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCAGGATCAGTAAGACAACTATG 3' T=60 °C</p>
<p>Intró 8 gen <i>SMARCE1</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTCTTAGCAGCTTTGAGATAAG 3' T=56 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCGTTAAACAAACGACTGGCTC 3' T=58 °C</p>	<p>Exó 3 gen <i>COX7B</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTAGTCTCCGAGAGAGTGAC 3' T=58 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCCTAATGGGTACACAGTGC 3' T=56 °C</p>
<p>Exó 10 gen <i>CHMP7</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTCATTAATGCATTTCTCCCTTG 3' T=58 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCATCAACTCTGTGACGCCAG 3' T=58 °C</p>	<p>Exó 11 gen <i>SYNGAP1</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTCTCAGCATGTTCCCTGGAAG 3' T=62 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCGACTCAGGAGACCCCTCTTG 3' T=62 °C</p>
<p>Exó 11 gen <i>SATB2</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTCCAGTGTAGATAGATAACAC 3' T=56 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCGAAGAAGCTTGTATGGTGTG 3' T=56 °C</p>	<p>Exó 5 gen <i>HSF2</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTGTGTAGTTGTTGTTCTGG 3' T=56 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCGAAGAAGCTTGTATGGTGTG 3' T=56 °C</p>
<p>Exó 12 gen <i>ASXL3</i></p> <p>For 5' GTTGTA AAAACGACGGCCAGTGGAGAGAATGCAGCTAGTC 3' T=58 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCCTTGTGCTAGGAGATCTAC 3' T=58 °C</p>	

Primerament, es realitza l'amplificació de la regió d'estudi, la mix consisteix en 12,5 µL *PCR Master Mix 2x* (Promega), 1,5 µL *primer forward*, 1,5 µL *primer reverse*, 8,5 µL H₂O estèril i 1 µL ADN. Les condicions del termociclador són: 95 °C 5 min, 35 cicles de 95 °C 30 s, T^a *melting* 30 s i 72 °C 45 s, i l'últim pas de 72 °C 20 min. Tot seguit, els productes de PCR es purifiquen utilitzant ExoSAP-IT (Affymetrix), que són enzims que digereixen l'excés d'oligonucleòtids i defosforilen nucleòtids. Seguint les instruccions del fabricant, s'utilitzen 8 µL del producte de PCR i 2 µL de ExoSAP-IT, les condicions del termociclador són: 37 °C 15 min i 80 °C 15 min.

La reacció de seqüenciació es realitza amb el kit *Big Dye Terminator* (Applied Biosystems). La mix de seqüenciació consisteix en 0,5 µL *BigDye Terminator*, 2 µL *Tampó 2x*, 3,2 µL *primer M13 forward* (1 pmol/µL), 3,5 µL H₂O estèril (Braun) i 3 µL producte de PCR purificat. Les condicions del termociclador són: 25 cicles de 96 °C 10 s, 50 °C 5 s i 60 °C 4 min. Tot seguit, es realitza la precipitació dels 12 µL obtinguts de la reacció de seqüenciació amb etanol al 100%, es porta a terme un rentat amb etanol al 70% i es deixa assecar durant 15 min a T^a amb.

En darrer lloc, es resuspenen les seqüències amb 10 µL de HiDi formamida (Applied Biosystems) i es carreguen en una placa de l'analitzador genètic ABI3130 (Applied Biosystems) que realitza una electroforesi capil·lar. Els resultats s'han analitzat amb el programa SeqScanner (Applied Biosystems).

3.2.6 Anàlisi de les CNVs exòniques a partir de les dades de l'exoma

Les variants en el número de còpies (CNVs) s'han detectat utilitzant el programa XHMM, que porta a terme un anàlisi de la cobertura de les diferents regions estudiades en l'exoma, per així poder realitzar una normalització i a través de l'aplicació del *Hidden Markov Model* identificar les CNVs (226). Les CNVs identificades s'han filtrat per un model d'herència dominant *de novo*. S'han descartat aquelles CNVs que solapen un 50% amb una altra CNV que es troba present en més del 10% de totes les mostres de la cohort d'estudi. Finalment, les CNVs *de novo* han estat filtrades considerant només aquelles que compleixen els següents criteris: i) el valor de *Phred score* que indica la presència de la CNV en el pacient havia de ser ≥ 10 , ii) el valor de *Phred score* que indica l'absència de la CNV en els progenitors havia de ser ≥ 10 i iii) s'ha filtrat per aquelles CNVs que estan situades en gens implicats en el neurodesenvolupament.

3.2.6.1 Multiplex Ligation-dependent Probe Amplification (MLPA) customitzada

La MLPA és una tècnica que permet avaluar de manera semiquantitativa canvis en el número de còpies (CNVs) de diferents regions cromosòmiques al mateix temps. En el cas de la MLPA customitzada a diferència dels kits comercials, es dissenyen sondes de manera personalitzada seguint les guies del fabricant (MRC-Holland), amb el fi d'estudiar regions específiques que no es troben cobertes en kits comercials. En el nostre cas, l'objectiu ha estat validar CNVs predites *de novo* pel programa XHMM a partir de les dades de la seqüenciació de l'exoma.

Les sondes de MLPA consisteixen en dos oligonucleòtids adjacents, l'esquerra (*Left Probe Oligonucleotid*, LPO) i el dret (*Right Probe Oligonucleotid*, RPO). Cadascun d'ells consisteix en una seqüència que hibrida amb la regió d'interès més la seqüència universal *forward* i *reverse* dels primers de PCR, respectivament (Figura 13).



Figura 13 | Disseny d'una sonda de MLPA, on poden observar la nomenclatura dels diferents components de la sonda (MRC-Holland).

3. MATERIAL I MÈTODES

Per dissenyar les sondes cal tenir en compte: i) les sondes LPO i RPO han de ser adjacents i la sonda RPO ha de portar en l'extrem 5' un grup fosfat per a què es pugui portar a terme la reacció de lligació, ii) no han de solapar amb cap SNP conegut i la regió que hibrida amb l'ADN ha de tenir una llargada mínima de 21 nucleòtids, iii) la T^a *melting* dels oligonucleòtids esquerra i dreta per separat ha de ser >71 °C i el contingut en GC ha de ser d'aproximament 50%, iv) preferiblement en el punt on han de lligar l'oligonucleòtid esquerra i el dret ha d'haver-hi una A o una T, v) entre dues sondes diferents ha d'haver-hi una diferència mínima de quatre nucleòtids i vi) la sonda dissenyada no ha d'hibridar en altres regions cromosòmiques.

En el nostre cas, s'han dissenyat tres sondes de MLPA que es troben situades dins de les CNVs candidates (Taula 6).

Taula 6 | Sondes de MLPA dissenyades en les regions d'interès.

Gen	Banda cromosòmica	Inici	Final	Seqüència LPO 5'-3' / Seqüència RPO 5P'-3'
MDGA2	14q21.3	47530697	47530759	GGGTTCCCTAAGGGTTGGAGATCACACCAGATCCTTATCACAAGATGA / CAACATCCAGATTGGCCGTGAGGTGAAAATATCTCTAGATTGGATCTTGCTGGCAC
ATAD3A	1p36.33	1454139	1454205	GGGTTCCCTAAGGGTTGGACCTGCACATTCGGGCACCGTCACCCCGCAA / ACGGGCACCGTACACCCCGCAAACGGGCACACTCTAGATTGGATCTTGCTGGCAC
TM4SF20	2q36.6	228243853	228243904	GGGTTCCCTAAGGGTTGGAGGGTTTTGAGAAAATTGGTCTTCTCA / ACTAAGCTGACAATTAGAGGTATCTCTAGATTGGATCTTGCTGGCAC

La tècnica de MLPA s'ha portat a terme seguint les instruccions del fabricant: desnaturalització de l'ADN 10 min a 98 °C, hibridació de les sondes 16 h a 60 °C, lligació 15 min a 54 °C i finalment PCR amb *primers* universals. Un cop finalitzada la PCR, es realitza una electroforesi capil·lar en l'analitzador genètic ABI3130. Per passar les mostres per l'analitzador genètic es porta terme una dilució 1/50 de les mostres amb H₂O i es prepara la següent mix: 9 µL HiDi formamida, 0,2 µL *Genescan 500 Liz dye Size Standar* (Applied Biosystems) i 1 µL producte de PCR.

L'anàlisi dels fragments es realitza amb el programa *Genemapper* (Applied Biosystems). El valor de l'alçada dels pics s'utilitza per realitzar la normalització intraindividual i interindividual, donant com a resultat el quocient de dosi (QD). El valor del QD indica si hi ha canvis a nivell del número de còpies: i) si és inferior a 0,7 es tracta d'una deleció, ii) si es troba entre 0,7 i 1,29 és normal i iii) si és superior a 1,29 es tracta d'una duplicació.

3.2.7 Quantificació de l'al·lel amb la variant amb el programa *Minor Variant Finder*

El programa *Minor Variant Finder* (MVF, Applied Biosystems) utilitza un algorisme que elimina el soroll de fons fent servir les dades de seqüenciació Sanger d'una mostra control i permet

identificar variants que es troben en un percentatge baix (fins al 5%). Les variants en un percentatge baix s'han de detectar en les dues cadenes (*forward* i *reverse*), permetent confirmar que les variants trobades són reals. En el nostre cas, aquest programa s'ha utilitzat amb l'objectiu de quantificar el percentatge de l'al·lel amb la variant a partir de la seqüenciació Sanger en ADN genòmic i ADNc.

S'han analitzat les variants *nonsense* i *frameshift* amb l'objectiu de comprovar si activaven el mecanisme de *Nonsense-Mediated mRNA Decay* (NMD). Les variants estudiades han estat: c.1861C>T en el gen *SYNGAP1* (pacient 3), c.1826delA en el gen *SATB2* (pacient 6), c.3106C>T en el gen *ASXL3* (pacient 10), c.456_459delTGAG en el gen *HSF2* (pacient 13) i c.1165C>T en el gen *CHMP7* (pacient 17). En la Taula 7 es troben les seqüències dels primers dissenyats per realitzar la quantificació de l'al·lel en ADNc, els primers utilitzats per a la quantificació en ADN genòmic són els mateixos que s'han fet servir per a la validació de les variants i es troben en la Taula 5. Tant les mostres problema com les control s'han seqüenciat amb els primers M13 *forward* i *reverse*.

Taula 7 | Primers dissenyats per quantificar l'al·lel amb la variant en ADNc dels gens *CHMP7*, *SYNGAP1*, *SATB2*, *HSF2* i *ASXL3*. En color blau fosc està marcada la seqüència dels primers M13 *forward* i *reverse*.

<p>Exons 11-12 del gen <i>SYNGAP1</i> en ADNc</p> <p>For 5' GTTGTAAAACACGACGGCCAGTGACAGGCCTTATCAGCGCCTC 3' T=64 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCGATGCAGTGTGGAGAGCAC3' T=64 °C</p>	<p>Exons 7-10 del gen <i>CHMP7</i> en ADNc</p> <p>For 5' GTTGTAAAACACGACGGCCAGTGCACAGTGGAGAAGGCAG 3' T=60 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCAAACCTCCCTCTGATAAAGG 3' T=60 °C</p>
<p>Exons 10-11 del gen <i>SATB2</i> en ADNc</p> <p>For 5' GTTGTAAAACACGACGGCCAGTCATGCAACACGTGGTCCAG 3' T=60 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCAGAGTGTGGATGGCTTCTCG 3' T=62 °C</p>	<p>Exons 3-6 del gen <i>HSF2</i> en ADNc</p> <p>For 5' GTTGTAAAACACGACGGCCAGTGGACAGGATGACTTGTGG 3' T=58 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCATGTGCTGAAACAGGTTCC 3' T=58 °C</p>
<p>Exons 11-12 del gen <i>ASXL3</i> en ADNc</p> <p>For 1 5' GTTGTAAAACACGACGGCCAGTGAAGATGATCAGTCAACCCG 3' T=60 °C</p> <p>For 2 5' GTTGTAAAACACGACGGCCAGTCTCTCAAGATTCACTTTCC 3' T=62 °C</p> <p>Rev 5' CACAGGAAACAGCTATGACCCAGAGACAATGCTCGCTG 3' T=60 °C</p>	

3.2.8 Real Time-Polymerase Chain Reaction (RT-PCR) amb sondes Taqman

La RT-PCR també coneguda com PCR quantitativa (qPCR) és una tècnica que permet mesurar l'amplificació d'una o més seqüències d'ADN o ADNc a mesura que aquesta es va generant. És una tècnica altament precisa i permet estudiar l'expressió gènica, analitzant si hi ha un augment o una disminució en l'expressió d'un gen o d'un grup de gens mesurant l'abundància del transcrit. Per analitzar l'expressió gènica en aquest estudi s'ha fet servir la tecnologia TaqMan (Taula 8), que consisteix en un parell de primers i una sonda Taqman marcada amb FAM o VIC en l'extrem 5' i en l'extrem 3' conté el MGB (*Minor Groove Binder*) i el supressor no fluorescent NFQ. Durant

la replicació, l'ADN polimerasa separa el supressor NFQ del fluoròfor, permetent així l'emissió de fluorescència.

Aquesta tècnica s'ha portat a terme per analitzar si hi havia canvis en l'expressió dels gens *SYNGAP1*, *CHMP7*, *SATB2* i *HSF2* que presenten variants *nonsense* i *frameshift* que poden activar el mecanisme de control post-transcripcional NMD, induint la degradació de transcrits amb codons stop prematurs.

Taula 8 | Sondes TaqMan utilitzades per analitzar l'expressió gènica.

Sondes problema	Sondes control (<i>housekeeping genes</i>)
Hs01006395_m1 <i>SYNGAP1</i> (exons 3-4)	Hs02758991_g1 <i>GAPDH</i> (exons 7-8)
Hs00378422_g1 <i>CHMP7</i> (exons 6-7)	
Hs00988309_g1 <i>HSF2</i> (exons 2-3)	
Hs01546832_m1 <i>SATB2</i> (exons 5-6)	

Amplificació per RT-PCR

La RT-PCR s'ha realitzat en l'aparell QuantStudio 7 Flex System (Applied Biosystems). Les mostres problema i les control s'han posat per triplicat. La mix de la RT-PCR consisteix en: 10 µL Mix *SensiFAST™ Probe Lo-ROX* (Bioline), 1 µL sonda *Taqman* (Thermofisher Scientific), 7 µL H₂O *RNase free* (Promega) i 2 µL ADNc. El programa d'amplificació consta d'una desnaturalització 2 min a 95 °C i 40 cicles de 10 s a 95 °C i 30 s a 60 °C.

Anàlisi dels resultats

Els resultats s'han analitzat seguint el mètode $2^{-\Delta\Delta CT}$ (227) amb el programa QuantStudio RT-PCR (Applied Biosystems). Per analitzar si hi ha diferències significatives en els gens analitzats s'ha utilitzat el test estadístic T-Student (IC, 95%).

3.2.9 Clonatge del domini motor de KIF1A *wild type* (WT) i mutat (R169T) en el plàsmid d'expressió pQE-80L

Els constructes del domini motor de KIF1A (aminoàcids 3-362, NP_001230937.1) *wild type* (WT) i amb la variant c.506G>C, p.Arg169Thr (R169T) clonats en el plàsmid pMA-RQ (resistent a ampicil·lina) s'han obtingut de la casa comercial *GeneArt Gene Synthesis* (Thermofisher Scientific). Prèviament a fer el clonatge, s'ha realitzat una amplificació dels constructes WT i R169T pMA-RQ. Es transformen 25 µL de cèl·lules competents DH5-α amb 1 µL de WT/R169T pMA-RQ (que es

troben a una concentració de 100 ng/ μ L). S'incuba 30 min en gel, tot seguit, es realitza un xoc tèrmic a 42 °C 45 s i es posa 1-2 min en gel. A continuació, s'afegeixen 500 μ L de medi SOC (ThermoFisher Scientific) i s'incuba durant 1 h a 37 °C a 220 rpm. Posteriorment, es sembren 5 μ L en una placa de LB amb l'antibiòtic ampicil·lina (100 μ g/mL) i es deixen en l'estufa a 37 °C durant tota la nit. Al dia següent, es realitza un pre-cultiu, picant 1 de les colònies WT/R169T i fent-les créixer amb 3 mL de LB i 3 μ L d'ampicil·lina (100 μ g/mL) durant 6 h. S'inoculen 200 μ L del pre-cultiu en 200 mL de LB i 200 μ L ampicil·lina (100 μ g/mL) que s'incubarà tota la nit a 37 °C i 220 rpm. Al dia següent, es realitza l'extracció de l'ADN plasmídic fent servir el kit comercial *QIAGEN Plasmid Midi*, seguint les instruccions del fabricant.

3.2.9.1 Subclonatge en el vector d'expressió pQE-80L

El subclonatge en el vector pQE-80L (cedit pel grup de *Molecular Genetics, University of Konstanz* del Dr. Thomas U. Mayer) s'ha realitzat seguint l'esquema de la Figura 14. Es digereixen 4 μ g dels constructes WT/R169T pMA-RQ i 4 μ g del vector pQE-80L amb 1 μ L dels enzims de restricció FseI (2.000 U/mL) i AscI (10.000 U/mL), 4 μ L CutSmart (New England Biolabs) i H₂O fins a un volum final de 40 μ L. S'incuba 1 h en el bany a 37 °C.

Per evitar que el plàsmid pQE-80L torni a relligar sense l'insert es realitza una desfosforilació utilitzant una fosfatasa alcalina amb els reactius del kit comercial *Rapid DNA Dephos & Ligation Kit* (Sigma-Aldrich). La reacció de desfosforilació consisteix en 3 μ L de *rAPid Alkaline Phosphatase* i 4,5 μ L de *rAPid Alkaline Phosphatase Buffer* 10x i s'incuba en el bany a 37 °C durant 30 min.

El producte de la reacció de restricció i de la reacció de defosforilació es carrega en un gel d'agarosa a l'1%. Es tallen les bandes corresponents al pes molecular del domini motor que és 1114 pb (WT/R169T) i 4751 pb del plàsmid pQE-80L. La purificació de les bandes del gel s'ha portat a terme amb el kit comercial *Zymoclean™ Gel DNA Recovery* (Zymogen Research) seguint les instruccions del fabricant.

La reacció de lligació s'ha realitzat amb el kit comercial *Rapid DNA Dephos & Ligation Kit* (Sigma-Aldrich). Es necessiten 100 ng del plàsmid pQE-80L per portar a terme la reacció de lligació, la proporció entre plàsmid i insert és 1:2, per tant, es posa el doble de quantitat d'insert, 4 μ L de *DNA dilution Buffer* i H₂O fins a un volum final de 9 μ L. A continuació, s'afegeixen 9 μ L de *T4 DNA dilution buffer* 2x i 1 μ L de *T4 DNA ligase*. S'incuba 15 min a T³ amb.

Un cop finalitzada la reacció de lligació es transformen 5 μ L en cèl·lules competents DH5- α amb el protocol esmentat anteriorment. Al dia següent es piquen 4 colònies de la placa amb el plàsmid pQE-80L amb el domini motor WT i 4 colònies de la placa que conté el plàsmid pQE-80L amb el

domini motor R169T. Aquestes colònies es fan créixer en 3 mL de LB amb 3 µL d'ampicil·lina (100 µg/mL) a 37 °C i en agitació (220 rpm) durant tota la nit i posteriorment, es purifica l'ADN amb el kit comercial *QIAprep® Spin Miniprep* (Qiagen) i es quantifica la concentració amb l'equip Nanodrop.

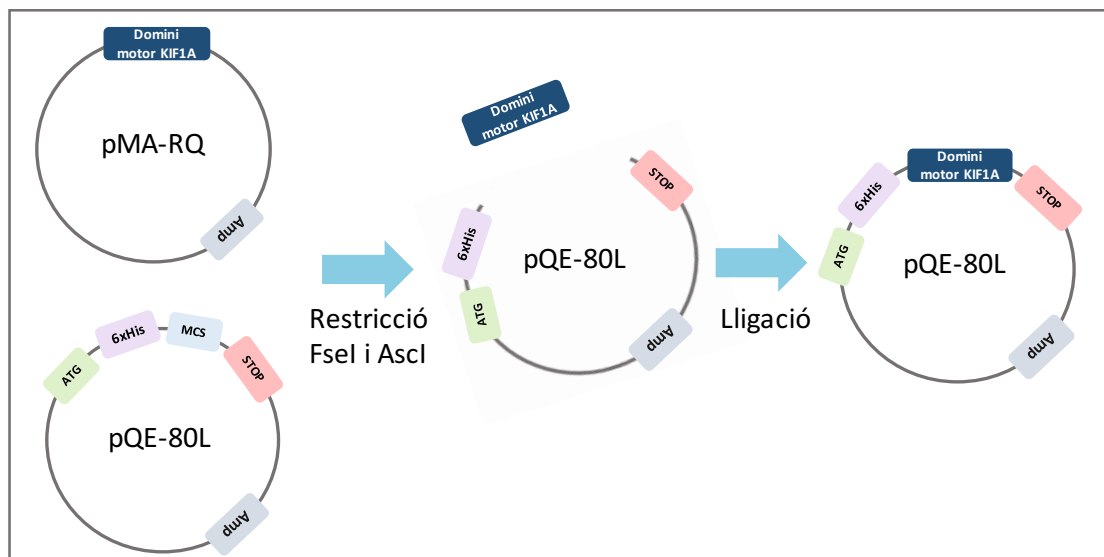


Figura 14| Esquema del clonatge realitzat per introduir el domini motor de KIF1A WT i R169T en el plàsmid d'expressió pQE-80L. MCS, *Multicloning site*; 6xHis, cua de 6 aminoàcids d'histidina; Amp, Ampicil·lina.

Finalment, es comprova que l'insert s'ha clonat correctament en el plàsmid pQE-80L realitzant una reacció de restricció amb els enzims FseI i AscI i per seqüenciació Sanger utilitzant els primers de la Taula 9.

Taula 9| Primers dissenyats per seqüenciar l'insert clonat en el vector pQE-80L.

pQE-80L For		pQE-80L Rev	
5' GTGAGCGGATAACAATTTC 3'	54 °C	5' GAGTTCTGAGGTCATTACTG 3'	56 °C
KIF1A C-term Domini motor			
5' GCAATGGTTGCAGCACTGTC 3'	62 °C		

Una de les colònies amb el domini motor de KIF1A WT/R169T correctament clonat en el vector d'expressió pQE-80L, s'ha transformat en cèl·lules DH5-α i s'ha realitzat una *MidiPrep*, per obtenir una major quantitat del vector clonat.

3.2.10 Producció de proteïna recombinant

La producció de proteïnes recombinants en cèl·lules competents d'*Escherichia Coli* (E. Coli) es basa en què aquestes cèl·lules tenen inserit en el cromosoma bacterià la seqüència codificant de la T7 ARN polimerasa sota el control de l'operó lacUV5. El repressor lac controla l'accés a la

seqüència codificant de la T7 ARN polimerasa unint-se a l'operó lacUV5. La producció de la proteïna recombinant es dona quan s'indueix l'expressió de la T7 ARN polimerasa a l'afegir en el medi *isopropyl-β-D-1-thiogalactopyranoside* (IPTG), que és un anàleg de l'allolactosa. La T7 ARN polimerasa produïda transcriurà la seqüència codificant de la proteïna d'interès que està sota el control del promotor T7 (228,229) (Figura 15).

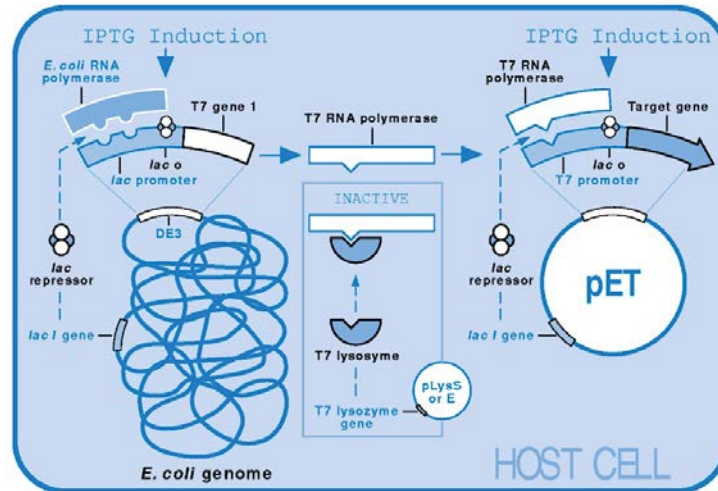


Figura 15| Sistema de producció de proteïnes recombinants en E.Coli. Imatge extreta de Mierendorf, R.C. et al., 1994. The pET System : Your Choice for Expression.

El primer pas per a la producció de proteïna és la transformació dels vectors pQE-80L amb el domini motor de KIF1A WT/R169T en les cèl·lules competents E. Coli BL21-Codon Plus-RIL. Aquestes cèl·lules competents contenen còpies extres de gens que codifiquen per ARNs de transferència (tARNs), que permeten obtenir alts nivells d'expressió de proteïnes heteròlogues. Les còpies extres dels tARNs es troben en plàsmids pACYC que són resistents a cloramfenicol. Es transformen 10 µL de cèl·lules BL21-Codon Plus-RIL amb 100 ng dels vectors d'expressió pQE-80L WT i R169T seguint el protocol esmentat en l'apartat 3.2.9. Al dia següent es piquen 10 colònies i s'inoculen en un mateix tub que conté 10 mL de LB, 10 µL d'ampicil·lina (100 µg/mL) i 18,75 µL cloramfenicol (50 µg/mL) i es fan créixer durant tota la nit a 37 °C i 220 rpm. Al dia següent s'extreu 200 µL del cultiu i es mesura la densitat òptica a 600 nm (DO600). S'inocula la quantitat de cèl·lules corresponents a DO600=0,05, que són 10 mL del cultiu de tota la nit en 250 mL de LB fresc amb 250 µL d'ampicil·lina (100 µg/mL) i 625 µL de cloramfenicol (50 µg/mL). A les 2-4 h es mesura la DO600, que ha d'estar entre 0,4-0,6. S'extreu la mostra corresponent al pellet sense induir que consisteix en una alíquota equivalent a volum (mL)=0,8/DO600 que es centrifuga a màxima velocitat durant 1 min. Tot seguit, s'afegeix IPTG (Sigma-Aldrich) als 250 mL de cultiu a una concentració final de 1 mM (2,5 mL) i es deixa créixer a 18 °C durant tota la nit a 220 rpm. Al dia següent es mesura la DO600 i es fa una alíquota

equivalent a volum (mL)=0,8/DO600. Es centrifuga 1 min a màxima velocitat, es descarta el sobrenedant i aquest serà el pellet induït.

3.2.10.1 Anàlisi de l'expressió de les proteïnes recombinants per *Western Blot*.

La tècnica *western blot* es basa en la detecció de proteïnes d'interès a través d'anticossos en lisats cel·lulars. Primer es separen les proteïnes per pes molecular en un gel de poliacrilamida. Tot seguit, les proteïnes es transfereixen a una membrana on poden ser identificades per immunodetecció. Aquesta tècnica permet valorar l'expressió de proteïnes d'interès.

Per analitzar l'expressió de la proteïna recombinant corresponent al domini motor KIF1A WT/R169T, les mostres del pellet no induït i induït s'han resuspès en 50 µL de la solució de lisi del kit de *Qiagen Ni-NTA Fast Start*. S'incuben 30 min en gel i es centrifuga 30 min a 14.000 g a 4 °C. Es recull el sobrenedant, que és la fracció que conté la proteïna soluble, i s'afegeixen 50 µL de *Laemmli Buffer 2x*. Es prepara una alíquota de 10 µL de cada mostra, es desnaturalitzen a 95 °C 5 min i es centrifuguen a màxima velocitat 5 min. Es carreguen 10 µL de totes les mostres (WT/R169T) no induïdes i induïdes en un gel pre-Cast Novex Tris-Glicina (Invitrogen) de poliacrilamida 4-12% i es deixa córrer a 150 mV en *MES-running buffer* (Thermofisher Scientific) fins que les proteïnes de la mida d'interès queden ben separades.

Un cop finalitzada l'electroforesi, es realitza la transferència de les proteïnes del gel a una membrana de PVDF fent servir el mètode de transferència semiseca. El gel i la membrana es troben formant un *sandwich* entre papers de filtre Whatman (Fisher Scientific). Es deixa transferir a 110 mV durant 90 min.

Posteriorment, es fa una tinció amb *Naphtol Blue* (10% metanol, 2% àcid acètic, 0,1% *Naphtol Blue Black* (Sigma Aldrich)) durant 2 min i s'aplica la solució decolorant (50% metanol i 7% àcid acètic). Aquesta tinció permet observar les proteïnes transferides a la membrana i identificar el domini motor WT/R169T segons el pes molecular.

També s'ha comprovat per immunodetecció que hi ha expressió de la proteïna recombinant. Es bloqueja la membrana amb BSA al 3% en TBS (solució salina) durant 1 h a T^a amb. Seguidament es realitzen dos rentats de la membrana durant 10 min amb TBS-Tween i un rentat de 10 min amb TBS. S'incuba la membrana amb l'anticòs penta-His, que reconeix la cua d'histidines present en la proteïna recombinant, durant 1 h a T^a amb. S'ha realitzat una dilució de l'anticòs 1/1000 en BSA al 3% en TBS. Es realitzen dos rentats de 10 min amb TBS-T i un rentat de 10 min amb TBS i s'incuba la membrana amb l'anticòs secundari, que en aquest cas és un anticòs *anti-mouse IgG-peroxidase* del qual es fa una dilució 1/10.000 i es deixa incubar a T^a amb durant 1 h. En darrer lloc, es fan tres rentats de 10 min amb TBS-T, per així poder eliminar l'excés d'anticòs de la

membrana. El revelat de la membrana s'ha portat a terme amb el producte ECL (*Enhanced Chemiluminescence*, GE Healthcare Life Sciences™ Amersham™). La detecció de la proteïna es basa en una reacció quimioluminiscent, ja que l'anticòs secundari té unit una peroxidasa (*Horseradish Peroxidase*) i a l'afegir un substrat quimioluminiscent, en aquest cas, el luminol combinat amb un fort agent oxidant, com el peròxid d'hidrogen, emet un senyal a una longitud d'ona de 450 nm. Les emissions de llum són capturades per autoradiografia.

3.2.10.2 Purificació de la proteïna recombinant.

La purificació en estat natiu de la proteïna recombinant WT i R169T s'ha portat a terme amb el sistema *Qiaexpress Ni-NTA Fast Start* (Qiagen), seguint les instruccions del fabricant. Aquest sistema es basa en que la proteïna recombinant conté una cua de sis residus d'histidina consecutius que s'uneix de manera selectiva i amb una alta afinitat a una columna de cromatografia que conté la resina Ni-NTA (àcid níquel-nitrilotriacetic). Es fan dos rentats per eliminar les unions inespecífiques de proteïnes i dues elucions de la proteïna (Figura 16). A continuació, es carreguen 5 µL de la primera fracció del lisat de les cèl·lules, 5 µL de cadascun dels rentats i 5 µL de cada elució amb 5 µL de *Laemml buffer 2x* en un gel de poliacrilamida. Es transfereixen les proteïnes a una membrana de PVDF i es tenyeix amb *Naphtol blue*.

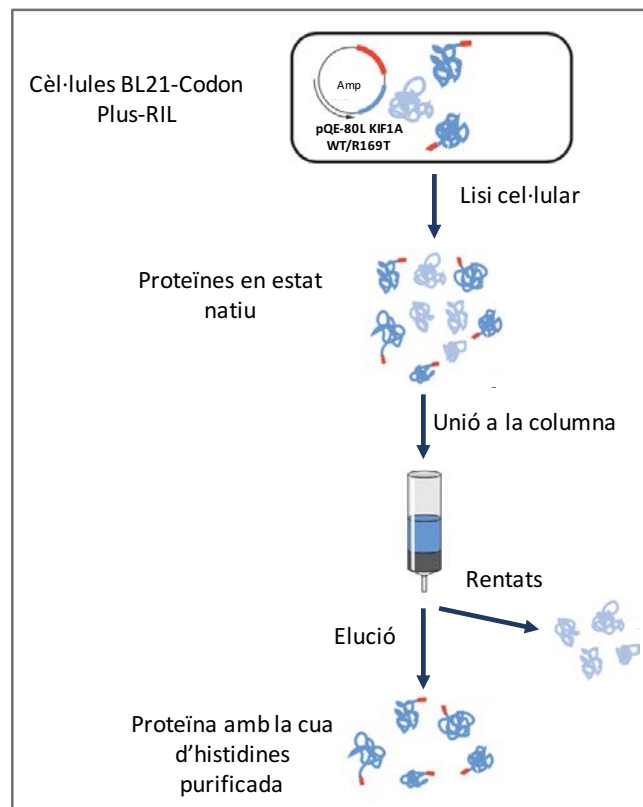


Figura 16 | Esquema del procés de purificació de les proteïnes recombinants. Imatge modificada de *QIAexpress Ni-NTA Fast Start Handbook*.

Un cop purificada la proteïna corresponent al domini motor de KIF1A WT/R169T, s'ha realitzat una diàlisi amb l'objectiu d'eliminar la solució d'elució (50 mM Na-fosfat, 300 mM NaCl i 250 mM imidazole pH 8), que conté altes concentracions d'imidazole que pot donar lloc a la precipitació de la proteïna i a interferències en la seva activitat biològica. La solució de diàlisi consisteix en un volum final de 3 L de: 20 mM Tris-HCl, 300 mM NaCl, 10 mM β -mercaptoetanol, 10% de glicerol i 150 μ M d'ATP, pH 7,4. El volum de les dues fraccions d'elució de cadascun dels dominis motors WT i R169T de KIF1A, s'inocula en cassetes de diàlisi i s'introdueixen en la solució de diàlisi a 4 °C. Es deixa en agitació tota la nit a 4 °C. Al dia següent, s'extreuen els cassetes de diàlisi i el volum es passa a un *ependorf*. Es realitza una centrifuga a màxima velocitat durant 5 min a 4 °C, d'aquesta manera s'elimina la proteïna que pot haver precipitat en el procés de diàlisi. En darrer lloc, per conservar la proteïna recombinant s'han realitzat alíquotes de 20 μ L, que han estat congelades de manera ràpida en nitrogen líquid (*snap freeze*) i emmagatzemades a -80 °C.

3.2.11 Assaig ATPasa del domini motor WT i R169T de KIF1A

L'assaig de l'activitat ATPasa s'ha realitzat amb el kit comercial *Kinesin End Point Assay* (Cytoskeleton) seguint les instruccions del fabricant. Un dels millors mètodes per avaluar o monitoritzar la funció motora de les quinesines és l'avaluació de l'activitat ATPasa. En l'assaig portat a terme es valora la quantitat de fosfat inorgànic (Pi) generat pel domini motor de la quinesina activada per la presència de microtúbuls.

Abans de realitzar l'assaig s'ha quantificat la concentració de la proteïna amb el mètode Bradford fent servir el kit *Bio-Rad Protein Assay* (Biorad). A continuació, es realitza una dilució a 0,5 μ g/ μ L del domini motor de KIF1A WT/R169T i d'aquesta dilució s'ha fet una altra 1/16. A partir d'aquesta última dilució, per realitzar l'assaig s'ha agafat 0,25 μ L, 0,5 μ L, 1 μ L, 2 μ L, 4 μ L i 8 μ L (Figura 17).

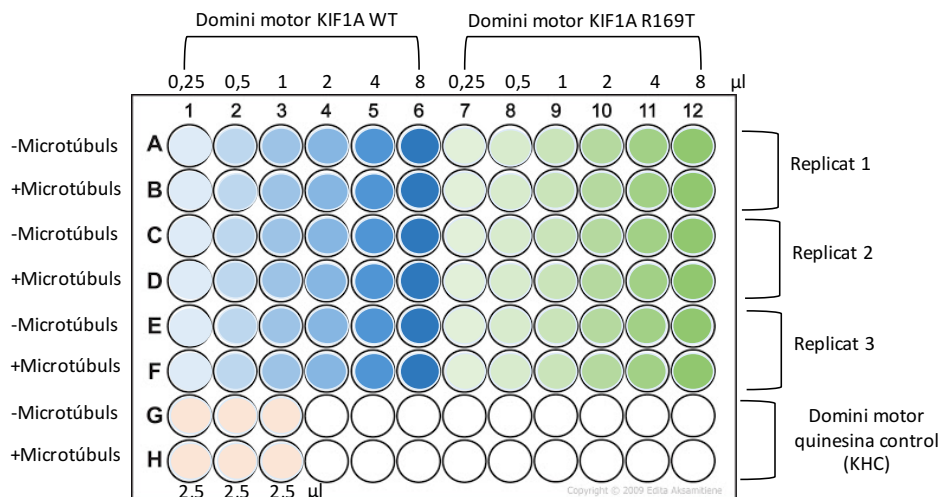


Figura 17 | Disseny de la placa de 96 pous on s'ha portat a terme l'assaig ATPasa.

D'aquesta manera s'obté un rang de 0,078 µg a 0,24 µg de proteïna que va incrementant en un factor 2x i ens permet observar com canvia l'activitat ATPasa a mesura que augmenta la concentració de proteïna.

En l'assaig també s'ha fet servir com a control el domini motor de la quinesina KHC (*Kinesin Heavy Chain*). S'han afegit 2,5 µL de la quinesina control KHC a una concentració de 0,08 µg/µL en els pous corresponents. En les reaccions sense microtúbuls s'afegeix *kinesin reaction buffer* fins a un volum final de 30 µL. En canvi, en les reaccions amb microtúbuls s'afegeix *kinesin reaction buffer* fins a un volum final de 20 µL i 10 µL de microtúbuls polimeritzats i estabilitzats amb taxol a una concentració de 0,2 µg/µL.

Per iniciar la reacció s'afegeix 10 µL d'ATP a cada pou i s'incuba la reacció 5 min a T^a amb. Prèviament, l'ATP s'ha preparat diluint una alíquota de 20 µL d'ATP 100 mM en PIPES *buffer* pH 7 en 1 mL d'aigua freda. Tot seguit, s'afegeixen 70 µL de la solució *CytoPhos* i s'incuba durant 10 min a T^a amb. A l'incorporar en l'assaig la solució *Cytophos* el que es produeix és una reacció colorimètrica que permet quantificar el Pi alliberat. Un cop han transcorregut els 10 min d'incubació, es llegeix la placa en un espectrofotòmetre a una longitud d'ona de 650 nm.

Per establir la correspondència dels valors d'absorbància amb la quantitat de Pi alliberat es realitza una corba estàndard de fosfat. La corba té un rang de concentració de Pi que va de 0,1 nmols a 1,5 nmols. Es llegeixen els valors d'absorbància a 650 nm. A partir de la corba estàndard es calcula l'equació de regressió lineal ($y = 0,1591x + 0,0142$) que servirà per extrapolar els valors d'absorbància obtinguts en l'assaig ATPasa.

Per analitzar si hi ha diferències significatives entre l'activitat ATPasa del domini motor WT i R169T s'ha utilitzat el test estadístic T-Student (IC, 95%).

3.2.12 Anàlisi de xarxes d'interacció

Per a l'anàlisi de les interaccions a nivell de proteïna i de vies moleculars s'han seleccionat les eines STRING, GeneMANIA i Reactome.

STRING (*Search tool for Retrieval of Interacting Genes/Proteins*) (<https://string-db.org/>) és una base de dades que obté la informació de les interaccions proteïna-proteïna integrant altres bases de dades experimentals, de prediccions, d'informació sobre vies moleculars, *text mining* i inter-organismes. Les evidències d'associació es troben categoritzades en set canals que l'usuari pot activar o inactivar: i) tres canals es basen en la informació del context genòmic (*gene neighbourhood*, *gene fusions* i *gene cooccurrence*), ii) co-expressió, iii) *text mining*, iv) dades experimentals bioquímiques/genètiques i v) dades sobre vies moleculars i complexes proteics. El *score* de confiança va de 0 a 1, com més alt és més probable que l'associació entre dues proteïnes

sigui certa, l'usuari pot escollir el *score* de confiança (230). En el nostre cas, s'ha fet servir un *score* de confiança de 0,7 i s'ha provat a inactivar i activar el canal *text mining*, que consisteix en anàlisi estadístic d'un gran nombre de texts científics on es citen a la vegada dues proteïnes, incloent tots els resums de la base de dades de Pubmed i OMIM. D'altra banda, STRING també realitza un anàlisi de les funcions biològiques que es troben enriquides en la xarxa, fent servir el sistema de classificació GO (*Gene Ontology*) i aplica el test exacte de Fisher seguit de la correcció de tests múltiples (231).

Un altre eina que s'ha utilitzat per a l'anàlisi d'interaccions és GeneMania (<http://genemania.org>), que porta a terme la construcció d'una xarxa a partir d'una llista de gens. Per a la construcció de la xarxa es fan servir diferents bases de dades: 1) interacció física i genètica utilitzant BIOGRID, 2) la predicció de la interacció entre proteïnes es basa en l'anàlisi d'ortòlegs de I2D, 3) dades d'interaccions a nivell molecular i en vies de senyalització utilitzant BIOGRID, *Human Protein Reference Database*, HumanCyc, IntAct, MINT i Reactome, entre d'altres. A diferència de STRING, permet personalitzar els resultats, ja que s'escullen quines bases de dades s'utilitzen per realitzar l'anàlisi (232). GeneMANIA fa servir un algoritme que integra diferents xarxes d'interacció. L'algoritme consisteix en dues parts: i) un model de regressió lineal que calcula l'associació funcional de la xarxa obtenint la informació de múltiples bases de dades i ii) un algoritme per predir la funció d'un gen, tenint en compte l'associació funcional d'aquest en la xarxa (233).

A banda de l'anàlisi d'interacció, s'han avaluat les vies moleculars enriquides utilitzant la base de dades Reactome (<https://reactome.org>), que és una base de dades sobre vies moleculars i reaccions biològiques humanes. Les reaccions es consideren com passos de la via molecular que canvien l'estat biològic d'una molècula. Reactome conté dades experimentals i fa servir altres bases de dades com NCBI, Ensembl, Uniprot, KEGG, CHEBI, Pubmed i GO, entre d'altres. Reactome associa proteïnes amb les seves funcions i les uneix en una xarxa. Les vies moleculars es troben organitzades de manera jeràrquica, agrupant les vies moleculars relacionades en dominis més generals de funcions biològiques, fent servir la jerarquia de processos biològics de GO. La base de dades de Reactome té una eina que realitza un càlcul de sobrerrepresentació, donada una llista de proteïnes calcula si aquest llistat conté més proteïnes anotades a cada via molecular de Reactome del que seria esperat per atzar: i) donat el número de proteïnes introduït, ii) el nombre de proteïnes anotades en la via molecular i iii) el total de proteïnes anotades en tota la base de dades de Reactome. Com a resultat, s'obté un p-valor que es troba corregit amb el procediment *Benjamini-Hockberg*, donant lloc al *False Discoverey Rate* (FDR) (234,235) .

En tots tres enfocaments, s'ha introduït un llistat de gens format pel gens responsables i candidats al fenotip SA-like identificats en la nostra cohort (*KIF1A*, *SYNGAP1*, *TBL1XR1*, *SLC6A1*, *ASXL3*, *SATB2*,

KCNQ3, SMARCE1, SPTAN1, LAS1L, VAMP2, CHMP7 i HSF2), els gens del diagnòstic diferencial de la SA (*ADSL, ATRX, CDKL5, EHMT1, FOXP1, KANSL1, MBD5, MECP2, SHANK3, SLC9A6, STXBP1, TCF4, ZEB2, MEF2C, MAPK8IP2, PLXNB2 i HERC2*) i el gen *UBE3A*.

3.2.13 Re-avaluació clínica dels pacients

La re-avaluació clínica s'ha portat a terme en 11 dels 12 pacients amb diagnòstic molecular, amb el fi d'identificar si complien les característiques clíniques associades al gen identificat i si les característiques clíniques no associades a la SA identificades en alguns dels pacients, estan associades al gen identificat. La re-avaluació clínica dels pacients s'ha realitzat per la genetista clínica, la neuropediatra i la psicòloga clínica de la CSPT.



Resultats

Els resultats d'aquesta tesi s'han estructurat en set capítols, el capítol 1 correspon als dos manuscrits derivats d'aquest treball i que responen a l'objectiu general. Els capítols 2, 3, 4, 5, 6 i 7 responen als altres objectius específics plantejats.

El capítol 1 correspon als resultats obtinguts de la seqüenciació de l'exoma complet. Estan escrits en forma de manuscrits, el primer està pendent de sotmetre i el segon està publicat a la revista *Molecular Genetics and Genomic Medicine*.

- *New genes involved in Angelman syndrome-like: expanding the genetic spectrum.*
-Annex I.
- *Identification of a de novo splicing variant in the Coffin-Siris gene, SMARCE1, in a patient with Angelman-like syndrome.*

El capítol 2 inclou l'anàlisi de l'exoma segons els models d'herència lligat al cromosoma X i autosòmic recessiu en els pacients en els quals no s'han identificat variants responsables del fenotip *SA-like de novo*.

El capítol 3 conté els resultats de l'avaluació de la càrrega genètica de variants *de novo* en els pacients de la nostra cohort.

El capítol 4 recull els resultats obtinguts dels anàlisis complementaris (canvis en el número de còpia i variants en mosaic) realitzats als pacients sense diagnòstic molecular, després de l'anàlisi dels models d'herència clàssics.

El capítol 5 correspon a l'estudi de l'activació del mecanisme de control post-transcripcional *Nonsense-Mediated mRNA Decay* (NMD) per variants *nonsense* i *frameshift*.

El capítol 6 recull els resultats obtinguts dels estudis funcionals *in vitro* per avaluar l'impacte de la variant R169T sobre la funció de la quinesina KIF1A.

El capítol 7 consisteix en la construcció de xarxes d'interacció, l'anàlisi de les funcions i vies moleculars enriquides dels gens identificats en la nostra cohort responsables del fenotip *SA-like*, els gens del diagnòstic diferencial de la SA i el gen *UBE3A*.

Capítol 1. Manuscrits

Manuscrit I

New genes involved in Angelman syndrome-like: expanding the genetic spectrum.

Cinthia Aguilera, Elisabeth Gabau, Ariadna Ramirez, Carme Brun, Jana Dominguez, Veronica Delgado, Steve Laurie, Sophia Derdak, Núria Capdevila, Nino Spataro, Neus Baena, Miriam Guitart, Anna Ruiz.

Pendent de sotmetre a una revista.

Aquest manuscrit recull els resultats de la seqüenciació de l'exoma (WES) en 16 dels pacients de la nostra cohort que presenten un fenotip SA, però que no tenien diagnòstic molecular (pacients SA-like). L'aplicació de WES ha permès la identificació de 12 variants patogèniques/probablement patogèniques (11 *de novo* i una variant lligada al cromosoma X) en 11 gens prèviament relacionats amb trastorns del neurodesenvolupament (*KIF1A*, *VAMP2*, *SYNGAP1*, *TBL1XR1*, *SATB2*, *KCNQ3*, *SMARCE1*, *SPTAN1*, *ASXL3*, *LAS1L* i *SLC6A1*) i una variant deletèria en un nou gen candidat (*HSF2*). Aquests resultats indiquen que del 10-15% dels pacients amb un diagnòstic clínic de SA però amb causa molecular desconeguda, el 70,5% presenten variants en altres gens diferents d'*UBE3A*, posant de manifest l'existència d'una gran heterogeneïtat genètica. A més, s'ha demostrat que la seqüenciació de l'exoma en trios és una bona eina diagnòstica i la seva aplicació hauria de ser considerada en pacients amb un fenotip SA-like. Finalment, es proposa que els gens identificats siguin inclosos en el diagnòstic diferencial de la SA.

New genes involved in Angelman syndrome-like: expanding the genetic spectrum

Running title: Expanding the genetic spectrum of AS-like

Cinthia Aguilera¹, Elisabeth Gabau², Ariadna Ramirez-Mallafre², Carme Brun-Gasca^{2,3},
Jana Dominguez-Carral², Veronica Delgadillo², Steve Laurie⁴, Sophia Derdak⁴, Natàlia
Padilla⁵, Xavier de la Cruz^{5,6}, Núria Capdevila², Nino Spataro¹, Neus Baena¹, Miriam
Guitart^{1*} and Anna Ruiz^{1*}

¹Genetics Laboratory, UDIAT-Centre Diagnòstic. Parc Taulí Hospital Universitari.

Institut d'Investigació i Innovació Parc Taulí I3PT. Universitat Autònoma de Barcelona.
Sabadell, Spain.

²Paediatric Unit. Parc Taulí Hospital Universitari. Institut d'Investigació i Innovació Parc
Taulí I3PT. Universitat Autònoma de Barcelona. Sabadell, Spain.

³Department of Clinical Psychology and Health Psychology, Universitat Autònoma de
Barcelona, 08193 Bellaterra, Barcelona, Spain.

⁴CNAG-CRG, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science
and Technology, Baldri Reixac 4, Barcelona 08028, Spain.

⁵Neurosciences Area, Vall d'Hebron Institute of Research (VHIR). Universitat Autònoma
de Barcelona. Barcelona, Spain.

⁶Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

*Corresponding authors:

Dr Anna Ruiz, e-mail: aruizn@tauli.cat, Dr Miriam Guitart, e-mail: mguitart@tauli.cat

Postal address:

Genetics Laboratory, UDIAT-Centre Diagnòstic, Parc Taulí Hospital Universitari. Institut d'Investigació i Innovació Parc Taulí I3PT. Universitat Autònoma de Barcelona.

Parc del Taulí 1, 08208 Sabadell, Barcelona, Spain.

Tel: 0034936933200. Fax: 0034937162004

ABSTRACT

Angelman syndrome (AS) is a neurogenetic disorder characterized by severe developmental delay with absence of speech, happy disposition, frequent laughter, hyperactivity, stereotypies, ataxia and seizures with specific EEG abnormalities. There is a 10-15% of patients with an AS phenotype whose genetic cause remains unknown (Angelman-like syndrome, AS-like). Whole-exome sequencing (WES) was performed on a cohort of 16 patients with clinical features of AS and no molecular diagnosis. As a result, we identified 11 *de novo* and 1 X-linked pathogenic/likely pathogenic variants in 11 neurodevelopmental genes (*SYNGAP1*, *VAMP2*, *TBL1XR1*, *ASXL3*, *SATB2*, *SMARCE1*, *SPTAN1*, *KIF1A*, *KCNQ3*, *SLC6A1* and *LAS1L*) and one deleterious *de novo* variant in a candidate gene (*HSF2*). Our results highlight the wide genetic heterogeneity in AS-like patients and expands the differential diagnosis. New AS-like genes do not interact directly with *UBE3A* gene product but are involved in synapsis and neuron system development.

Keywords: Angelman-like syndrome (AS-like), Whole exome sequencing (WES), *de novo*, genetic heterogeneity, neurodevelopment.

INTRODUCTION

Angelman syndrome (AS, OMIM #105830) is a neurogenetic disorder with a prevalence of about 1/15000 births. AS is characterized by severe developmental delay/intellectual disability (DD/ID) with absence of speech and distinctive dysmorphic craniofacial features such as microcephaly and wide mouth. Neurological problems include ataxia and seizures with specific EEG abnormalities. The behavioral phenotype is characterized by happy disposition, frequent laughter, hyperactivity and stereotypies [1]. The consensus criteria for the clinical diagnosis of AS was proposed in 2005 by Williams et al.,[1] which included a list of (i) consistent, (ii) frequent and (iii) associated features. However, clinical manifestations of AS can overlap with other diseases.

AS is caused by the loss of function in neuronal cells of the ubiquitin protein ligase E6-AP (E6-Associated Protein) encoded by the *UBE3A* gene, which is located on chromosome 15q11-q13 imprinted region. Methylation study of this region identifies 75–80% of AS patients including maternal deletion, paternal uniparental disomy (UPD) and imprinting center defects. Pathogenic or likely pathogenic variants in the *UBE3A* gene identify a further 10% of cases. However, for approximately 10-15% of clinically diagnosed AS patients, the genetic cause remains unknown (AS-like)[2].

Some of these AS-like patients present alternative clinical and molecular diagnoses in syndromes that have overlapping clinical phenotypes and that should be considered in the differential diagnosis of AS. AS differential diagnosis include single gene disorders such as Christianson syndrome (*SLC9A6*), Rett syndrome (*MECP2*), Pitt Hopkins syndrome (*TCF4*), Kleefstra syndrome (*EHMT1*) and Mowat-Wilson syndrome (*ZEB2*). Individuals affected by the above mentioned syndromes present severe DD,

seizures, postnatal microcephaly, absent or minimal speech and sleep disturbances as AS patients [3,4].

In order to further identify the molecular defects in AS-like patients, whole exome sequencing (WES) was performed in a cohort of 15 parent-patient trios and one single patient with clinical features of AS and no molecular diagnosis. Pathogenic/likely pathogenic variants in known neurodevelopmental genes were found in 75% of patients while a deleterious variant in a new candidate gene was identified in another patient. Overall, our results show that 10-15% of patients with a clinical but with no molecular diagnosis of AS present alternative genetic alterations in genes not previously associated to AS, expanding its genetic spectrum.

MATERIALS AND METHODS

Patient samples and clinical description of the cohort

16 patients (8 girls and 8 boys) from the Parc Taulí Hospital Universitari (Sabadell, Spain) who met the consistent clinical features of AS [1] and lacked a molecular diagnosis of AS were selected. Patient 2 had also been included in another study [5]. The corresponding informed consent was obtained from all parents for each participant in the study.

Clinical characteristics of the enrolled patients are available in Table 1. Consistent features were present in 100% of patients except for the ataxia of gait which was present in 10 of 16 patients. Even though the ataxia of gait is considered a consistent feature in AS patients a recent review shows that it ranges from 72,7% to 100 % depending on the genetic etiology [6]. All the cases were sporadic and no other relevant findings were present in their family history.

Whole-exome sequencing and variant interpretation

Trio WES of 15 patients and their parents was performed using the SureSelect Human All Exon V5+UTR kit (Agilent technologies). In patient 5, WES was performed only in the patient sample. Variant discovery was performed following the GATK best practices [7]. All exome variants were first checked against a *de novo* followed by an X-linked and autosomal recessive model of inheritance. The impact of missense and splice site variants was assessed using several *in silico* tools (Supplementary Table 1). Variants were classified following the ACMG guidelines [8]. Pathogenic and likely pathogenic variants have been submitted to ClinVar.

Real time quantitative PCR (RTqPCR) analysis

RNA was extracted using the Biostic Blood Total RNA Isolation Kit sample (MO BIO laboratories, Inc) and cDNA was obtained using the PrimeScriptTMRT reagent Kit (Takara). RTqPCR gene expression analysis was performed using the Taqman probes HSF2-Hs00988309_g1 and GADPH-Hs02758991_g1 for normalization (Applied Biosystems).

Network and pathway enrichment analysis in AS-like genes

Leveraging the STRING database [9], a network analysis was performed in order to examine if the new identified AS-like genes interact among themselves, with the known AS-like genes [3] or with the *UBE3A* gene. Network analysis was carried out without taking into account the text mining interaction option and using a minimum required interaction score of 0.7 (high confidence) (Supplementary Fig.1).

RESULTS

Identified variants were first filtered according to a dominant *de novo* model of inheritance. Variants in genes known to be involved in neurodevelopmental diseases were selected and confirmed to be *de novo*. Overall, 11 *de novo* (*SYNGAP1*, *VAMP2*, *TBL1XR1*, *ASXL3*, *SATB2*, *SMARCE1*, *SPTAN1*, *KIF1A*, *KCNQ3*, *SLC6A1*) and 1 X-linked (*LAS1L*) protein altering variants were confirmed in 12 patients, leading to a diagnostic yield of 75%. These variants were located in 11 different genes previously reported to be associated with neurodevelopmental disorders [5,10–18].

Pathogenic and likely pathogenic variants identified in this study are summarized in Table 2. Most of the genes identified in our cohort are involved in synapsis (*VAMP2*, *SYNGAP1*, *SLC6A1*, *KCNQ3* and *KIF1A*) and chromatin remodeling or transcription regulation (*TBL1XR1*, *SATB2*, *SMARCE1*, *ASXL3* and *LAS1L*) as has been described before in neurodevelopmental diseases [19,20].

Additional clinical features of patients were analyzed taking into account the clinical phenotype described for the genes identified. The presence of specific clinical features associated to the new genes were confirmed for some of them. In short, cerebellar atrophy in *SPTAN1* [11], motor sensory neuropathy, optic nerve and cerebellar atrophy in *KIF1A* [10], hypoplasia of the corpus callosum, hypoplasia of the 5th finger nail, hypertrichosis, sparse scalp hair and aggressive behavior in *SMARCE1* [21], truncal obesity and short stature in *LAS1L* [14], myoclonic atonic seizures in *SLC6A1* [13], aggressive behavior in *SYNGAP1* [15], dysmorphic features and dental anomalies in *ASXL3* [12] and aggressive behavior and dental anomalies in *SATB2* [17] (Fig.1). However, not all patients presented all the clinical features associated with the genes identified. Indeed, unsteady gait and hypotonia were not present in patient carrying the pathogenic

variant in *SYNGAP1* [15]; similarly the patient harboring a pathogenic variant in *SATB2* did not show sialorrhea and feeding difficulties [17]. Finally, the ataxia of gait, stereotypies and hypotonia were not observed for the patient with a pathogenic variant in *KCNQ3* [18].

A novel candidate variant was identified in a gene not previously associated with neurodevelopmental disorders. The identified variant is a *de novo* frameshift deletion c.456_459delTGAG (NM_004506.3), p.(Ser152Argfs*40) in *HSF2* gene. The variant has not been reported before and is not present in the gnomAD database. Quantification of mRNA transcripts showed a reduction in the allele carrying the frameshift variant suggesting the activation of the nonsense-mediated mRNA decay (NMD) machinery [22], supporting a loss of function mechanism of disease for the *HSF2* gene (Supplementary Fig.2).

To assess whether the genes identified in our study are functionally related to previously known AS-like genes or *UBE3A*, we tested their connectivity in the human protein-protein interaction network. No direct protein interactions were detected between the genes identified in our study, the known AS-like genes nor *UBE3A*. However, a significant enrichment in GO terms related to the nervous system development (biological process), protein N-terminus binding (molecular function) and synapse (cellular component) were observed. Overall, these results suggest that the genes found in our cohort, the *UBE3A* gene and the known AS-like genes may not interact directly but carry out biological functions involved in synapsis and brain development, leading to overlapping phenotypes (Supplementary Fig.1).

DISCUSSION

We identified causal variants in 12 out of 16 patients with an AS-like phenotype. The global yield diagnostic of WES in this study is 75%, which is higher to what has been reported in the literature for other neurodevelopmental disorders (24-68%) [23]. The results of WES led to the identification of 11 new genes to cause an AS-like phenotype (*SYNGAP1*, *VAMP2*, *TBL1XR1*, *ASXL3*, *SATB2*, *SMARCE1*, *SPTAN1*, *KIF1A*, *KCNQ3*, *SLC6A1* and *LAS1L*), all of them previously associated with other neurodevelopmental disorders. In addition, we propose *HSF2* (Heat Shock Factor) as a new candidate gene for the AS-like phenotype. Although *HSF2* has not been previously associated with any human disease, the gene is highly expressed in the brain and highly intolerant to loss of function variation (pLI 0.92). *HSF2* knockout mice show defects in spermatogenesis and in the development of the central nervous system [24,25]. The identification of additional patients with loss of function variants in *HSF2* and functional studies in neural cells will contribute to elucidate the role of *HSF2* in the AS-like phenotype.

De novo variants have been described to account for approximately half of the genetic architecture of severe developmental disorders [26]. In our cohort, 11 of the 12 pathogenic and likely pathogenic variants were *de novo*, accounting for 68% of diagnosis and highlighting the power of using trio-WES for the molecular diagnosis of severe developmental disorders. Only in one case, the X-linked variant in *LAS1L* was inherited from the mother, who was a healthy mosaic carrier (20%).

All patients had received an initial diagnosis of AS, supported by the presence of consistent and frequent clinical features. In the majority of our patients (13/16) the initial diagnosis was done during infancy or early childhood (before five years old). At

the time of initial diagnosis, all of them presented severe global DD and speech impairment in addition to ataxia of gait or the characteristic happy disposition. The review of patient's clinical reports upon WES results showed the presence of additional clinical features generally not described for AS, but that were then associated with the new identified genes. Pathogenic/likely pathogenic variants in *SMARCE1*, *SATB2*, *SYNGAP1*, *SLC6A1*, *KIF1A*, *ASXL3*, *SPTAN1* and *LAS1L* genes are associated with neurodevelopment disorders that overlap with AS and with some differential features that were present in our patients (Fig.1). On the other hand, *VAMP2*, *KCNQ3* and *TBL1XR1* genes are associated with ID, autism spectrum disorder and epilepsy [5,16,18], features that are shared by AS-like patients.

Lack of molecular diagnosis in 10-15% of clinically diagnosed AS patients has been used to define the AS-like group. Our results indicate that the 75% of AS-like patients are carriers of pathogenic variants in genes involved in neurodevelopmental disorders whose features overlap with AS, showing the wide genetic heterogeneity in AS-like (Fig. 1). AS-like new genes do not interact directly with *UBE3A* gene product but are involved in synapsis and nervous system development. Except for the *SYNGAP1* gene, none of the genes identified here have been previously described in the differential diagnosis of AS [27]. We propose the genes identified in this study should be included in the AS differential diagnosis and that trio WES should be considered as first line approach for the molecular diagnosis of AS-like patients. A high rate of diagnosis in patients with AS-like is essential, contributing to more appropriate clinical patient surveillance as well as allowing family genetic counseling and making reproductive decisions.

ACKNOWLEDGEMENTS

We thank the patients and families for their participation in this study and Instituto de Salud Carlos III (PI16/01411), Asociación Española de Síndrome de Angelman, Institut d'investigació i innovació Parc Taulí I3PT (CIR2016/025, CIR2018/021) and Ministerio de Economía y Competitividad (SAF2016-80255-R) for their financial support.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

1. Williams C, Beaudet AL, Clayton-Smith J, Knoll JH, Kyllerman M, Laan LA, et al. Angelman Syndrome 2005: Updated Consensus for Diagnostic Criteria Charles. *Am J Med Genet Part A*. 2006;140:413–8.
2. Beygo J, Buiting K, Ramsden SC, Ellis R, Clayton-Smith J, Kanber D. Update of the EMQN/ACGS best practice guidelines for molecular analysis of Prader-Willi and Angelman syndromes. *Eur J Hum Genet*. 2019 Sep;27:1326–40.
3. Tan WH, Bird LM, Thibert RL, Williams CA. If not Angelman, what is it? A review of Angelman-like syndromes. *Am J Med Genet Part A*. 2014;164:975–92.
4. Luk H-M. Angelman-Like Syndrome: A Genetic Approach to Diagnosis with Illustrative Cases. *Case Rep Genet*. 2016;2016:1–6.
5. Salpietro V, Malintan NT, Llano-Rivas I, Spaeth CG, Efthymiou S, Striano P, et al. Mutations in the Neuronal Vesicular SNARE VAMP2 Affect Synaptic Membrane Fusion and Impair Human Neurodevelopment. *Am J Hum Genet*. 2019 Apr;104:721–30.
6. Bell L, Wittkowski A, Hare DJ. Movement Disorders and Syndromic Autism: A Systematic Review. *J Autism Dev Disord*. 2019 Jan;49:54–67.
7. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinforma*. 2013;43:11.10.1-33.
8. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics

- and the Association for Molecular Pathology. *Genet Med.* 2015;17:405–23.
9. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 2017;45:D362–8.
 10. Esmaeeli Nieh S, Madou MRZ, Sirajuddin M, Fregeau B, Mcknight D, Lexa K, et al. De novo mutations in KIF1A cause progressive encephalopathy and brain atrophy. *Ann Clin Transl Neurol.* 2015;2:623–35.
 11. Syrbe S, Harms FL, Parrini E, Montomoli M, Mutze U, Helbig KL, et al. Delineating SPTAN1 associated phenotypes: from isolated epilepsy to encephalopathy with progressive brain atrophy. *Brain.* 2017 Sep;140:2322–36.
 12. Balasubramanian M, Willoughby J, Fry AE, Weber A, Firth H V, Deshpande C, et al. Delineating the phenotypic spectrum of Bainbridge-Ropers syndrome: 12 new patients with de novo , heterozygous, loss-of-function mutations in ASXL3 and review of published literature. *J Med Genet.* 2017;54:537–43.
 13. Johannesen KM, Gardella E, Linnankivi T, Courage C, Martin A de Saint, Lehesjoki A-E, et al. Defining the phenotypic spectrum of SLC6A1 mutations. *Epilepsia.* 2018;59:389–402.
 14. Hu H, Haas S, Chelly J, Esch H Van, Raynaud M, Brouwer A de, et al. X-exome sequencing of 405 unresolved families identifies seven novel intellectual disability genes. *Mol Psychiatry.* 2015;21:133–48.
 15. Mignot C, von Stulpnagel C, Nava C, Ville D, Sanlaville D, Lesca G, et al. Genetic and neurodevelopmental spectrum of SYNGAP1-associated intellectual disability and epilepsy. *J Med Genet.* 2016 Aug;53:511–22.
 16. Laskowski RA, Thornton JM, Tyagi N, Johnson D, McWilliam C, Kinning E, et al.

- Integrating population variation and protein structural analysis to improve clinical interpretation of missense variation: application to the WD40 domain. *Hum Mol Genet.* 2016;25:927–35.
17. Zarate YA, Smith-Hicks CL, Greene C, Abbott MA, Siu VM, Calhoun ARUL, et al. Natural history and genotype-phenotype correlations in 72 individuals with SATB2-associated syndrome. *Am J Med Genet Part A.* 2018;176:925–35.
 18. Sands TT, Miceli F, Lesca G, Beck AE, Sadleir LG, Arrington DK, et al. Autism and developmental disability caused by KCNQ3 gain-of-function variants. *Ann Neurol.* 2019 Jun;
 19. Hormozdiari F, Penn O, Borenstein E, Eichler EE. The discovery of integrated gene networks for autism and related disorders. *Genome Res.* 2015 Jan;25:142–54.
 20. Kochinke K, Zweier C, Nijhof B, Fenckova M, Cizek P, Honti F, et al. Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am J Hum Genet.* 2016 Jan;98:149–64.
 21. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Aguilera MA, Meyer R, et al. VarSome: the human genomic variant search engine. *Bioinformatics.* 2018;1–3.
 22. Lykke-Andersen S, Jensen TH. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol.* 2015 Nov;16:665–77.
 23. Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *npj Genomic Med.* 2018;3:1–10.

24. Wang G, Zhang J, Moskophidis D, Mivechi NF. Targeted disruption of the heat shock transcription factor (hsf)-2 gene results in increased embryonic lethality, neuronal defects, and reduced spermatogenesis. *Genesis*. 2003;36:48–61.
25. Kallio M, Chang Y, Manuel M, Alastalo T-P, Rallu M, Gitton Y, et al. Brain abnormalities, defective meiotic chromosome synapsis and female subfertility in HSF2 null mice. *EMBO J*. 2002 Jun;21:2591–601.
26. McRae J, Clayton S, Fitzgerald T, Kaplanis J, Prigmore E, Rajan D, et al. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017 Feb;542:433–8.
27. Parker MJ, Magee AC, Maystadt I, Benoit V, study D, FitzPatrick DR, et al. De novo, heterozygous, loss-of-function mutations in SYNGAP1 cause a syndromic form of intellectual disability . *Am J Med Genet Part A*. 2015;167:2231–7.

FIGURE LEGENDS

Fig.1. Schematic representation of the phenotypic overlap between the patients with pathogenic/likely pathogenic variants genes and the AS phenotype. In the middle of the figure there are the core features of AS that show all the patients, while in the tips there are the clinical features present in the patients of our cohort and that are associated with the gene identified.

SUPPORTING INFORMATION

Supplementary Fig.1. Functional and enrichment network analysis of protein-protein interactions of the candidate genes identified in our cohort, the known AS-like genes

and *UBE3A*, using **STRING**. **A)** The figure shows the interaction pattern between the different genes; circumferences and edges represent the genes and their interactions, respectively. Edges are colored according to whether there is a known interaction, a predicted interaction or others. **B)** The table shows the GO terms significantly represented amongst the candidate genes. The analysis is repeated for the three GO domains: Biological Process, Molecular Function and Cellular Component. Overall, we see that the significant terms refer to different aspects of the nervous system, from development to synapses.

Supplementary Fig.2. Quantification of *HSF2* mRNA transcripts suggest that variant c.456_459delTGAG is sensitive to NMD. A) qPCR analysis of *HSF2* gene expression in patient 13 and a control sample normalized to GAPDH shows less *HSF2* expression in patient 13 (* p-value 0.014). **B)** Sanger sequencing of a fragment encompassing variant c.456_459delTGAG from patient 13 and a control sample shows a reduction in the percentage of the allele with the variant in the cDNA compared to DNA.

Table 1. Clinical characteristics of AS-like patients.

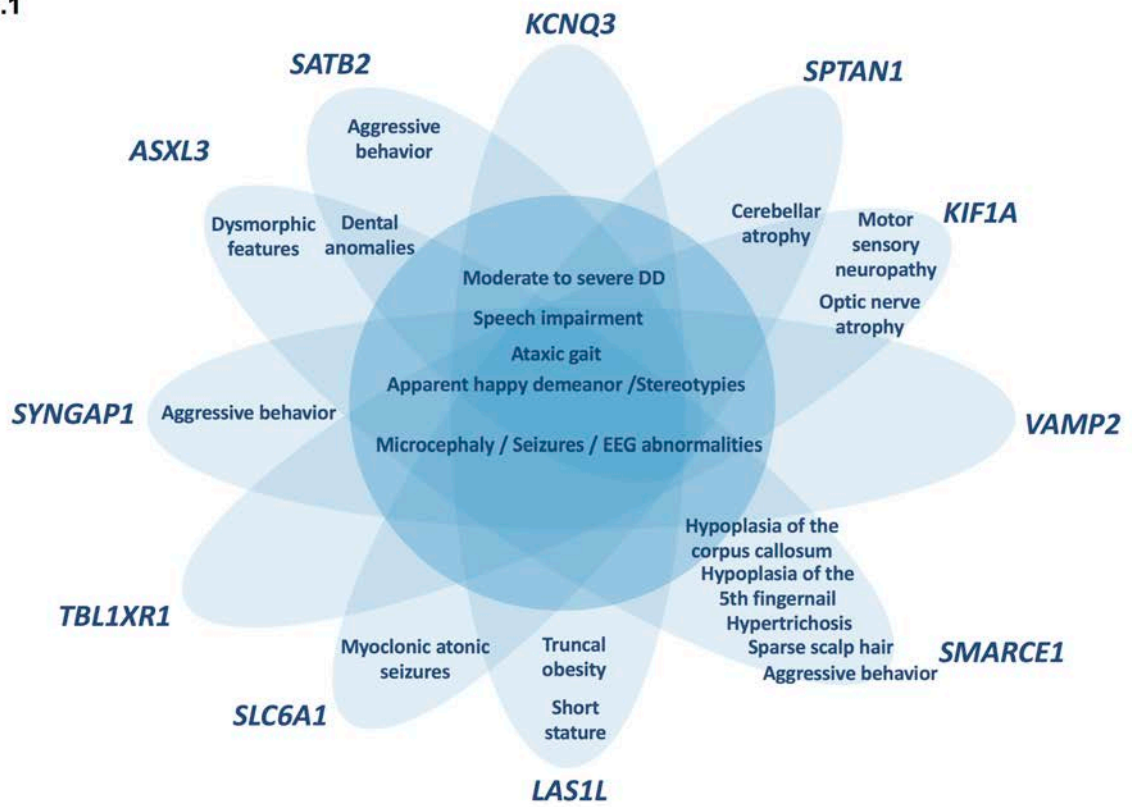
Patient	Gender	Age at clinical diagnosis (years)	Age at molecular diagnosis (years)	Consistent features present in 100% of patients with AS				Frequent features present in more than 80% of AS affected individuals			Associated features present in 20-80% of AS affected individuals	Other additional relevant clinical findings not associated to AS
				Severe developmental delay	Speech impairment	Ataxia or unsteady gait	Ataxia or happy demeanor/ Stereotypies	Microcephaly	Seizures	Abnormal EEG		
1	M	10 months	8	+	+(More than 20 words)	+	+/-	+	+	+	+	Optic and progressive cerebellar atrophy, motor sensory neuropathy, hyperreflexia, congenital torticollis
2	M	1	14	+	+(5-10 words)	+	+/-	-	+	+	+	Congenital torticollis; bruxism, aggressive behavior
3	F	8	19	+	+(less than 5 words)	-	+/-	-	+	+	+	Aggressive behavior
4	F	1	12	+	+(Absent speech)	+	-/+	-	+	+	+	Aggressive behavior
5	F	11 months	9	+	+(Absent speech)	+	+/-	+(Relative)	-	+	+	Aggressive behavior
6	F	19 months	20	+	+(5-10 words)	+	+/-	-	-	+	-	Dental anomalies, auto and hetero-aggressive behavior
7	F	4	18	+	+(Absent speech)	-	+/-	+(Relative)	+	+	NA	-
8	M	5	15	+	+(less than 5 words)	+	+/-	-	-	+	+	Sparse scalp hair, hypertrichosis in the back and hypoplasia of the corpus callosum, hypoplastic 5 th fingernail, auto and hetero-aggressive behavior
9	F	1	14	+	+(More than 20 words)	+	+/-	+	-	-	-	Cerebellar atrophy
10	M	8	38	+	+(Absent speech)	+	+/-	+(Relative)	+	+	+	Dental anomalies, bruxism, episodic hyperventilation
11	M	1	7	+	+(less than 5 words)	-	+/-	+(Relative)	-	-	-	Truncal obesity, short stature
12	F	2	14	+	+(less than 5 words)	-	+/-	+(Relative)	+	+	NA	-
13	M	3	24	+	+(5-10 words)	-	+/-	+	+	+	NA	Colpocephaly and hypoplasia of the corpus callosum, abnormal behavior, hypothyroidism and bruxism
14	F	4	28	+	+(More than 20 words)	-	+/-	+(Relative)	+	+	+	Dental abnormalities, global brain atrophy, nonspecific leukoencephalopathy, episodic hyperventilation
15	M	3	9	+	+(Absent speech)	+	+/-	+(Relative)	+	-	-	Episodic hyperventilation, mild subcortical atrophy
16	M	2	13	+	+(Absent speech)	+	+/-	-	+	+	+	Myoclonic atonic seizures, bruxism

M, Male; F, Female; +, present; -, not present; NA, non-available data.

Table 2. Pathogenic and likely pathogenic variants identified in AS-like patients.

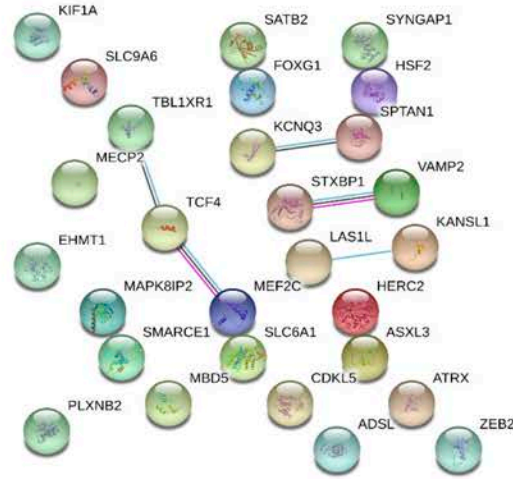
Patient	Gene	NIM number	Nucleotide change	Amino acid change	Variant Type	Pattern of inheritance	Varsome Classification	Described before	Protein function
1	<i>KIF1A</i>	NM_001244008.1	c.506G>C	p.Arg169Thr	Missense	De novo	Likely pathogenic	No	<i>KIF1A</i> belongs to the kinesin family and acts as an anterograde motor protein.
2	<i>VAMP2</i>	NM_014232.2	c.128_130delTTGG	p.Val43del	In-frame	De novo	Pathogenic	Yes Salpietro et al., 2019	<i>VAMP2</i> is a member of the SNARE family of proteins, which are involved in membrane fusion of synaptic vesicles.
3	<i>SYNGAP1</i>	NM_006772.2	c.1861C>T	p.Arg621*	Nonsense	De novo	Pathogenic	No	<i>SYNGAP1</i> is a RAS-GTPase-activating protein with a critical role in synaptic development, structure, function and plasticity.
4	<i>TBL1XR1</i>	NM_024665.5	c.1000T>C	p.Cys334Arg	Missense	De novo	Likely pathogenic	No	<i>TBL1XR1</i> is part of the repressive NCoR/SMRT complex acting as a transcriptional regulator.
5	<i>TBL1XR1</i>	NM_024665.5	c.1043A>G	p.His348Arg	Missense	De novo	Likely pathogenic	No	<i>SATB2</i> participates in chromatin remodeling and transcription regulation.
6	<i>SATB2</i>	NM_001172509.1	c.1826delA	p.Asp609Alafs*15	Frameshift	De novo	Pathogenic	No	<i>KCNQ3</i> is a voltage-gated potassium channel subunits that underlay the neuronal M-Current.
7	<i>KCNQ3</i>	NM_004519.3	c.688C>T	p.Arg230Cys	Missense	De novo	Pathogenic	Yes Decipher and Miceli F et al., 2015, Sands TT et al., 2019	<i>SMARCE1</i> is part of the SWI/SNF chromatin remodeling complex involved in transcriptional activation.
8	<i>SMARCE1</i>	NG_032163.1 (NM_003079.4)	c.237+1G>T	p.Ala53_Lys79del	Splice site	De novo	Likely pathogenic	No	<i>SPTAN1</i> is an α -II spectrin involved in stabilization and activation of membrane channels, transporters and receptors.
9	<i>SPTAN1</i>	NM_001130438.2	c.6592_6597dupCTGCAG	p.Leu2198_Gln2199dup	In-frame	De novo	Likely pathogenic	No	<i>ASXL3</i> plays a role in the regulation of gene transcription and histone deubiquitination.
10	<i>ASXL3</i>	NM_030632.2	c.3106C>T	p.Arg1036*	Nonsense	De novo	Pathogenic	Yes Kuechler A et al., 2016	<i>LAS1L</i> is involved in the 60S ribosomal subunit synthesis, maturation of 28S rRNA and regulation of transcription.
11	<i>LAS1L</i>	NM_031206.4	c.1237G>A	p.Gly413Arg	Missense	X-linked	Likely pathogenic	No	<i>SLC6A1</i> gene encodes for the GAT-1 GABA transporter.
16	<i>SLC6A1</i>	NM_003042.3	c.889G>A	p.Gly297Arg	Missense	De novo	Pathogenic	Yes Carvill GL et al., 2015	

Fig.1



Supplementary Fig.1

A)



Node Legend

Node Color

- colored nodes: query proteins and first shell of interactors
- white nodes: second shell of interactors

Node Content

- empty nodes: proteins of unknown 3D structure
- filled nodes: some 3D structure is known or predicted

Edge Legend

Known Interactions

- from curated databases
- experimentally determined

Predicted Interactions

- gene neighborhood
- gene fusions
- gene co-occurrence

Others

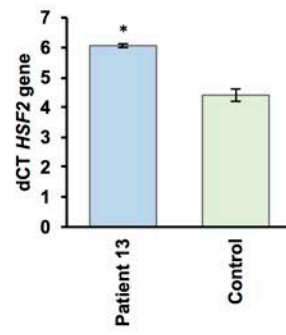
- co-expression
- protein homology

B)

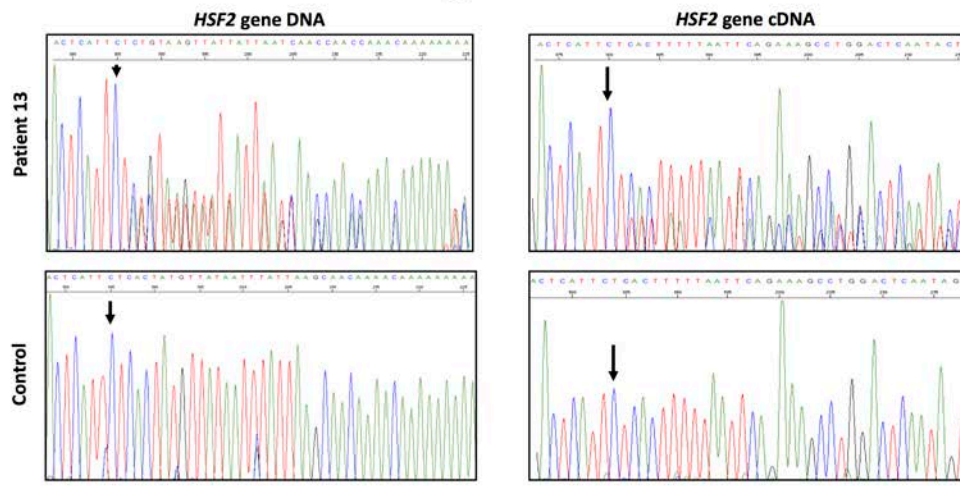
Biological Process (Top 20 GO terms)			
GO-term	Description	Count in gene set	False discovery rate
GO:0007399	Nervous system development	17 of 2206	7.64e-07
GO:0022008	Neurogenesis	14 of 1519	3.23e-06
GO:0048699	Generation of neurons	13 of 1422	1.10e-05
GO:0030182	Neuron differentiation	10 of 940	0.00015
GO:0016043	Cellular component organization	20 of 5163	0.00022
GO:0048666	Neuron development	9 of 758	0.00022
GO:0048731	System development	18 of 4144	0.00023
GO:0007268	Chemical synaptic transmission	7 of 402	0.00024
GO:0031175	Neuron projection development	8 of 616	0.00024
GO:0048667	Cell morphogenesis involved in neuron differentiation	7 of 400	0.00024
GO:0030154	Cell differentiation	16 of 3457	0.00032
GO:0048468	Cell development	11 of 1493	0.00032
GO:0006325	Chromatin organization	8 of 683	0.00033
GO:0048167	Regulation of synaptic plasticity	5 of 164	0.00033
GO:0051962	Positive regulation of nervous system development	7 of 488	0.00036
GO:0050804	Modulation of chemical synaptic transmission	6 of 316	0.00038
GO:0045666	Positive regulation of neuron differentiation	6 of 337	0.00047
GO:0016358	Dendrite development	4 of 90	0.00054
GO:0000122	Negative regulation of transcription by RNA polymerase II	8 of 809	0.00073
GO:0051960	Regulation of nervous system development	8 of 817	0.00076
Molecular Function (GO terms)			
Go term	Description	Count in gene set	False discovery rate
GO:0047485	Protein N-terminus binding	4 of 109	0.0048
GO:0003682	Chromatin binding	6 of 501	0.0083
GO:0000981	DNA-binding transcription factor activity, RNA polymerase II-specific	9 of 1631	0.0220
GO:0001228	DNA-binding transcription activator activity, RNA polymerase II-specific	5 of 408	0.0220
GO:0003677	DNA binding	11 of 2457	0.0220
GO:0005515	Protein binding	18 of 6605	0.0220
GO:0017075	Syntaxin-1 binding	2 of 20	0.0220
GO:0019904	Protein domain specific binding	6 of 706	0.0220
GO:0043565	Sequence-specific DNA binding	7 of 1047	0.0220
GO:0140110	Transcription regulator activity	10 of 2069	0.0220
Cellular Component (Top 20 GO terms)			
GO-term	Description	Count in gene set	False discovery rate
GO:0045202	Synapse	9 of 849	0.00045
GO:0044456	Synapse part	8 of 705	0.00059
GO:0097458	Neuron part	10 of 1449	0.0016
GO:0098794	Postsynapse	6 of 435	0.0019
GO:0005694	Chromosome	8 of 950	0.0020
GO:0031981	Nuclear lumen	15 of 4030	0.0046
GO:0043005	Neuron projection	8 of 1142	0.0046
GO:0043232	Intracellular non-membrane-bounded organelle	15 of 4005	0.0046
GO:0044427	Chromosomal part	7 of 819	0.0046
GO:0005654	Nucleoplasm	13 of 3446	0.0085
GO:0000785	Chromatin	5 of 489	0.0107
GO:0070013	Intracellular organelle lumen	16 of 5162	0.0107
GO:0030424	Axon	5 of 530	0.0125
GO:0071339	MLL1 complex	2 of 31	0.0138
GO:0000790	Nuclear chromatin	4 of 333	0.0156
GO:0033267	Axon part	4 of 341	0.0162
GO:0005634	Nucleus	18 of 6892	0.0177
GO:0098793	Presynapse	4 of 354	0.0177
GO:0036477	Somatodendritic compartment	5 of 731	0.0361
GO:0044451	Nucleoplasm part	6 of 1073	0.0361

Supplementary Fig.2

A)



B)



Annex I. Manuscrit I

El trio 17 (pacient 17, mare i germana) s'ha exclòs del manuscrit perquè no es disposa de les dades genòmiques del pare (èxitus) i no és possible demostrar que les variants candidates identificades siguin *de novo*.

L'anàlisi del model d'herència *de novo* en el trio 17 s'ha realitzat filtrant per totes aquelles variants presents en el pacients però absents tant en la mare com en la germana (genotip 0/0). S'han obtingut un total de 1054 variants *de novo*, que han estat filtrades per: i) variants amb un efecte potencial sobre la proteïna (*nonsense*, *frameshift*, *splicing* i *missense*) i ii) una freqüència poblacional en les bases de dades d'ExAC i gnomAD inferior a 1/10.000. Com a resultat del filtratge, s'han obtingut 37 variants (Taula 10). D'aquestes variants, s'ha seleccionat la variant c.1165C>T, p.Gln389* en el gen *CHMP7* com a candidata perquè és una variant *nonsense* que trunca la proteïna i no es troba en les bases de dades de població control ExAC i gnomAD. Es tracta d'un gen que és intolerant als canvis de pèrdua de funció (pLI 0.86) i s'expressa principalment en el còrtex cerebral, sistema digestiu, teixits endocrins, pulmó i múscul (236,237). Malgrat que el gen *CHMP7* no ha estat associat prèviament amb trastorns del neurodesenvolupament (TND), aquest gen és un bon candidat a ser responsable del fenotip SA-like perquè pertany a la família de complexes ESCRT (*Endosomal Sorting Complexes Required for Transport*), els quals estan implicats en neurodegeneració i en el neurodesenvolupament. Seria necessari portar a terme estudis funcionals que demostrin els efectes de la variant p.Gln389* sobre la funció de la proteïna i la seva implicació en el neurodesenvolupament.

Taula 10| Variants obtingudes del filtratge segons un model d'herència dominant *de novo* en el pacient 17.

Posició cromosòmica	Tipus de variant	Gen	Isoforma	Canvi nucleòtid	Canvi d'aminoàcid	ExAC/gnomAD
1:43637165	Missense	<i>EBNA1BP2</i>	NM_001159936.1	c.473A>C	p.Glu158Pro	0,0000335/ 0,00008642
1: 114682293	Missense	<i>SYT6</i>	NM_001270805.1	c.201C>G	p.His67Gln	- / 0,000004069
1: 149290670	Missense	<i>FAM231D</i>	NM_001348147.1	c.142G>C	p.Gly48Arg	- / -
1: 152277924	Missense	<i>FLG</i>	NM_002016.2	c.9438G>C	p.Gln3146His	- / -
1:155263070	Missense	<i>PKLR</i>	NM_000298.6	c.1334C>T	p.Ala445Val	0,00002472/ 0,00003655
1:161201242	Missense	<i>NR1I3</i>	NM_001077482.2	c.571A>T	p.Ile191Phe	- / -
2:9002794	Missense	<i>MBOAT2</i>	NM_138799.4	c.1111A>G	p.Ile371Val	0,000008244 / 0,00004468
2:43458445	Missense	<i>THADA</i>	NM_001083953.1	c.5504A>C	p.Asn1835Thr	0,00001272 / 0,00000491
3:38748862	Missense	<i>SCN10A</i>	NM_006514.3	c.4294G>A	p.Asp1432Asn	0,00002473/ 0,00003251

4. RESULTATS

Taula 10| Continuació.

3:52326388	Missense	GLYCK	NM_145262.4	c.818G>A	p.Arg273His	0,000008267 / 0,0000122
3:52366244	Missense	DNAH1	NM_015512.5	c.1120C>T	p.Arg374Cys	0,00002921/ 0,0000326
3:123301174	Nonsense	HACD2	NM_001329783.2	c.158G>A	p.Trp53*	- / 0,000004692
3:132244526	Missense	DNAJC13	NM_001329126.2	c.6148G>A	p.Val2050Ile	0,00004119/ 0,00002031
6:110112598	Missense	FIG4	NM_014845.6	c.2200G>A	p.Glu734Lys	0,00007509 / 0,00005718
8:23117701	Nonsense	CHMP7	NM_152272.5	c.1165C>T	p.Gln389*	- / -
9:129642414	Missense	ZBTB34	NM_001099270	c.724G>A	p.Asp242Asn	0,00001664 / 0,00000817
10:101816831	Frameshift	CPN1	NM_001308.3	c.949delG	p.Glu317Lysfs*15	- / 0,000008123
10:106075653	Missense	ITPRIP	NM_033397.3	c.157T>A	p.Leu53Met	- / -
11:8669593	Missense	TRIM66	NM_014818.2	c.331C>T	p.Leu111Phe	- / -
12:69652785	Missense	CPSF6	NM_001300947.2	c.1221G>A	p.Met407Ile	0,00001656 / 0,000008188
12:77421879	Missense	E2F7	NM_203394.3	c.1924A>G	p.Met642Val	0,00001649 / 0,00001218
12:108145600	Missense	PRDM4	NM_012406.4	c.718G>C	p.Val240Leu	0,00001648 / 0,00001218
12:120613012	Missense	GCN1	NM_006836.2	c.1046C>T	p.Ser349Leu	0,00003313 / 0,00002438
14:73137989	Missense	DPF3	NM_001280544.1	c.1094C>T	p.Thr365Met	0,00003319 / 0,00005315
15:45397846	Missense	DUOX2	NM_014080.4	c.2329G>A	p.Ala777Thr	- / 0,000004065
15:90167586	Missense	TICRR	NM_001308025.1	c.4042G>A	p.Val1348Ile	0,00008244 / 0,00008127
15:91030800	Missense	IQGAP1	NM_003870.4	c.4234G>A	p.Ala1412Thr	0,0000083 / 0,00002032
16:1657233	Missense	IFT140	NM_014714.4	c.35C>T	p.Pro12Leu	0,00002472 / 0,00002843
17:19642904	Missense	ALDH3A1	NM_001135168.1	c.1033G>A	p.Val345Met	0,000008925 / 0,0000204
17:26976197	Missense	SDF2	NM_006923.4	c.446G>A	p.Arg149Gln	0,00003295 / 0,00003248
17:35300196	Missense	LHX1	NM_005568.5	c.989C>T	p.Pro330Leu	- / -
17:38349645	Missense	RAPGEFL1	NM_001303534.2	c.1369C>T	p.Arg457Cys	0,00002471 / 0,00001218
17:73038316	Missense	ATP5H	NM_006356.3	c.179A>T	p.Asn60Ile	- / -
18:8720366	Missense	MTCL1	NM_015210.4	c.229C>G	p.His77Asp	- / -
20:56137869	Missense	PCK1	NM_002591.4	c.524G>C	p.Arg175Pro	0,000008246 / 0,000004062
20:56138726	Missense	PCK1	NM_002591.4	c.904G>A	p.Gly302Arg	0,0000495 / 0,00004877
22:29130622	Missense	CHEK2	NM_007194.4	c.88G>A	p.Gly30Ser	- / -

Manuscrit II

Identification of a *de novo* splicing variant in the Coffin-Siris gene, *SMARCE1*, in a patient with Angelman-like syndrome


Cinthia Aguilera, Elisabeth Gabau, Steve Laurie, Neus Baena, Sophia Derdak, Nuria Capdevila, Ariadna Ramírez, Verónica Delgadillo, María Jesús García-Catalan, Carme Brun, Miriam Guitart, Anna Ruiz.

Molecular Genetics and Genomic Medicine 2019 Jan, 7(1):e00511.

DOI: 10.1002/mgg3.511

Aquesta publicació es centra en el pacient 8 de la nostra cohort en el qual s'ha identificat la variant *de novo* c.237+1G>T en el gen *SMARCE1*. Variants *missense* en el gen *SMARCE1* donen lloc a la síndrome de Coffin-Siris (CSS), mentre que variants truncants causen meningiomes cranials o espinals. L'anàlisi de l'ADNc del gen *SMARCE1* mostra que la variant de *splicing* c.237+1G>T dóna lloc a la deleció de l'exó 5 (*exon skipping*) sense alterar la pauta de lectura. En la proteïna, s'eliminen 27 aminoàcids situats al final del domini ric en prolina i l'inici del domini HMG, aquest últim domini és essencial pel correcte funcionament de la proteïna. La re-avaluació clínica del pacient ha confirmat la presència de característiques clíniques de la CSS, moltes de les quals solapen amb la SA. Es proposa que la CSS sigui inclosa en el diagnòstic diferencial de la SA.

Identification of a de novo splicing variant in the Coffin–Siris gene, *SMARCE1*, in a patient with Angelman-like syndrome

Cinthia Aguilera¹  | Elisabeth Gabau² | Steve Laurie³ | Neus Baena¹ | Sophia Derdak³ | Núria Capdevila² | Ariadna Ramirez² | Veronica Delgadillo² | Maria Jesus García-Catalan² | Carme Brun² | Miriam Guitart¹ | Anna Ruiz¹

¹Genetics Laboratory, UDIAT-Centre Diagnòstic, Parc Taulí Hospital Universitari, Institut d'Investigació i Innovació Parc Taulí I3PT, Universitat Autònoma de Barcelona, Sabadell, Spain

²Paediatric Unit, Parc Taulí Hospital Universitari, Institut d'Investigació i Innovació Parc Taulí I3PT, Universitat Autònoma de Barcelona, Sabadell, Spain

³CNAG-CRG, Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology, Barcelona, Spain

Correspondence

Miriam Guitart and Anna Ruiz, Genetics Laboratory, UDIAT-Centre Diagnòstic, Parc Taulí Hospital Universitari, Institut d'Investigació i Innovació Parc Taulí I3PT, Universitat Autònoma de Barcelona, Sabadell, Spain.
Emails: mguitart@tauli.cat or aruizn@tauli.cat

Funding information

Institut d'Investigació i Innovació Parc Taulí I3PT, Grant/Award Number: CIR2016/025; Asociación Española de Síndrome de Angelman; Instituto de Salud Carlos III, Grant/Award Number: PI16/01411

Abstract

Background: Patients affected by Angelman syndrome (AS) present severe intellectual disability, lack of speech, ataxia, seizures, abnormal electroencephalography (EEG), and a characteristic behavioral phenotype. Around 10% of patients with a clinical diagnosis of AS (AS-like) do not have an identifiable molecular defect. Some of these patients harbor alternative genetic defects that present overlapping features with AS.

Methods: Trio whole-exome sequence was performed on patient and parent's DNA extracted from peripheral blood. Exome data were filtered according to a *de novo* autosomal dominant inheritance. cDNA analysis was carried out to assess the effect of the splice site variant.

Results: We identified a novel heterozygous *SMARCE1* splicing variant that leads to an exon skipping in a patient with an Angelman-like phenotype. Missense variants in the *SMARCE1* gene are known to cause Coffin–Siris syndrome (CSS), which is a rare congenital syndrome. Clinical reevaluation of the patient confirmed the presence of characteristic clinical features of CSS, many of them overlapping with AS.

Conclusions: Taking into account the novel finding reported in this study, we consider that CSS should be added to the expanding list of differential diagnoses for AS.

KEYWORDS

Angelman syndrome (AS), Coffin–Siris syndrome (CSS), exome sequencing, *SMARCE1*

1 | INTRODUCTION

Angelman syndrome (AS) is a neurodevelopmental disorder characterized by severe intellectual disability, lack of speech, ataxia of gait, seizures, a characteristic electroencephalography (EEG), and a unique behavior that includes

any combination of happy demeanor, easily excitable personality, frequent laughter, and stereotypes (Bird, 2014; Buiting, 2010). AS is caused by the lack of expression of the maternally inherited *UBE3A* gene (OMIM 601623) in neurons. Around 10% of patients with a clinical diagnosis of AS are not molecularly confirmed (Dagli, Buiting, & Williams,

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2018 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals, Inc.

Mol Genet Genomic Med. 2019;7:e511.
<https://doi.org/10.1002/mgg3.511>

wileyonlinelibrary.com/journal/mgg3 | 1 of 6

2012). Some of these Angelman-like syndrome patients harbor alternative genetic defects that present overlapping clinical features with AS (Tan, Bird, Thibert, & Williams, 2014).

Genomic approaches such as array comparative genomic hybridization and whole-exome sequencing have already been useful to identify alternative genes responsible for other heterogeneous genetic diseases such as Rett, Kleeftstra, and Smith–Magenis syndromes (Berger et al., 2017; Kleeftstra et al., 2012; Sajan et al., 2016). Here, we identified with exome sequencing a novel heterozygous *SMARCE1* (OMIM 603111) splicing variant in a patient with an Angelman-like phenotype.

2 | MATERIALS AND METHODS

2.1 | Ethical compliance

The protocol for the study has been approved by the institutional Ethics Committee of Institut d'Investigació i Innovació Parc Taulí I3PT and the corresponding informed consent has been obtained from the parents.

2.2 | Patient

The proband is a 14-year-old boy who was born at term to non-consanguineous parents following a normal pregnancy. The

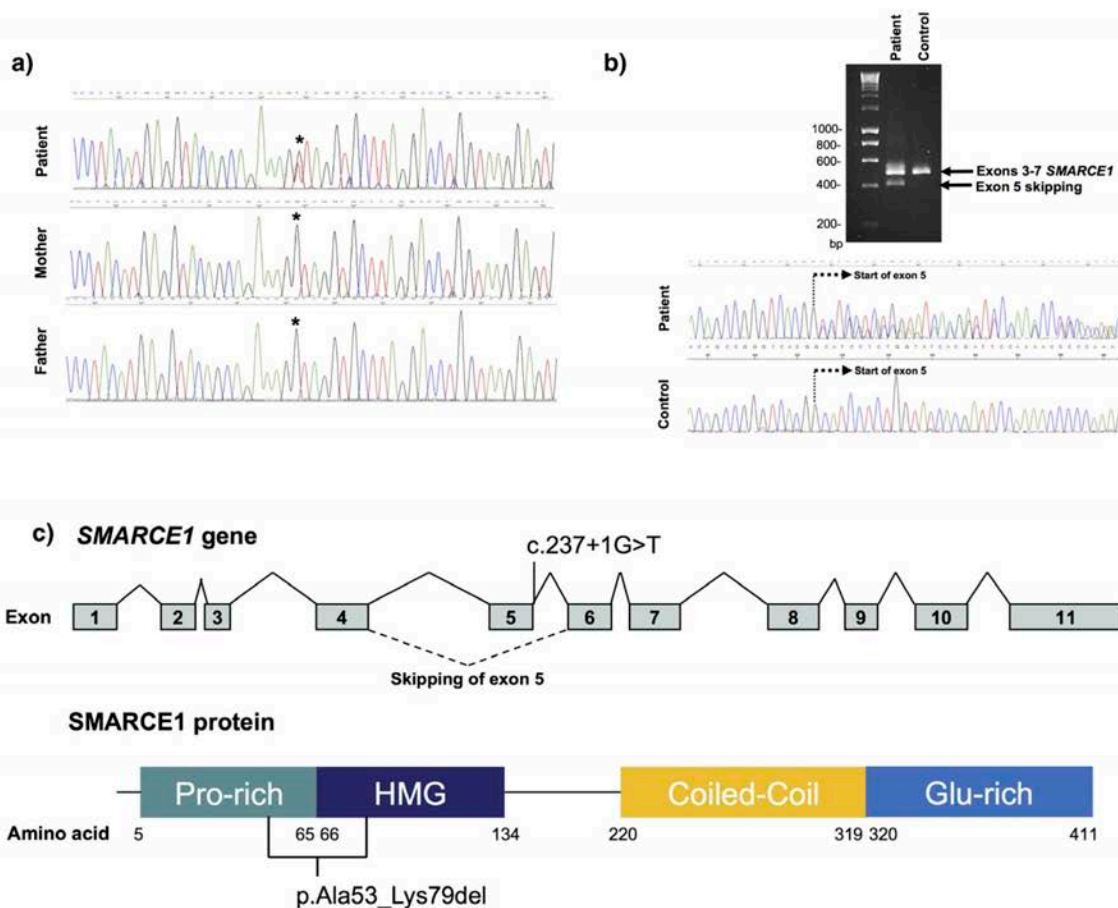


FIGURE 1 Molecular characterization of the *SMARCE1* (NCBI RefSeq NM_003079.4) c.237+1G>T splicing variant. (a) Sanger sequencing chromatographs showing the *SMARCE1* c.237+1G>T variant in the patient but not in his unaffected parents. The variant is indicated by black asterisks. (b) cDNA analysis of the *SMARCE1* c.237+1G>T variant in peripheral blood. PCR amplification products of exons 3 to 7 were run in a gel electrophoresis. The patient showed two bands compared to the negative control. Sanger sequencing chromatographs indicate skipping of exon 5. The start of the exon is indicated by a dashed line. (c) Schematic structure of the human *SMARCE1* gene and protein illustrating the predicted variant effect on splicing and protein. Exons are shown as boxes and introns as lines. Protein functional domains are shown as boxes. The amino acid deletion is delimited by black lines. Proline-rich domain (Pro-rich) 5–65 aa, High Mobility Group (HMG) 66–134 aa, Coiled-Coil domain 220–319 aa, glutamic acid-rich domain (Glu-rich) 320–411 aa

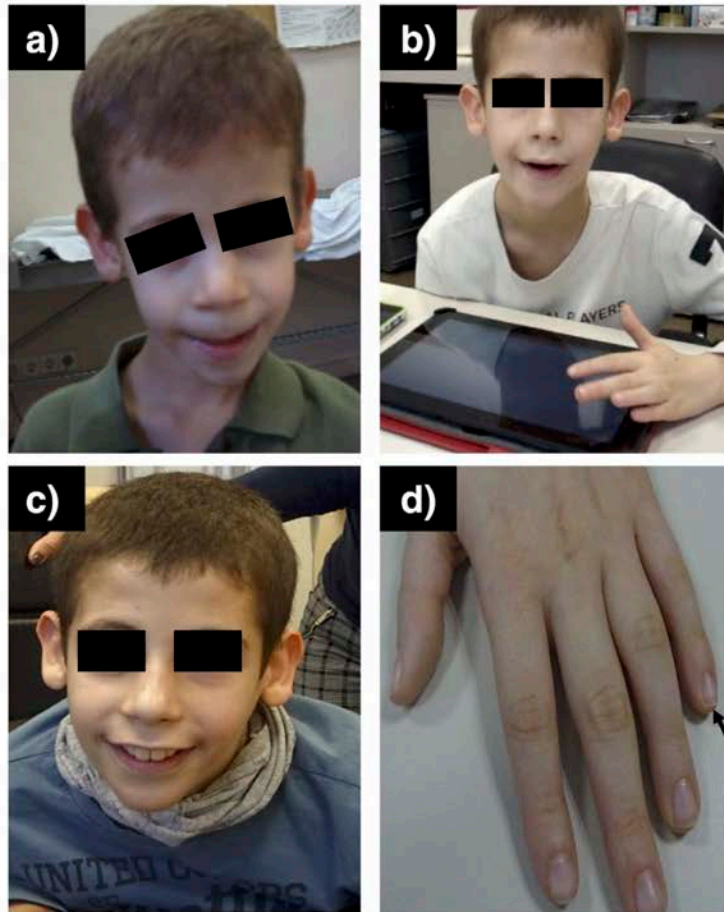


FIGURE 2 Patient clinical features. (a) Patient at 5 years of age, (b) 9 years of age, (c) 14 years of age, (d) Left hand showing small and hypoplastic 5th fingernail

dominant *de novo* model of inheritance, a population allele frequency of $<1/1,000$ and a predicted impact on the protein. Variants in the *SMARCE1* gene are known to cause Coffin–Siris syndrome (CSS5; Coffin–Siris syndrome 5, OMIM 616938), which is a rare congenital syndrome affecting many organs, characterized by moderate to severe intellectual disability (Kosho & Okamoto, 2014; Santen, Emmelien, Vulto-van Silfhout, Pottinger, & Van Bon, 2013).

The presence of the variant was confirmed by Sanger sequencing in the index patient, whereas the variant was not detected in his unaffected parents (Figure 1a). In order to analyze the effect of the splice site variant on mRNA processing, cDNA analysis was performed on the patient and a control sample. Amplification of exons 3 to 7 resulted in an additional smaller fragment in the patient suggesting exon skipping. Sanger sequencing of the cDNA confirmed the skipping of exon 5 in the patient sample (Figure 1b). Deletion of exon 5 results in an in-frame deletion of 27 amino acids, removing the last part of the Pro-rich (Proline-rich) domain and the

start of the HMG (High Mobility Group) domain (Figure 1c), which is essential for the proper functioning of the protein (Lomelí & Castillo-Robles, 2016).

4 | DISCUSSION

Using trio whole-exome sequencing, we identified a novel splice site variant resulting in an in-frame deletion in the *SMARCE1* gene in a patient with an AS-like phenotype. Pathogenic variants in *SMARCE1* and another five genes (*SMARCB1*, *SMARCA4*, *SMARCA2*, *ARID1A*, and *ARID1B*) encoding subunits of the switch/sucrose non-fermenting (SWI/SNF) ATP remodeling complex cause CSS (Miyake, Tsurusaki, & Matsumoto, 2014) which is a rare congenital syndrome characterized by developmental delay, moderate to severe intellectual disability, hypoplastic or absent fifth fingernails or toenails, distinctive facial features, hypertrichosis, sparse scalp hair, and hypotonia (Kosho & Okamoto,

2014; Santen et al., 2013; Tsurusaki et al., 2014; Zarate et al., 2016).

Germline *SMARCE1* heterozygous loss-of-function variants have been found in young patients with cranial and spinal meningiomas, consistent with a tumor suppressor mechanism (Lomeli & Castillo-Robles, 2016; Smith et al., 2014) while missense variants cause Coffin–Siris syndrome (Kosho & Okamoto, 2014).

Here, we describe a splicing variant (c.237+1G>T) in the *SMARCE1* gene leading to an in-frame deletion of 27 amino acids, removing part of the HMG domain (Figure 1c). The patient presents CSS clinical features. No spinal or intracranial meningiomas, which are characteristic of pathogenic loss-of-function variants in *SMARCE1*, were detected in a recent MRI, which showed a dysgenesis and hypoplasia of the corpus callosum and a global dilatation of the ventricular system, characteristic of Coffin–Siris patients. Interestingly, Smith et al., (2013) reported a splicing variant similar to the variant found in our patient, c.237+2T>C. It was identified in two members of a family with multiple spinal meningiomas with no clinical symptoms of CSS. RNA analysis of the affected individuals presented two alternatively spliced *SMARCE1* transcripts, one leading to a premature STOP codon and a less abundant second transcript leading to the same in-frame deletion as that in our patient;. In our patient, the c.237+1G>T variant leads only to one alternative splice variant, the 27 amino acid in-frame deletion, suggesting that the abundance of the in-frame deletion transcript leads to a gain of function or dominant negative effect (Tsurusaki et al., 2012) and the appearance of a CSS phenotype.

The patient described here shows almost all the consistent and frequent clinical features associated to AS (Williams, Driscoll, & Dagli, 2010) except for the microcephaly and the frequent laughter/smiling (Table 1). Also, other associated AS features such as attraction to water, a happy demeanor, or frequent drooling are present. This led to the clinical diagnosis of AS which was not confirmed molecularly.

Clinical reevaluation of the patient after the identification of the *SMARCE1* pathogenic variant showed the presence of clinical features associated to CSS but not to AS (Table 1). Among them, the characteristic hypoplastic nail on the 5th finger of the left hand, a coarse facies, sparse scalp hair, hypertrichosis in the back and dysgenesis and hypoplasia of the corpus callosum (Figure 2). CSS patients, like AS patients, present severe developmental delay, speech impairment with expressive language more severely affected than receptive language, moderate to severe intellectual disability and behavioral abnormalities such as hyperactivity and autistic features (Table 1) reflecting the clinical overlap between the two syndromes. However, the patient described here presents some other features characteristic of AS which have not been described in CSS before such as the ataxia of gait and the stereotypies.

To date, only six individuals with *SMARCE1* missense pathogenic variants have been reported (Zarate et al., 2016). Here, we describe the first patient with a pathogenic splicing variant in the CSS gene, *SMARCE1*, who had a diagnosis of AS-like and who presents some clinical features characteristic of AS, which have not been previously associated to CSS. Taking into account these results, we believe that CSS should be added to the expanding list of differential diagnoses for AS, probably accounting for some of the molecularly undiagnosed AS-like patients.

The increasing use of exome sequencing in diagnostic laboratories will allow the analysis of all those genes that are involved in severe neurodevelopment disorders in patients who present AS clinical features, improving the diagnostic rate, and providing knowledge of the phenotypic spectrum of AS-like causative genes, among them those responsible for Coffin–Siris syndrome.

ACKNOWLEDGMENTS

We thank the patient and his family for their participation in this study. We also thank Instituto de Salud Carlos III (PI16/01411), Institut d'Investigació i Innovació Parc Taulí I3PT (CIR2016/025) and Asociación Española de Síndrome de Angelman for their financial support.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ORCID

Cynthia Aguilera  <http://orcid.org/0000-0002-0363-8590>

REFERENCES

- Berger, S. I., Ciccone, C., Simon, K. L., Malicdan, M. C., Vilboux, T., Billington, C., ... Smith, A. C. M. (2017). Exome analysis of Smith–Magenis-like syndrome cohort identifies de novo likely pathogenic variants. *Human Genetics*, *136*, 409–420. <https://doi.org/10.1007/s00439-017-1767-x>
- Bird, L. (2014). Angelman syndrome: review of clinical and molecular aspects. *The Application of Clinical Genetics*, *7*, 93–104. <https://doi.org/10.2147/TACG>
- Buiting, K. (2010). Prader-Willi syndrome and Angelman syndrome. *American Journal of Medical Genetics Part C*, *154C*, 365–376. <https://doi.org/10.1002/ajmg.c.30273>
- Dagli, A., Buiting, K., & Williams, C. A. (2012). Molecular and clinical aspects of Angelman syndrome. *Molecular Syndromology*, *2*, 100–112.
- DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., ... Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, *43*, 491–498. <https://doi.org/10.1038/ng.806>

- Kleefstra, T., Kramer Jamie, M., Neveling, K., Willemsen Marjolein, H., Koemans Tom, S., Vissers Lisenka, E. L. M., ... van Bokhoven, H. (2012). Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *American Journal of Human Genetics*, *91*, 73–82. <https://doi.org/10.1016/j.ajhg.2012.05.003>
- Kosho, T., & Okamoto, N. (2014). Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. *American Journal of Medical Genetics Part C*, *166*, 262–275. <https://doi.org/10.1002/ajmg.c.31407>
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., ... MacArthur, D. G. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*, 285–291. <https://doi.org/10.1038/nature19057>
- Lomelí, H., & Castillo-Robles, J. (2016). The developmental and pathogenic roles of BAF57, a special subunit of the BAF chromatin-remodeling complex. *FEBS Letters*, *590*, 1555–1569. <https://doi.org/10.1002/1873-3468.12201>
- Miyake, N., Tsurusaki, Y., & Matsumoto, N. (2014). Numerous BAF complex genes are mutated in Coffin-Siris syndrome. *American Journal of Medical Genetics Part C*, *166*, 257–261. <https://doi.org/10.1002/ajmg.c.31406>
- Sajan, S. A., Jhangiani, S. N., Muzny, D. M., Gibbs, R. A., Lupski, J. R., Glaze, D. G., ... Neul, J. L. (2016). Enrichment of mutations in chromatin regulators in people with Rett syndrome lacking mutations in MECP2. *Genetics in Medicine*, *19*, 13–19.
- Santen, G. W. E., Emmelien, A., Vulto-van Silfhout, A. T., Pottinger, C., & Van Bon, B. W. M. (2013). Coffin-Siris syndrome and the BAF complex: Genotype-phenotype study in 63 patients. *Human Mutation*, *34*, 1519–1528. <https://doi.org/10.1002/humu.22394>
- Schrier Vergano, S., Santen, G., Wieczorek, D., Wollnik, B., & Matsumoto, N. (2018). Coffin-siris syndrome. In M. P. Adam, H. H. Ardinger, R. A. Pagon, S. E. Wallace, L. J. H. Bean, K. Stephens, & A. Amemiya (Eds.), *GeneReviews*® [Internet]. (pp. 1993-2018). Seattle, WA: University of Washington.
- Smith, M. J., O'Sullivan, J., Bhaskar, S. S., Hadfield, K. D., Poke, G., Caird, J., ... Evans, D. G. (2013). Loss-of-function mutations in SMARCE1 cause an inherited disorder of multiple spinal meningiomas. *Nature Genetics*, *45*, 295–298. <https://doi.org/10.1038/ng.2552>
- Smith, M. J., Wallace, A. J., Bennett, C., Hasselblatt, M., Elert-Dobkowska, E., Evans, L. T., ... Evans, D. G. (2014). Germline SMARCE1 mutations predispose to both spinal and cranial clear cell meningiomas. *The Journal of Pathology*, *234*, 436–440. <https://doi.org/10.1002/path.4427>
- Tan, W.-H., Bird, L. M., Thibert, R. L., & Williams, C. A. (2014). If not Angelman, what is it? A review of Angelman-like syndromes. *American Journal of Medical Genetics. Part A*, *164*, 975–992. <https://doi.org/10.1002/ajmg.a.36416>
- Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y., ... Matsumoto, N. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nature Genetics*, *44*, 376. <https://doi.org/10.1038/ng.2219>
- Tsurusaki, Y., Okamoto, N., Ohashi, H., Mizuno, S., Matsumoto, N., Makita, Y., ... Miyake, N. (2014). Coffin-Siris syndrome is a SWI/SNF complex disorder. *Clinical Genetics*, *85*, 548–554. <https://doi.org/10.1111/cge.12225>
- Williams, C. A., Driscoll, D. J., & Dagli, A. I. (2010). Clinical and genetic aspects of Angelman syndrome. *Genetics in Medicine*, *12*, 385–395.
- Zarate, Y. A., Bhoj, E., Kaylor, J., Li, D., Tsurusaki, Y., Miyake, N., ... Schrier Vergano, S. A. (2016). SMARCE1, a rare cause of Coffin-Siris syndrome: Clinical description of three additional cases. *American Journal of Medical Genetics. Part A*, *170*, 1967–1973. <https://doi.org/10.1002/ajmg.a.37722>

How to cite this article: Aguilera C, Gabau E, Laurie S, et al. Identification of a de novo splicing variant in the Coffin-Siris gene, *SMARCE1*, in a patient with Angelman-like syndrome. *Mol Genet Genomic Med*. 2019;7:e511. <https://doi.org/10.1002/mgg3.511>

Capítol 2. Anàlisi d'altres models d'herència mendelians

En els pacients en els quals no s'han identificat variants patogèniques/probablement patogèniques en l'anàlisi del model d'herència *de novo* (pacients 11, 12, 13, 14, 15 i 17), s'ha portat a terme l'anàlisi dels models d'herència lligat al cromosoma X i autosòmic recessiu.

2.1 Herència lligada al cromosoma X

L'anàlisi del model d'herència lligat al cromosoma X ha identificat un total de 103-111 variants per individu. Posteriorment, aquestes variants s'han filtrat per una freqüència al·lèlica en ExAC i gnomAD $<1/1000$ i l'impacte sobre la proteïna (*nonsense*, *frameshift*, *splicing* i *missense*), obtenint-se entre 0-4 variants per individu (Taula 11). S'ha realitzat una cerca en les bases de dades d'OMIM, SysID, DDG2P i Pubmed dels gens identificats, per analitzar si intervenen en el desenvolupament del sistema nerviós o estan associats a TND. Finalment, s'han considerat rellevants dues variants localitzades en els gens *LAS1L* i *COX7B* en els pacients 11 i 12, respectivament.

Taula 11 | Variants resultants de l'anàlisi del model d'herència lligat al cromosoma X.

	Gen	Isoforma	Tipus de variant	Canvi de nucleòtid	Canvi d'aminoàcid	ExAC/ gnomAD	Classificació (guia ACMG)
Pacient 11	<i>LAS1L</i>	NM_031206.4	<i>Missense</i>	c.1237G>A	p.Gly413Arg	- / -	Probablement patogènica
	<i>MAMLD1</i>	NM_001177465.3	<i>Missense</i>	c.1597C>A	p.Pro533Thr	- / -	Probablement benigna
Pacient 12	<i>COX7B</i>	NM_001866.3	<i>Missense</i>	c.220C>A	p.Pro74Thr	- / -	Significat incert
Pacient 15	<i>MAGEA10</i>	NM_021048.5	<i>In-frame</i>	c.1097_1099delCCT	p.Ser366del	- / -	Significat incert
Pacient 17	<i>DMD</i>	NM_004006.2	<i>Missense</i>	c.821A>G	p.Tyr274Cys	0,0001/ 0,0001	Benigna
	<i>FAM47B</i>	NM_152631.3	<i>Missense</i>	c.346G>A	p.Ala116Thr	-/ 0,005617	Significat incert
	<i>SRPX</i>	NM_006307.5	<i>Missense</i>	c.593C>T	p.Thr198Ile	0,0004 / 0,0004	Significat incert
	<i>SLC25A5</i>	NM_001152.5	<i>Missense</i>	c.662C>G	p.Thr221Ser	0.0007 / -	Significat incert

En el pacient 11 s'ha identificat la variant *missense* c.1237G>A, p.Gly413Arg en el gen *LAS1L*. La mare és portadora de la variant en mosaic (20%, segons les dades de WES) (Figura 18). Aquesta variant no ha estat descrita prèviament i només s'han identificat dues famílies portadores de variants en aquest gen, associades a la síndrome de Wilson-Turner (WTS, MIM # 309585) (238).

El gen *LAS1L* codifica per una subunitat d'un complex proteic que es troba al nuclèol que es requereix per a la síntesi de la subunitat 60S del ribosoma i la maduració de l'ARN ribosòmic 28S. *LAS1L* també està implicat en la regulació de la transcripció i forma part del complex repressor coREST/HDAC1 (239–241). En la re-avaluació clínica del pacient, s'ha observat que a part de presentar les característiques clíniques de la SA, també mostra estatura baixa i obesitat troncal, que són característiques pròpies de la síndrome de Wilson-Turner. Tot i això, el pacient no presenta criptorquídia i l'hipogonadisme no s'ha pogut avaluar perquè el pacient és prepuberal. L'eina Varsome suggereix que es tracta d'una variant de significat incert que compleix els següents criteris de patogenicitat d'acord amb la guia ACMG: i) no es troba en bases de dades de població control (ExAC i gnomAD) i ii) segons els predictors *in silico* té un efecte deleteri sobre la proteïna.

Tenint en compte que el pacient 11 presenta característiques clíniques específiques de la WTS i els criteris de patogenicitat de la guia ACMG anteriorment esmentats, aquesta variant s'ha classificat com probablement patogènica, considerant-se la responsable del fenotip SA-like present en el pacient.

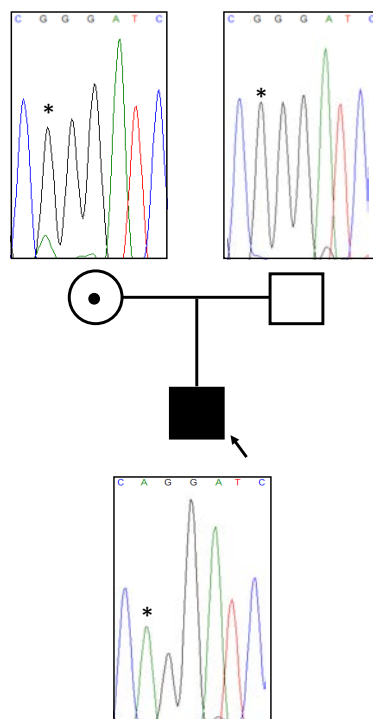


Figura 18| Pedigrí del trio 11 on es mostra els resultats de la validació per seqüenciació Sanger de la variant c.1237G>A en el gen *LAS1L*.

En la pacient 12 s'ha identificat la variant c.220C>A, p.Pro74Thr en el gen *COX7B*, no descrita prèviament. La mare és portadora en mosaic (26%, segons les dades del WES) (Figura 19). El gen *COX7B* codifica per la subunitat 7B de la citocrom c oxidasa mitocondrial i s'associa a la síndrome de microftàlmia amb defectes linears de la pell (MIM # 300887) lligada al cromosoma X dominant. És letal en els nens i en les nenes afectades hi ha molta variabilitat fenotípica. El fenotip clàssic consisteix en microftàlmia i/o anoftàlmia amb lesions linears a la pell, però també hi ha característiques clíniques que es troben en menys del 70% dels individus afectats entre les quals es troben la DI, retard del desenvolupament, microcefàlia i epilèpsia infantil (242). S'ha observat que hi ha una gran variabilitat fenotípica a causa de la inactivació preferencial del cromosoma X, algunes pacients no presenten microftàlmia, però sí lesions en la pell i a l'inrevés (243). L'eina Varsome suggereix que la variant c.220C>A és de significat incert. Els criteris de patogenicitat que compleix són els següents: i) els programes de predicció *in silico* la classifiquen com una variant deletèria i ii) no es troba en les bases de dades d'ExAC i gnomAD. Considerant que la pacient 12 no presenta els principals criteris clínics de la síndrome de microftàlmia amb defectes linears de la pell, s'ha classificat com una variant de significat incert.

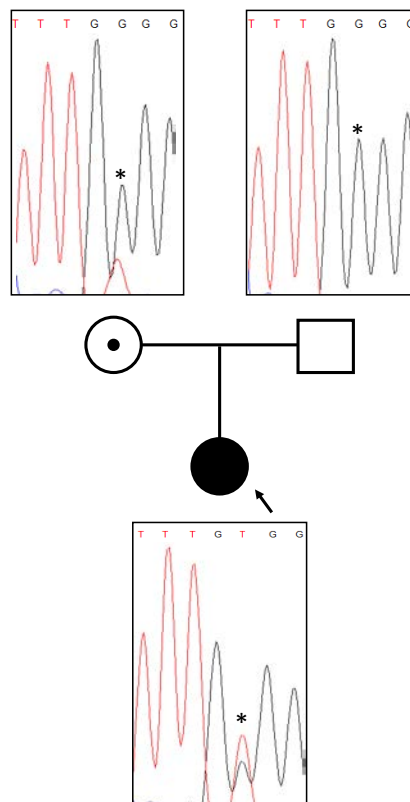


Figura 19] Pedigrí del trio 12 on es mostra els resultats de la validació per seqüenciació Sanger de la variant c.220C>A en el gen *COX7B*.

2.2 Herència autosòmica recessiva

L'anàlisi del model d'herència autosòmic recessiu, s'ha realitzat considerant només els heterozigots compostos, ja que no hi havia consanguinitat en els trios de la nostra cohort. Les variants s'han filtrat per una freqüència al·lèlica en ExAC i gnomAD $<1/200$ i per variants amb un impacte sobre la proteïna (*nonsense*, *frameshift*, *splicing* i *missense*). Com a resultat, s'han obtingut entre 349-522 variants per individu. S'ha analitzat si les variants es troben en gens implicats en el desenvolupament del sistema nerviós o en TND cercant en les bases de dades d'OMIM, SysID, DD2GP i Pubmed. Aquest anàlisi ha permès la identificació de dues variants *missense* en el gen *RMND1* (c.1286A>T, p.Glu429Val i c.1049T>C, p.Met350Thr) en la pacient 14, les quals no han estat descrites prèviament (Taula 12).

Taula 12 | Variants rellevants resultants de l'anàlisi del model d'herència autosòmic recessiu.

	Gen	Isoforma	Tipus de variant	Canvi de nucleòtid	Canvi d'aminoàcid	ExAC/ gnomAD	Herència	Classificació (guia ACMG)
Pacient 14	<i>RMND1</i>	NM_017909.4	<i>Missense</i>	c.1286A>T	p.Glu429Val	- / -	Heterozigot compost	Variante de significat incert
	<i>RMND1</i>	NM_017909.4	<i>Missense</i>	c.1049T>C	p.Met350Thr	- / -	Heterozigot compost	Variante de significat incert

Variants patogèniques en el gen *RMND1* s'associen a un trastorn mitocondrial amb herència autosòmica recessiva que es caracteritza per hipotonia neonatal, acidosi làctica, encefalopatia, sordesa neurosensorial i insuficiència renal (MIM # 614922). L'eina Varsome suggereix que les dues variants són de significat incert tenint en compte les següents evidències: i) no es troben en les bases de dades de població control ExAC i gnomAD, ii) els programes *in silico* prediuen un efecte deleteri sobre la proteïna i iii) variants truncants en el gen *RMND1* són el principal mecanisme de patogenicitat, mentre que més del 50% de les variants *missense* reportades són benignes. Considerant que la pacient 14 no presenta les principals característiques clíniques de la malaltia mitocondrial associada a variants patogèniques/probablement patogèniques en el gen *RMND1*, s'han classificat com variants de significat incert.

Capítol 3. Avaluació de la càrrega genètica *de novo* en els pacients amb un fenotip SA-like

S'ha reportat que la càrrega de variants genètiques d'un individu defineix el fenotip final que aquest presenta. En la síndrome de Rett i en la neuropatia Charcot-Marie-tooth s'ha descrit que hi ha un increment de variants *de novo* i a que a més, aquestes poden modular el fenotip (244,245). En el cas de l'autisme s'ha descrit que diferents variants *de novo* són les responsables del fenotip, indicant un model oligogènic per l'autisme (246). Per estudiar la càrrega de variants *de novo* en els pacients de la nostra cohort, s'han analitzat totes les variants que potencialment alteraven la proteïna, obtingudes de l'anàlisi del model d'herència *de novo*, que ha donat com a resultat entre 0-5 variants per individu. En la Taula 13 es mostren totes les variants *de novo* que potencialment alteren la proteïna. D'aquestes, d'una banda, s'han analitzat aquelles que es troben localitzades en gens associats a TND consultant les bases de dades DDG2P, SysID, OMIM i Pubmed i d'altra banda, les variants localitzades en gens que estan associats a altres malalties que no són TND. La classificació de les variants s'ha portat a terme amb el suport de l'eina Varsome.

Taula 13 | Variants *de novo* que potencialment alteren la proteïna en els pacients amb un fenotip SA-like.

	Posició cromosòmica	Tipus de variant	Gen	Isoforma	Canvi nucleòtid	Canvi d'aminoàcid	ExAC/GnomAD
Pacient 1	2:95847733	Missense	ZNF2	NM_021088.4	c.1160G>A	p.Arg387Gln	0,00002491/ 0,00001627
	2:241725854	Missense	KIF1A	NM_001244008.1	c.506G>C	p.Arg169Thr	- / -
Pacient 2	2:170557484	Frameshift	PHOSPHO2	NM_001199286.1	c.7delA	p.Ile3Phefs*3	- / -
	17:8065077	In-frame	VAMP2	NM_014232.3	c.128_130delTGG	p.Val43del	- / -
Pacient 3	1:145585510	Nonsense	PIAS3	NM_006099.3	c.1775G>A	p.Arg592Hys	0,00008635 / 0,0001
	6:33408690	Missense	SYNGAP1	NM_006772.2	c.1861C>T	p.Arg621*	- / -
	17:18221151	Nonsense	SMCR8	NM_144775.3	c.2048C>T	p.Ser683Phe	- / -
Pacient 4	2:220355247	Missense	SPEG	NM_005876.5	c.9038T>G	p.Val3013Gly	- / -
	3:176756148	Missense	TBL1XR1	NM_024665.5	c.1000T>C	p.Cys334Arg	- / -
	12:56384478	Nonsense	RAB5B	NM_002868.4	c.328C>T	p.Arg110*	- / -
	17:55184422	Missense	AKAP1	NM_003488.4	c.1597A>C	p.Thr533Pro	0,0621 / 0,043
Pacient 6	1:22186113	Frameshift	HSPG2	NM_001291860.1	c.5242A>C	p.Thr1748Pro	0,237 / 0,0813
	2:200137309	Frameshift	SATB2	NM_001172509.2	c.1826delA	p.Asp609Alafs*15	- / -
	7:925694	Missense	GET4	NM_015949.3	c.157T>G	p.Tyr53Asp	0,1032 / 0,0501
	8:105360983	Missense	DCSTAMP	NM_030788.4	c.203C>T	p.Thr68Met	- / -
Pacient 7	9:71866104	Missense	TJP2	NM_001170416.2	c.3238A>G	p.Ile1080Val	0,000008237/ 0,00001218
	6:76751645	Frameshift	IMPG1	NM_001563.4	c.265delA	p.Ile89Phefs*2	- / -

4. RESULTATS

Taula 13 | Continuació.

Pacient 8	X:153053039	Missense	IDH3G	NM_004135.4	c.554T>G	p.Val185Gly	0,0001/ 0,00005056
	8:55540933	Missense	RP1	NM_006269.1	c.4491G>C	p.Glu1497Asp	- / -
	8:133192493	Missense	KCNQ3	NM_004519.4	c.688C>T	p.Arg230Cys	- / -
	15:42439924	Missense	PLA2G4F	NM_213600.3	c.1096A>C	p.Thr366Pro	0,0006 / 0,0002
	17:38793743	Splicing	SMARCE1	NM_003079.5	c.237+1G>T	p.Ala53_Lys79del	- / -
Pacient 9	22:21081576	Missense	PI4KA	NM_058004.3	c.4883A>C	p.Tyr1628Ser	0,0005 / 0,0004
	22:38219579	Missense	GALR3	NM_003614.1	c.166A>C	p.Thr56Pro	- / -
	9:131389677	In-frame	SPTAN1	NM_001130438.2	c.6589_6590insAG CTGC	p.Leu2198_Gln219 9dup	- / -
Pacient 10	14:14752304	Nonsense	YLPM1	NM_019589.3	c.222C>A	p.Cys74*	- / -
	18:31322918	Nonsense	ASXL3	NM_030632.3	c.3106C>T	p.Arg1036*	- / -
Pacient 11	22:38610836	Indel	MAFF	NM_001161572.1	c.464_481dupCCCC GCCACGCGCCCGGA	p.Asp155_Pro160 dup	0,0001/ 0,00008688
	22:50721182	Missense	PLXNB2	NM_012401.4	c.2945A>C	p.Asn982Thr	0,0222 / 0,0115
	20:42328505	Missense	MYBL2	NM_002466.4	c.772G>A	p.Gly258Ser	0,00002484/ 0,00002033
	19:49636406	Splicing	PPFIA3	NM_003660.4	c.1026+2T>G	-	0,35 / 0,1672
Pacient 12	1:52255263	Missense	NRDC	NM_002525.2	c.3439G>A	p.Gly1147Arg	- / 0,000008134
	1:153657484	Missense	NPR1	NM_000906.4	c.1529T>G	p.Val510Gly	0,0019 / 0,0011
	2:196720641	Missense	DNAH7	NM_018897.3	c.8489G>A	p.Gly2830Asp	- / -
	8:25159972	Splicing	DOCK5	NM_024940.8	c.976+2T>G	-	0,0005 / 0,0011
Pacient 13	22:26709766	Missense	SEZ6L	NM_021115.5	c.1913T>G	p.Val638Gly	0,0239 / 0,0142
	1:205631176	Missense	SLC45A3	NM_033102.3	c.1037T>G	p.Val346Gly	0,0991 / 0,0562
	3:197273305	Missense	BDH1	NM_004051.4	c.10A>C	p.Thr4Pro	0,00004142/ 0,00001224
Pacient 14	6:110679444	Missense	METTL24	NM_001123364.2	c.32G>T	p.Cys11Phe	- / -
	6:122737363	Frameshift	HSF2	NM_004506.3	c.456_459delTGAG	p.Ser152Argfs*40	- / -
	9:96715295	Missense	BARX1	NM_021570.4	c.398G>C	p.Gly133Ala	- / -
	X:100075437	Missense	CSTF2	NM_001306206.1	c.32T>G	p.Val11Gly	- / -
	19:49132198	Missense	SPHK2	NM_020126.4	c.1133A>C	p.Tyr378Ser	0,0017 / 0,001
Pacient 16	19:51014372	Splicing	JOSD2	NM_001270639.1	c.-18+2T>G	-	0,0202 / 0,0083
	3:11067498	Missense	SLC6A1	NM_003042.4	c.889G>A	p.Gly297Arg	- / -
	12:108986032	Missense	TMEM119	NM_181724.3	c.128A>G	p.Glu43Gly	0,04469 / 0,0002

Pel que fa a les variants localitzades en gens implicats en el neurodesenvolupament, cal destacar les variants en els gens *SPEG*, *HSPG2*, *PI4KA* i *CSTF2* (pacients 4, 6, 8 i 14). El gen *SPEG* s'associa a una miopatia centronuclear (MIM # 615959), el gen *HSPG2* està associat a una displàsia disegmentària tipus Silverman Handmaker (MIM # 224410) i Schwartz-Jampel tipus 1 (MIM #

255800); el gen *PI4KA* s'associa a polimicrogiria, síndrome perisilvià, hipoplàsia del cerebel i artrogriposi (MIM # 616531) i finalment, el gen *CSTF2* s'ha associat recentment a DI lleu no sindròmica en homes (247). Els predictors *in silico* classifiquen aquestes variants com a deletèries, les variants en *HSPG2* i *PI4KA* es troben en les bases de dades ExAC i gnomAD, mentre que les variants en el gen *SPEG* i *CSTF2* no s'han descrit prèviament. En els casos de *SPEG*, *HSPG2* i *PI4KA*, el model d'herència és autosòmic recessiu, per tant, s'han classificat com a variants de significat incert, ja que es necessitarien dues variants patogèniques per donar el fenotip.

En el cas del gen *CSTF2*, codifica per un factor que està implicat en el processament del pre-ARNm, en concret, en la decapació de l'extrem 3' i poliadenilació. Regula l'expressió gènica de diferents teixits, en especial, en cervell i testicles (248–250). És un gen intolerant als canvis de pèrdua de funció (pLI 1) i té un valor significatiu de z-score de 3,49, observant-se menys variants *missense* de les que s'esperaria en aquest gen. S'expressa principalment en el tracte digestiu proximal, en la medulla òssia i teixits limfoides, però també s'expressa a alts nivells en cervell (237). Estudis en ratolins *knockout* (gen paràleg *Cstf2t*^{-/-}) han demostrat que aquest gen intervé en l'aprenentatge i la memòria, sobretot en les femelles, i en el cas dels mascles s'ha observat que la disrupció del gen afecta a l'espermatogènesi (251).

Recentment, s'ha identificat la variant *missense* p.Asp50Ala en quatre homes d'una família, el cas índex presenta retard del desenvolupament i els altres individus afectats presenten DI lleu i afectació del llenguatge, les dones portadores eren sanes. La variant D50A es troba localitzada en el motiu de reconeixement de l'ARN (RRM) de la proteïna CSTF2. Estudis funcionals han demostrat que la variant D50A afecta a l'eficiència de la decapació i poliadenilació del pre-ARNm i augmenta l'afinitat del domini RRM per l'ARN. A més, estudis en ratolins ^{D50A/Y} s'observen canvis en la decapació i poliadenilació de 1370 gens en el cervell (247).

La variant p.Val11Gly identificada en la pacient 14 de la nostra cohort, es troba localitzada abans de l'inici del domini RRM, per tant, es desconeix l'impacte que pot tenir sobre la funció de la proteïna (Figura 20). Hi ha discrepàncies entre els predictors bioinformàtics, alguns la consideren una variant deletèria, mentre que d'altres la consideren una variant neutra (Provean -4,08 (*Damaging*), MutPred 0,43, Mutation Assessor 0,16 (*Neutral*), Polyphen2 0,18 (*Benign*), SNP&GO *Neutral* RI:6 i PredictSNP 63% (*Neutral*)). Per tant, tenint en compte les raons esmentades anteriorment, que és un gen poc conegut i que només s'ha descrit una variant en una família associada a DI lleu no sindròmica lligada al cromosoma X, s'ha classificat com a variant de significat incert.

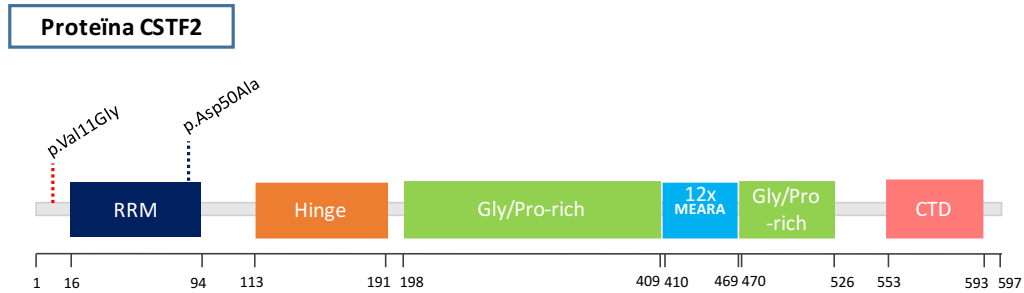


Figura 20 | Estructura de la proteïna CSTF2 on es troben indicades la variant identificada en la pacient 14 de la nostra cohort i la variant identificada per Grozdanov et al., 2020. RRM, RNA *recognition motif*; Hinge, *CSTF-77 interaction domain*; Gly/Pro-rich, *Glycine/Proline rich domain*; 12x MEARA, *repetition of amino acids MEARA*; CTD, *C-terminal domain*.

D'altra banda, s'han identificat variants en gens associats a altres malalties que no són TND. Entre elles, cal destacar la variant c.265delA, p.Ile89Phefs*2 en el gen *IMPG1* que és el component majoritari de la matriu d'interfotoreceptors de la retina i que es troba associat a retinitis pigmentosa (MIM # 613581) autosòmica recessiva i a una distròfia macular viteliforme (MIM # 616152) autosòmica dominant. L'eina Varsome suggereix que aquesta variant és patogènica considerant que: i) no es troba en les bases de dades d'ExAC i gnomAD, ii) és *de novo* i iii) és una variant que trunca la proteïna. Per tant, en aquest cas tenint en compte que la distròfia macular viteliforme té una herència autosòmica dominant i que és de presentació tardana, caldria realitzar un seguiment oftalmològic a la pacient 7.

En darrer lloc, s'han analitzat els gens no associats a TND ni a cap altre malaltia. S'han analitzat detalladament aquelles variants que: i) no estan presents en les bases de dades ExAC i gnomAD ii) tenen un efecte deleteri sobre la proteïna i iii) la funció biològica que realitzen (252). A més a més dels gens *HSF2* i *CHMP7*, que s'han considerat candidats al fenotip *SA-like* en els pacients 13 i 17 (Resultats, Capítol 1 i Annex I), s'han analitzat com a possibles candidates, les variants localitzades en: a) el gen *SMCR8* que forma part del complex C9orf72-SMCR8 que té una activitat com a factor d'intercanviador de nucleòtids de guanina i actua com a regulador de la via mTOR, b) *RAB5B* que està implicat en el transport de proteïnes i el tràfic de vesícules, c) *YLPM1* que es troba implicat en la reducció de l'activitat telomerasa durant la diferenciació de les cèl·lules mare embrionàries i d) el gen *METTL24* que codifica per una metiltransferasa. Les variants en aquests gens tenen una predicció deletèria, però es classifiquen com a variants de significat incert, degut a què aquests gens no han estat associats prèviament a malaltia. Les variants localitzades en els gens *SMCR8*, *RAB5B* i *YLPM1* s'han identificat en els pacients 3, 4 i 9, que presenten variants patogèniques/probablement patogèniques en gens responsables de TND (*SYNGAP1*, *TBL1XR1* i

SPTAN1). Tot i això, aquestes variants podrien estar contribuint o modulant el fenotip. D'altra banda, en el pacient 13, sense diagnòstic molecular, s'ha identificat la variant c.32G>T, p.Cys11Phe en el gen *METTL24*. El gen *METTL24* codifica per una proteïna que pertany a la superfamília de les metiltransferases que es secreta a l'espai extracel·lular (252). Les dades d'expressió segons la base de dades *The Human protein Atlas* (253) mostren que l'ARN s'expressa principalment en teixit muscular, en teixits sexuals femenins i en el tracte gastrointestinal, però en cervell també hi ha certa expressió del gen.

En la base de dades Pubmed només hi ha una entrada d'aquest gen, on canvis en la metilació d'aquest gen s'associa amb asma atòpic (254). Es desconeix si aquest gen té alguna implicació en el neurodesenvolupament. El gen *METTL24* té un pLI de 0 i el valor de z-score és de 0,49, ambdós no són significatius, per tant, no és un gen que sigui intolerant a canvis de pèrdua de funció ni a variants *missense*. Els predictors bioinformàtics indiquen que es tracta d'una variant que tindria un efecte neutre (Provean -1,23 (*Neutral*), Polyphen2 0,062 (*Benign*), Mutation Assessor 0,895 (*low*), SNP&GO *Neutral* RI:3, MutPred 0,445, PredictSNP 61% (*Deleterious*)). Per tant, s'ha classificat com una variant de significat incert. En aquest mateix pacient, s'ha identificat la variant truncant en el gen *HSF2* candidata al fenotip *SA-like*, per tant, la variant en *METTL24* podria tenir una funció moduladora del fenotip.

Capítol 4. Anàlisis complementaris realitzats als pacients sense diagnòstic molecular

4.1 Detecció de delecions/duplicacions intragèniques en gens implicats en el neurodesenvolupament

Un dels criteris d'inclusió per entrar en aquest estudi era tenir un resultat sense alteracions patogèniques/probablement patogèniques en l'array-CGH. A tots els pacients inclosos s'havia realitzat un array-CGH de 60K (Agilent Technologies), però una de les limitacions tècniques d'aquest array és que no permet detectar petites delecions o duplicacions intragèniques. Tenint això en compte, en quatre pacients (12, 13, 14 i 15) en els quals no s'havien detectat variants patogèniques/probablement patogèniques responsables del fenotip *SA-like*, s'ha portat a terme l'estudi de canvis en número de còpies (CNVs) a partir de les dades obtingudes de WES mitjançant el programa XHMM. L'anàlisi s'ha centrat en les CNVs *de novo* i s'han seleccionat aquelles que es troben en gens implicats en el neurodesenvolupament cercant en les bases de dades SysID, OMIM i en PubMed. En els quatre pacients estudiats, s'han obtingut entre 3-9 CNVs per individu de les quals s'han seleccionat tres com a possibles candidates (Taula 14) i s'han validat amb la tècnica MLPA.

Taula 14 | CNVs predites *de novo* en gens implicats en el neurodesenvolupament. DUP, duplicació, QD, quocient de dosi.

Pacient	Posició cromosòmica	Mida (Kb)	Tipus	Gens	Phred score	QD de la MLPA	Resultat validació
12	chr14:47528320-47530883	2,56	DUP	<i>MDGA2</i>	11	1	Fals positiu
13	chr1:1454063-1454505	0,44	DUP	<i>ATAD3A</i>	24	1,57	Heretada del pare
15	chr2:228243643-228244178	0,54	DUP	<i>TM4SF20</i>	11	1,46	Heretada de la mare

Com a resultat de la validació de les CNVs detectades, s'observa que una de les tres és un fals positiu (gen *MDGA2*). El quocient de dosi de la sonda dissenyada dins de la regió duplicada en el gen *MDGA2* és de 1. En els pacients 13 i 15 les dues CNVs detectades amb una herència predita *de novo*, en realitat han estat heretades d'un dels progenitors. La duplicació identificada en el gen *ATAD3A* en el pacient 13, consisteix en la duplicació de l'exó 5 del gen i part de la regió intrònica, que es prediu que provoca un canvi en la pauta de lectura introduint un codó stop prematur, p.Val222Serfs*2. Aquest gen s'associa amb la síndrome de Harel-Yoon (MIM #617183) amb una herència autosòmica dominant o recessiva. La forma dominant es caracteritza per retard del desenvolupament, hipotonia, atròfia del nervi òptic, neuropatia perifèrica i cardiopatia hipertròfica. La forma recessiva és més greu i s'associa a retard del desenvolupament, retard de

llenguatge, cataractes congènites, epilèpsia, atàxia, hipoplàsia del nervi òptic i atròfia del cerebel (255). En el pacient 13, es podria descartar la forma dominant, considerant que la duplicació que presenta és heretada del pare que és sa.

D'altra banda, la duplicació identificada en el gen *TM4SF20* en el pacient 15, inicialment va ser predita com una variant *de novo*, però en la MLPA de validació s'observa que en realitat aquesta variant ha estat heretada de la mare, la qual també la presenta en heterozigosi. Aquesta duplicació inclou l'exó 1, part de la regió no traduïda 5' i part de l'intró 1, que es prediu que introdueix un canvi en la pauta de lectura que dóna lloc a un codó stop prematur (p.Ala62Ilefs*12). Aquest gen està associat a un trastorn específic de llenguatge (MIM # 615432) amb una herència autosòmica dominant, a més, aquests pacients també presenten anomalies cerebrals, TEA, retard del desenvolupament, però també hi ha pacients que tenen una capacitat cognitiva normal (256). Considerant que la mare és portadora sana de la variant en heterozigosi, s'ha descartat com a causant del fenotip *SA-like* present en el pacient, tot i que no es pot excloure com una variant que estigui modulant o contribuint al fenotip.

El fet de què s'hagin identificat un fals positiu i variants *de novo* quan en realitat eren variants heretades d'un dels progenitors, és a causa de què hi ha una gran dispersió en la cobertura de les regions analitzades, fent que sigui difícil de predir amb certesa CNVs i de determinar el patró d'herència. En la Figura 21 es troba representat el gràfic de la profunditat de cobertura dels quatre exons del gen *TM4SF20*, en verd es pot observar la duplicació en el pacient, la línia lila correspon a la mare que es troba en el límit i el color gris correspon a altres mostres utilitzades per a la normalització, mostrant que hi ha una gran dispersió de la cobertura en aquesta regió. És per aquest motiu, que les CNVs predites s'han de validar amb una altra tècnica, en aquest cas la MLPA per confirmar que realment estan presents i el patró d'herència.

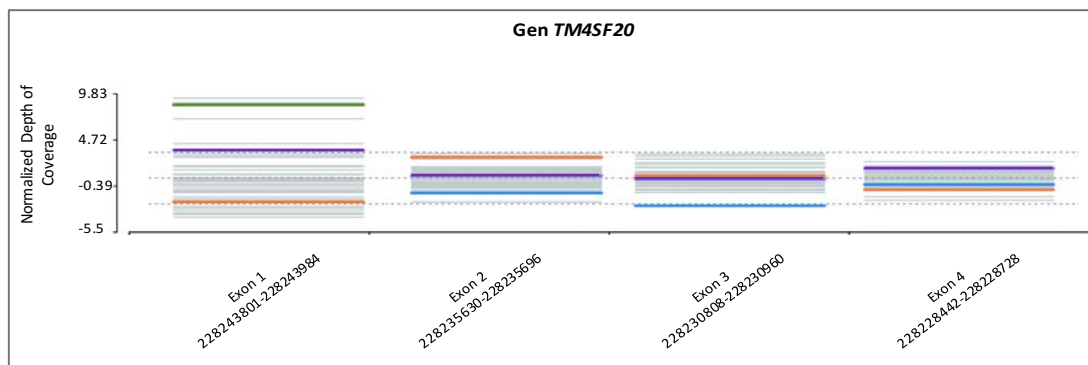


Figura 21 | Gràfic que mostra la normalització de la profunditat de cobertura dels exons 1-4 del gen *TM4SF20* del trio format pel pacient 15 i els seus progenitors. En color verd es troba representada la cobertura del pacient en la regió duplicada, en color blau la cobertura del pacient en les regions on no hi ha CNVs, el color lila correspon a la mare. El color taronja correspon al pare i el gris correspon a les altres mostres utilitzades per fer la normalització i que no corresponen al trio.

4.2 Avaluació de variants en mosaic

En els TND i concretament en les encefalopaties epilèptiques s'ha descrit variants en mosaic responsables del fenotip. En concret, en nou gens (*CDKL5*, *GABRA1*, *GABRG2*, *GRIN2B*, *KCNQ2*, *MECP2*, *PCDH19*, *SCN1A* i *SCN2A*) relacionats amb epilèpsia s'ha reportat un 3,5% de mosaicisme (257). En aquests casos el mosaic seria post-zigòtic i seria un fenomen que es produiria *de novo* en el pacient. D'altra banda, també podria trobar-se mosaicisme en els progenitors que no presentarien fenotip. De fet, en un altre estudi portat a terme en encefalopaties epilèptiques s'ha reportat que un 10% dels progenitors són mosaics per la variant patogènica present en el seu fill/filla i que suposadament era *de novo* (258). En quatre dels cinc pacients sense diagnòstic molecular s'ha portat a terme l'anàlisi dels mosaics, el trio 17 es va excloure de l'anàlisi. L'anàlisi dels mosaics en aquests quatre pacients (pacients 12, 13, 14 i 15) consistia per una banda, en analitzar variants *de novo* degudes a un mosaic post-zigòtic i per l'altra, variants en mosaic en un dels progenitors i que es trobessin en heterozigosi en els pacients.

Variants *de novo* en mosaic candidates al fenotip SA-like

L'anàlisi de variants *de novo* en mosaic en els quatre pacients ha resultat en la identificació de 18-28 variants per individu, filtrant per una freqüència al·lèlica entre el 10-30%, una freqüència poblacional <1/1000 i prioritant les variants amb un efecte deleteri sobre la proteïna. S'han considerat rellevants les variants localitzades en gens implicats en el neurodesenvolupament. En la Taula 15 es troben resumides les dues variants candidates identificades en els pacients 12 i 13.

Taula 15 | Variants candidates resultants de l'anàlisi de variants en mosaic *de novo*.

	Gen	Isoforma	Tipus de variant	Canvi de nucleòtid	Canvi d'aminoàcid	ExAC/gnomAD	Freqüència al·lèlica %	Herència
Pacient 12	<i>TUBB2B</i>	NM_178012.5	Missense	c.751C>A	p.Arg251Ser	- / -	12,5%	<i>De novo</i>
Pacient 13	<i>ATAD3A</i>	NM_018188.5	Missense	c.1301G>A	p.Arg434Gln	0,00004423/ 0,0001004	11.8%	<i>De novo</i>

El gen *TUBB2B* codifica per la isoforma β de la tubulina i s'associa a una displàsia cortical complexa amb altres malformacions cerebrals, presenta una herència autosòmica dominant (MIM # 610031). La variant *missense* c.751C>A identificada en la pacient 12 es troba en un mosaic del 12,5%. L'eina Varsome suggereix que aquesta variant és probablement patogènica, atès que no es troba en les bases de dades d'ExAC i gnomAD i els predictors *in silico* la classifiquen com una variant deletèria. Aquesta variant es troba en la base de dades de ClinVar (VCV000432230.2) classificada com una variant de significat incert. Considerant que la pacient 12 no presenta malformacions cerebrals que és una característica predominant en els pacients que presenten

variants patogèniques/probablement patogèniques en aquest gen (259,260), s'ha decidit classificar-la com una variant de significat incert.

La variant *missense* c.1301G>A, p.Arg434Gln en el gen *ATAD3A* identificada en el pacient 13 es troba en un mosaic del 11,8%. El programa Varsome suggereix que aquesta variant és probablement patogènica, atès que es troba en molt baixa freqüència en les bases de dades d'ExAC i gnomAD i els predictors *in silico* la classifiquen com a deletèria. L'anàlisi de CNVs ha identificat una duplicació de l'exó 5 del gen *ATAD3A* que es prediu que introdueix un canvi en la pauta de lectura. Per tant, s'hauria d'estudiar a nivell funcional si la variant c.1301G>A en un mosaic del 11,8% i la CNV són suficients per presentar la forma recessiva de la síndrome de Harel-Yoon. No obstant, el pacient 13 no presenta moltes de les característiques clíniques de la forma recessiva com les cataractes congènites, l'atròfia del cerebel, l'afectació del nervi òptic, l'atàxia o la hipotonia. Per aquest motiu, les dues variants identificades en el gen *ATAD3A* s'han classificat com a variants de significat incert.

En síntesi, l'anàlisi de variants en mosaic *de novo* ha donat lloc a la identificació de dues variants candidates en els pacients 12 i 13 classificades com a variants de significat incert, que hauran de ser revisades en el futur, pel fet que podrien tenir rellevància en el fenotip *SA-like* que presenten aquests pacients.

Variants en mosaic en els progenitors i en heterozigosi en els pacients candidates al fenotip *SA-like*

L'anàlisi de variants en mosaic (5-30%) en un dels progenitors i present en heterozigosi en el pacient (30-70%), filtrades per una freqüència al·lèlica <1/1000 en les bases de dades d'ExAC i gnomAD, ha donat com a resultat entre 13 i 24 variants en els quatre individus estudiats. Les variants obtingudes s'han filtrat per gens implicats en el neurodesenvolupament i segons l'impacte que tenen sobre la proteïna. Les variants amb una predicció neutra s'han descartat. Aquest anàlisi no ha permès identificar variants candidates a ser responsables del fenotip *SA-like*.

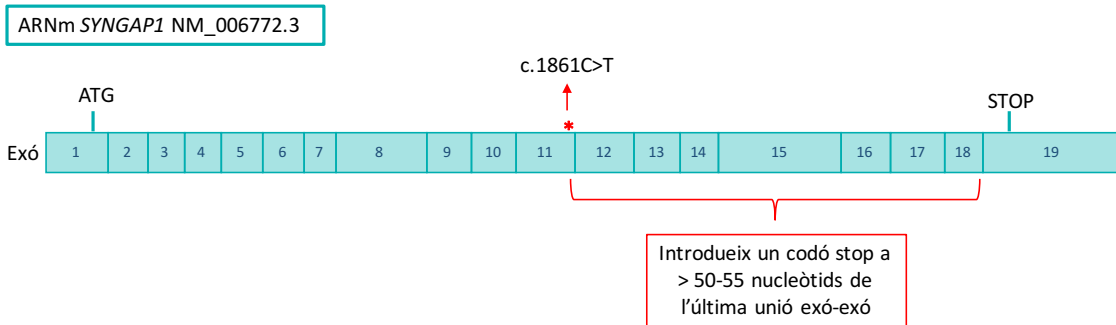
Capítol 5. Activació del mecanisme *Nonsense-mediated mRNA Decay* (NMD) per variants *nonsense* i *frameshift*

El mecanisme de control post-transcripcional NMD actua degradant l'ARNm de transcrits amb codons stop prematurs. Variants *nonsense* i *frameshift* que introdueixen un codó stop prematur que es troba localitzat a >50-55 nucleòtids de l'última unió exó-exó provoquen l'activació d'aquest mecanisme i com a conseqüència la degradació de l'ARNm. Amb l'objectiu de comprovar si les variants *nonsense* i *frameshift* identificades com a responsables o candidates al fenotip SA-like en la nostra cohort activen el mecanisme NMD, s'ha quantificat l'al·lel amb la variant en ADN genòmic i en ADNc i s'ha avaluat l'expressió gènica per qPCR amb sondes TaqMan.

Variante *nonsense* c.1861C>T, p.Arg621* en el gen *SYNGAP1*

La variant c.1861C>T identificada en la pacient 3 introdueix un codó stop prematur situat a >50-55 nucleòtids de l'última unió exó-exó (Figura 22A), per tant, es preveu que s'activi el mecanisme NMD degradant-se l'ARNm. Es van dissenyar primers per amplificar l'exó 11 on es troba situada la variant en ADN genòmic i ADNc en el pacient 3 i en una mostra control d'un individu sa, i es van analitzar les dades amb el programa *Minor Variant Finder*.

A)



B)

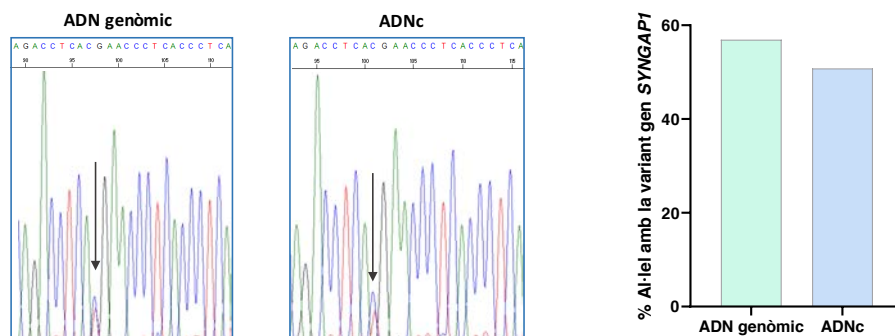


Figura 22 Quantificació de l'al·lel amb la variant c.1861C>T en el gen *SYNGAP1* en ADN genòmic i ADNc. **A)** Representació esquemàtica del gen *SYNGAP1* on es troba assenyalada la variant i la distància a la unió exó-exó més 3'. **B)** Seqüenciació Sanger de l'exó 11 del gen *SYNGAP1* i gràfic que indica el percentatge al qual es troba la variant c.1861C>T en ADN genòmic i ADNc.

El resultat mostren que no hi ha diferències en el percentatge de l'al·lel amb la variant, tant en ADN genòmic com en ADNc es troba al voltant del 50% (Figura 22B).

L'anàlisi de l'expressió gènica per qPCR indica que no hi ha diferències en l'expressió gènica del gen *SYNGAP1* en la pacient 3 respecte a la mostra control (p-valor 0,53) (Figura 23).

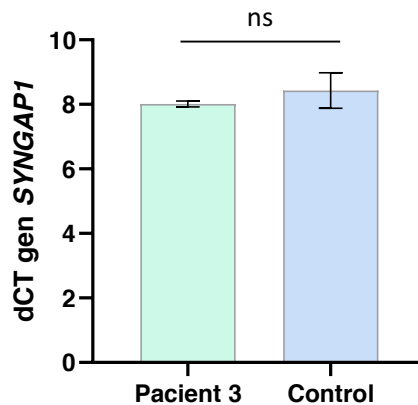


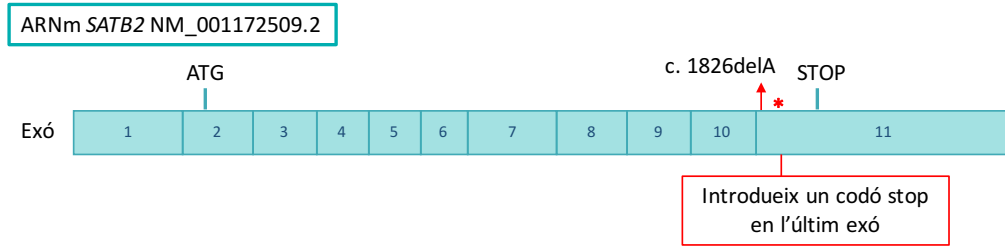
Figura 23 | Expressió del gen *SYNGAP1*. No hi ha diferències significatives entre la pacient 3 i la mostra control. Test estadístic T-Student (p-valor 0,53); ns, no significatiu.

Tenint en compte els resultats de la quantificació de l'al·lel en ADN genòmic i ADNc, juntament amb els de l'expressió gènica per qPCR, la variant c.1861C>T escapa el mecanisme NMD i per consegüent, no es degradaria l'ARNm, resultant en una proteïna truncada.

Variant *frameshift* c.1826delA, p.Asp609Alafs*15 en el gen *SATB2*

La variant c.1826delA es troba situada en l'últim exó i introdueix un codó stop prematur 332 nucleòtids abans del codó stop prematur natural (Figura 24A). S'ha descrit que les variants que introdueixen un codó stop en l'últim exó escapen el mecanisme NMD, per tant, no hi hauria degradació de l'ARNm, generant-se una proteïna truncada.

A)



B)

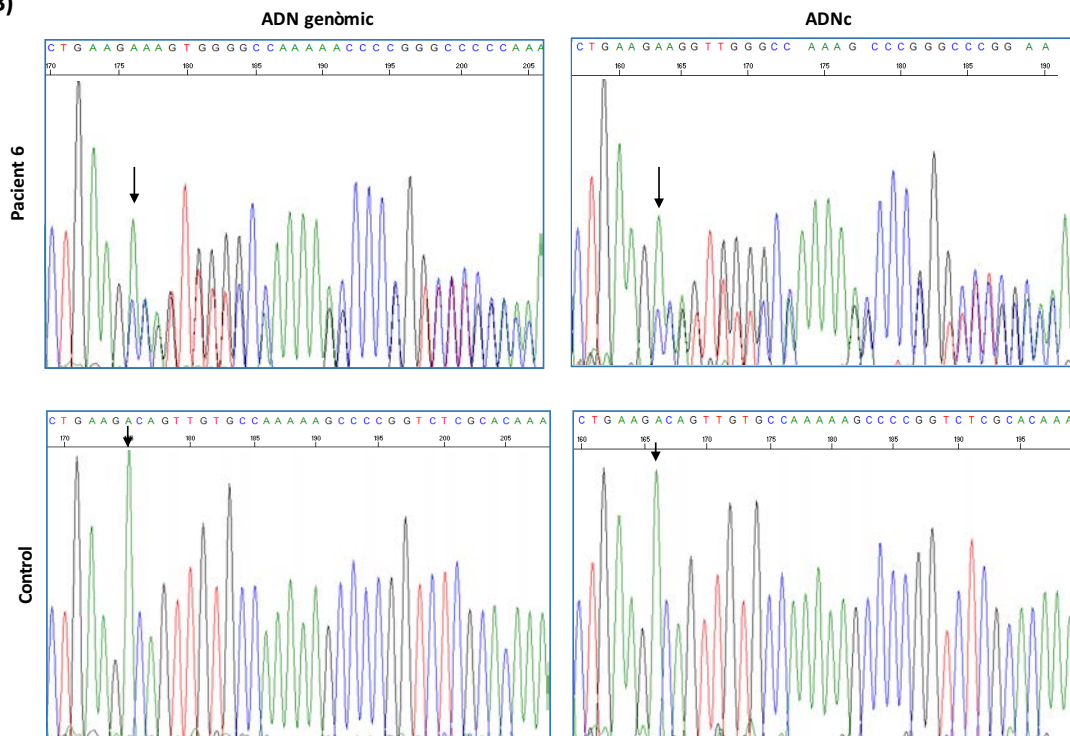


Figura 24 | Quantificació de l'al·lel amb la variant c.1826delA en el gen *SATB2* en ADN genòmic i ADNc. **A)** Representació esquemàtica del gen *SATB2* on es troba assenyalada la variant. **B)** Seqüenciació Sanger del gen *SATB2* en ADN genòmic i ADNc en la pacient 6 i en una mostra control.

Es va amplificar la regió de l'exó 11 on es troba localitzada la variant c.1826delA en el gen *SATB2* en ADN genòmic i en ADNc (Figura 24B). Al tractar-se d'una variant *frameshift*, el *software Minor Variant Finder* no permet calcular el percentatge al qual es troba l'al·lel amb la variant. Per calcular el percentatge s'ha dividit l'alçada a la qual es troben 10 nucleòtids corresponents al canvi en la pauta de lectura entre el valor dels mateixos 10 nucleòtids en la mostra control. Aquest càlcul s'ha portat a terme tant per la quantificació de l'al·lel en ADN genòmic com en ADNc. El resultat mostra que la variant en ADN genòmic es troba en un 40,2% i en ADNc en un 43,4%, indicant que no hi ha diferències en el percentatge de l'al·lel amb la variant (Figura 25).

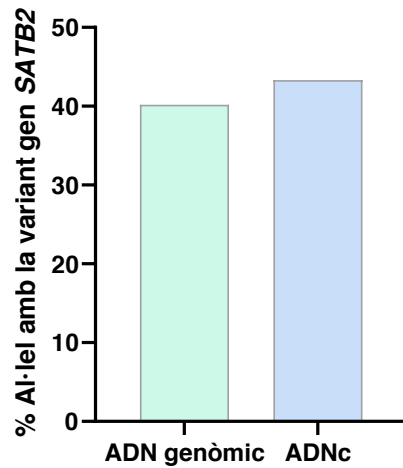


Figura 25 | Percentatge al qual es troba la variant c.1826delA en ADN genòmic i ADNc.

D'altra banda, l'anàlisi de l'expressió gènica del gen *SATB2* per qPCR mostra que no hi ha diferències significatives respecte al control negatiu (p-valor 0,0612) (Figura 26). Per tant, els resultats obtinguts de la qPCR estan d'acord amb els obtinguts en la quantificació de l'al·lel i no hi ha degradació de l'ARNm, predient-se la síntesi d'una proteïna truncada.

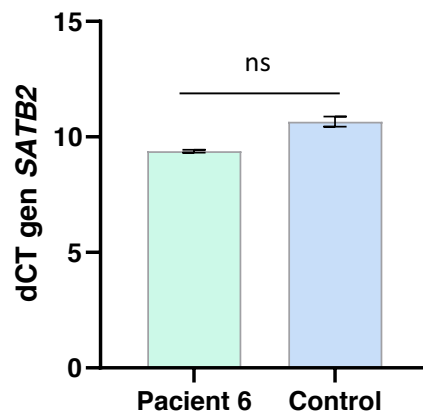
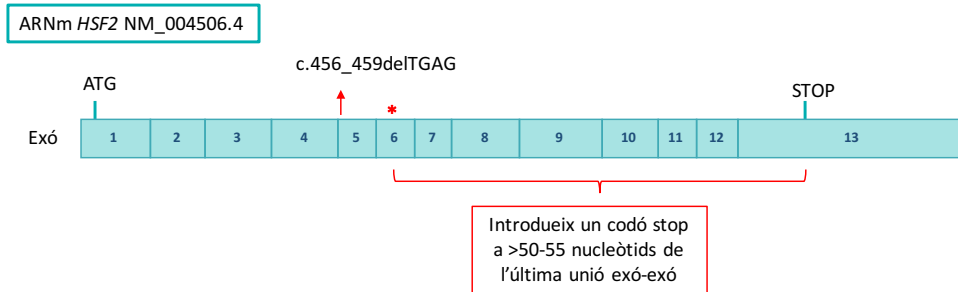


Figura 26 | Expressió del gen *SATB2*. No hi ha diferències significatives entre la pacient 6 i la mostra control. Test estadístic T-Student (p-valor 0,0612); ns, no significatiu.

Variant *frameshift* c.456_459delTGAG, p.Ser152Argfs*40 en el gen *HSF2*

La variant c.456_459delTGAG en el gen *HSF2* elimina quatre nucleòtids inicials de l'exó 5, produint un canvi en la pauta de lectura que introdueix un codó stop prematur en l'exó 6, situat a >50-55 nucleòtids de l'última unió exó-exó (Figura 27A). Es va amplificar l'exó 5 del gen *HSF2* on es troba situada la variant c.456_459delTGAG en ADN genòmic i ADNc en la pacient 13 i en un control negatiu (Figura 27B).

A)



B)

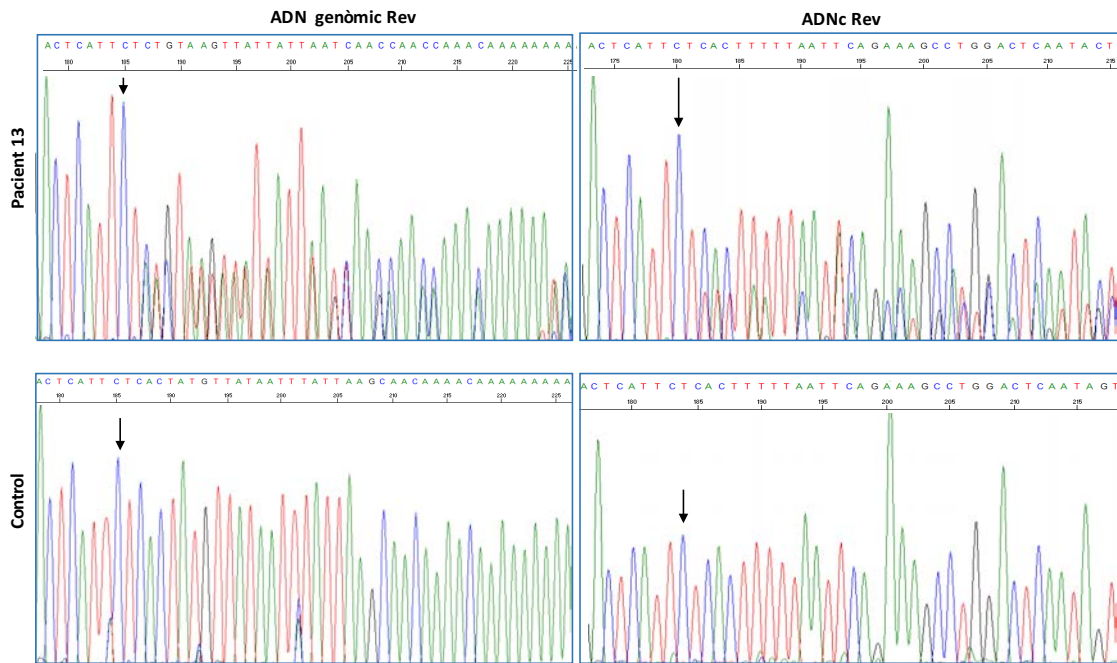


Figura 27 | Quantificació de l'al·lel amb la variant c.456_459delTGAG en el gen *HSF2* en ADN genòmic i ADNc. A) Representació esquemàtica del gen *HSF2* on es troba assenyalada la variant i on s'introdueix el codó stop prematur. B) Seqüenciació Sanger del gen *HSF2* en ADN genòmic i ADNc en la pacient 13 i en el control negatiu.

El càlcul del percentatge de l'al·lel amb la variant s'ha realitzat seguint el mateix procediment que per la variant *frameshift* en el gen *SATB2*. Els resultats mostren que en ADN genòmic l'al·lel amb la variant es troba en un 43%, mentre que en ADNc es troba en un 33% (Figura 28).

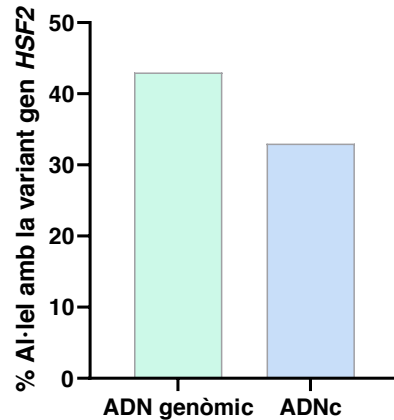


Figura 28 | Percentatge al qual es troba la variant c.456_459delTGAG en ADN genòmic i ADNc.

L'anàlisi de l'expressió gènica per qPCR del gen *HSF2* mostra que hi ha diferències significatives en l'expressió respecte la mostra control (p-valor 0,014) (Figura 29).

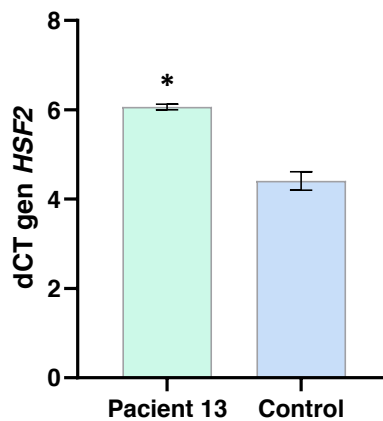


Figura 29 | Expressió del gen *HSF2*. En el pacient 13 hi ha una disminució significativa de l'expressió respecte a la mostra control. Test estadístic T-Student (p-valor 0,014); *p-valor<0,05.

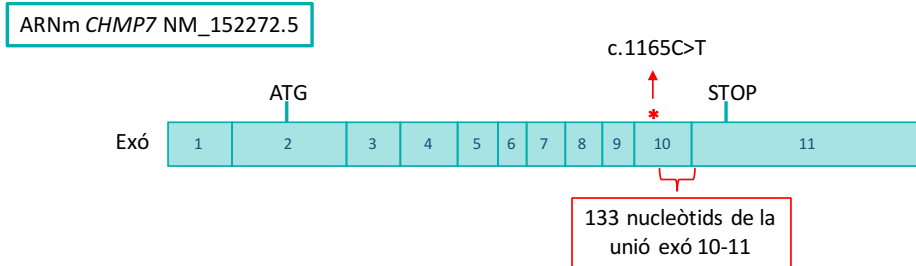
Tant els resultats de la quantificació de l'al·lel amb la variant com els resultats de la qPCR indiquen que la variant c.456_459delTGAG activa el mecanisme NMD.

Variant *nonsense* c.1165C>T, p.Gln389* en el gen *CHMP7*

La variant c.1165C>T es troba localitzada a 133 nucleòtids *upstream* de l'última unió exó-exó, per tant, hauria de ser subjecte de degradació pel mecanisme NMD (Figura 30A). Es va amplificar l'exó 10 on es troba situada la variant c.1165C>T en ADN genòmic i en ADNc. Els resultats mostren

que hi ha una disminució del percentatge de l'al·lel amb la variant, en ADN genòmic es troba al voltant del 50%, mentre que en ADNc es troba al 30% (Figura 30B).

A)



B)

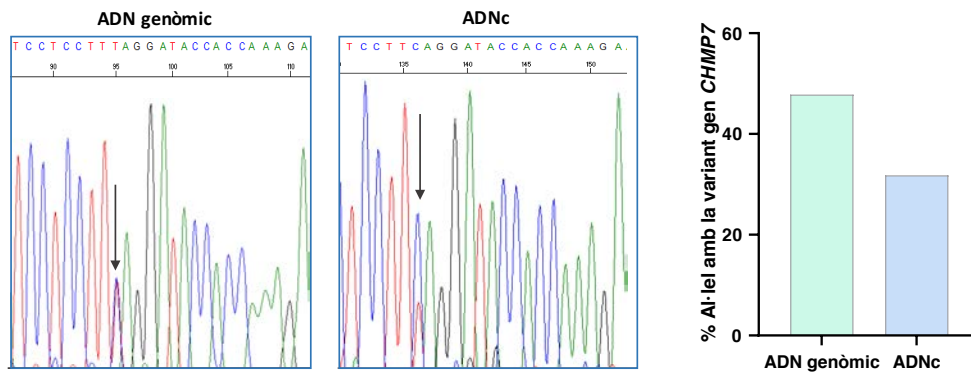


Figura 30 | Quantificació de l'al·lel amb la variant c.1165C>T en el gen *CHMP7* en ADN genòmic i ADNc. **A)** Representació esquemàtica del gen *CHMP7* on es troba assenyalada la variant i la distància a la unió exó-exó més 3'. **B)** Seqüenciació Sanger de l'exó 10 del gen *CHMP7* i gràfic que indica el percentatge al qual es troba la variant c.1165C>T en ADN genòmic i ADNc.

D'altra banda, en els resultats de la qPCR s'observa que hi ha una disminució significativa de l'expressió del gen *CHMP7* respecte a la mostra control (p-valor 0,016) (Figura 31).

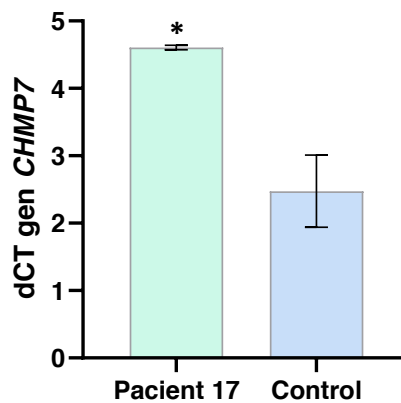


Figura 31 | Expressió del gen *CHMP7*. En el pacient 17 hi ha una disminució significativa de l'expressió respecte a la mostra control. Test estadístic T-Student (p-valor 0,016); *p-valor<0,05.

En conjunt, els resultats tant de la quantificació de l'al·lel com els del la qPCR indiquen que el transcrit amb la variant c.1165C>T indueix l'activació del mecanisme NMD.

Variant *nonsense* c.3106C>T, p.Arg1036* el gen *ASXL3*

La variant c.3106C>T, p.Arg1036* introdueix un codó de stop prematur en l'últim exó, per tant, l'ARNm no hauria de ser subjecte de degradació pel mecanisme NMD. Es van amplificar els exons 11-12 del gen *ASXL3* en ADN genòmic i en ADNc. Els resultats mostren que la variant c.3106C>T es troba aproximadament al 50% en ADN genòmic, mentre que no va haver amplificació en ADNc extret de limfòcits. Es va dissenyar un nou *primer forward* però tampoc es va observar amplificació. D'acord amb la base de dades *The Human Protein Atlas* (253), el gen *ASXL3* s'expressa principalment en cervell i en teixits sexuals masculins i femenins (Figura 32). Amb la finalitat de comprovar que la causa de l'absència d'amplificació en el pacient és deguda a que aquest gen s'expressa molt poc o no s'expressa en limfòcits, es van amplificar els exons 11-12 del gen *ASXL3* en les línies de neuroblastoma SKNBE(2) i CHLA-90 (cedides pel grup de recerca translacional en càncer en la infància i l'adolescència, VHIR).

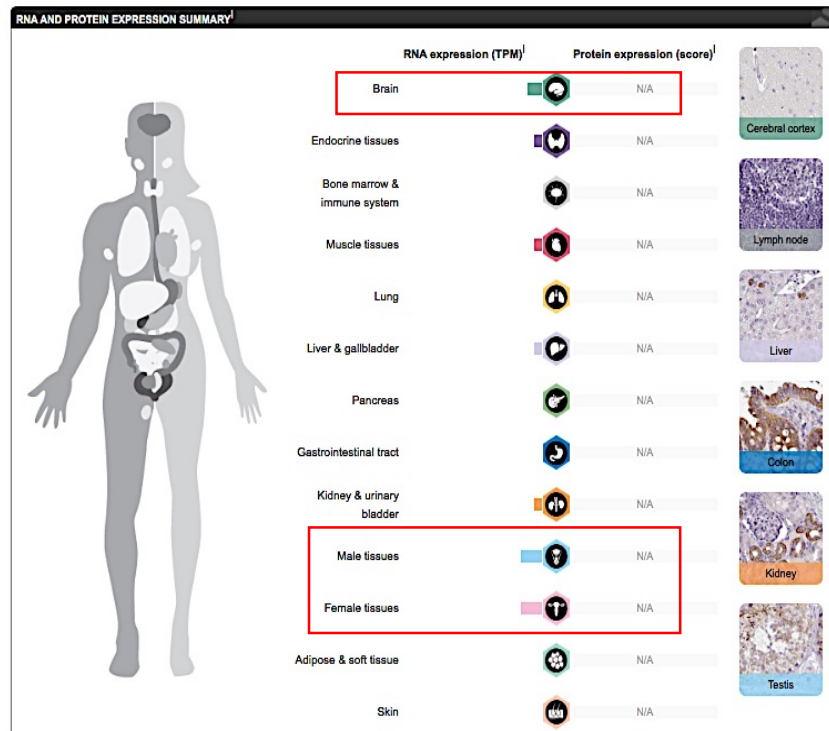


Figura 32 | Expressió de l'ARN i de la proteïna ASXL3 segons la base de dades *The Human Protein Atlas* (www.proteinatlas.org).

En la Figura 33 s'observa que hi ha amplificació dels exons 11-12 en les línies de neuroblastoma utilitzant les dues parelles de *primers*, mentre que no hi ha amplificació en ADNc extret de limfòcits. Aquest resultat confirma que el gen *ASXL3* no s'expressa o s'expressa molt poc en sang perifèrica. Per aquest motiu, es necessitaria mostra d'un teixit que derivi del neuroectoderm, per analitzar si el transcrit amb la variant c.3106C>T activa el mecanisme NMD.

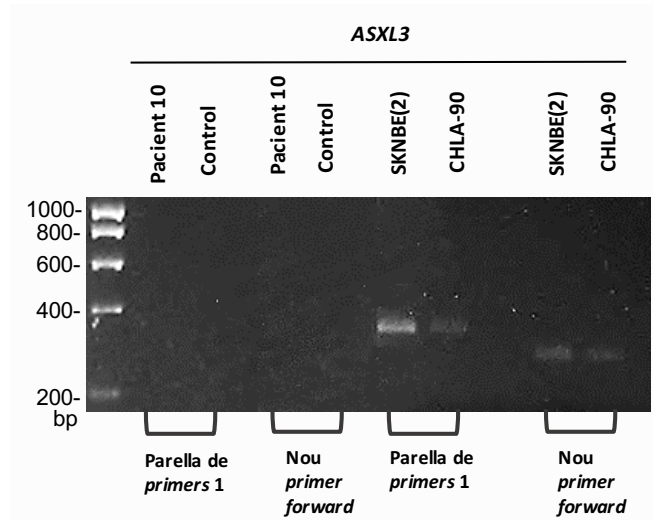


Figura 33 | Gel d'agarosa on s'han carregat els productes d'amplificació dels exons 11-12 del gen *ASXL3* en ADNc del pacient 10, de la mostra control i de les línies de neuroblastoma SKNBE(2) i CHLA-90, fent servir la parella de *primers 1* i amb un nou *primer forward*.

Capítol 6. Impacte de la variant R169T sobre la funció de KIF1A

En el pacient 1 s'ha identificat la variant c.506G>C, p.Arg169Thr (R169T) localitzada en el domini motor del gen *KIF1A* com a responsable del fenotip SA (Figura 34). Aquesta variant es va classificar com a probablement patogènica, ja que compleix els criteris de patogenicitat següents: i) no es troba en bases de dades de població control (ExAC i gnomAD), ii) els predictors *in silico* indiquen que es tracta d'una variant amb un efecte deleteri sobre la funció de la proteïna, iii) es troba en un domini funcional ben caracteritzat i conservat, on més del 80% de les variants són patogèniques i iv) és *de novo*. Tot i això, al tractar-se d'una variant no descrita anteriorment, segons la guia ACMG es necessiten portar a terme estudis funcionals que demostrin l'efecte deleteri de la variant per classificar-la com a patogènica. Per determinar com afecta la variant R169T a la funció de KIF1A s'ha realitzat l'estudi de l'activitat ATPasa del domini motor de KIF1A.

Proteïna KIF1A

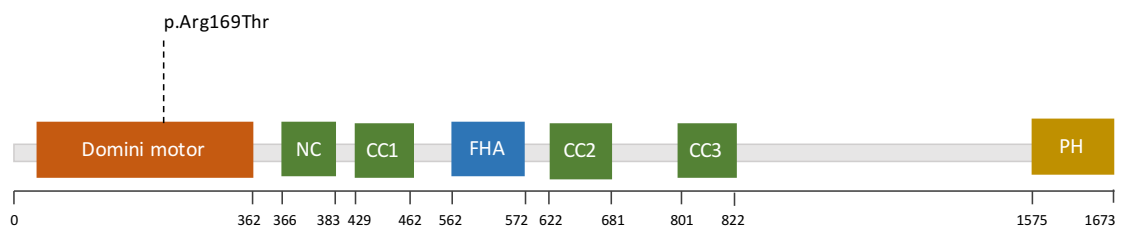


Figura 34 | Estructura de la proteïna KIF1A on es troba indicada la variant p.Arg169Thr identificada en el pacient 1. NC, Neck-coil; CC, Coiled Coil; FHA, forkhead-associated; PH, pleckstrin homology domain.

La nostra hipòtesi és que la variant R169T afectarà a l'activitat ATPasa, com a conseqüència de defectes en la unió als microtúbuls. L'Arg169 és un aminoàcid amb càrrega positiva que es troba localitzat en un domini d'unió a microtúbuls i estableix interaccions electrostàtiques amb els aminoàcids amb càrrega negativa de la tubulina. Per contra, l'aminoàcid Thr no té càrrega i probablement no interaccioni o tingui una menor afinitat per la tubulina.

6.1 Activitat ATPasa del domini motor de KIF1A amb la variant R169T

Un dels models desenvolupats per avaluar la motilitat de les quinesines són els assajos ATPasa, que permeten quantificar el Pi alliberat pel domini motor d'una quinesina en presència d'ATP i microtúbuls, i de forma indirecta indiquen si la motilitat de la quinesina està preservada o no. La finalitat de realitzar aquest assaig ha estat avaluar l'impacte de la variant *missense* R169T localitzada en el domini motor de KIF1A sobre la motilitat de la proteïna.

4. RESULTATS

Per fer aquest assaig s'ha produït i purificat *in vitro* la forma WT i mutada (R169T) del domini motor de KIF1A. Per determinar que hi havia expressió de la proteïna recombinant després de la inducció amb IPTG del cultiu de cèl·lules BL21 transformades amb el vector d'expressió pQE-80L KIF1A WT/R169T, el pellet de cèl·lules induïdes i no induïdes s'ha lissat i s'han carregat les mostres en un gel de poliacrilamida. A continuació, s'han transferit les proteïnes a una membrana de PVDF i s'ha realitzat una tinció amb *Naphtol Blue*. Es pot observar una banda a 43 KDa corresponent al domini de KIF1A WT/R169T després de la inducció amb IPTG (Figura 35A). També s'ha comprovat per *Western blot* que hi ha expressió de les proteïnes recombinants amb un anticòs que detecta la cua d'histidines (Figura 35B).

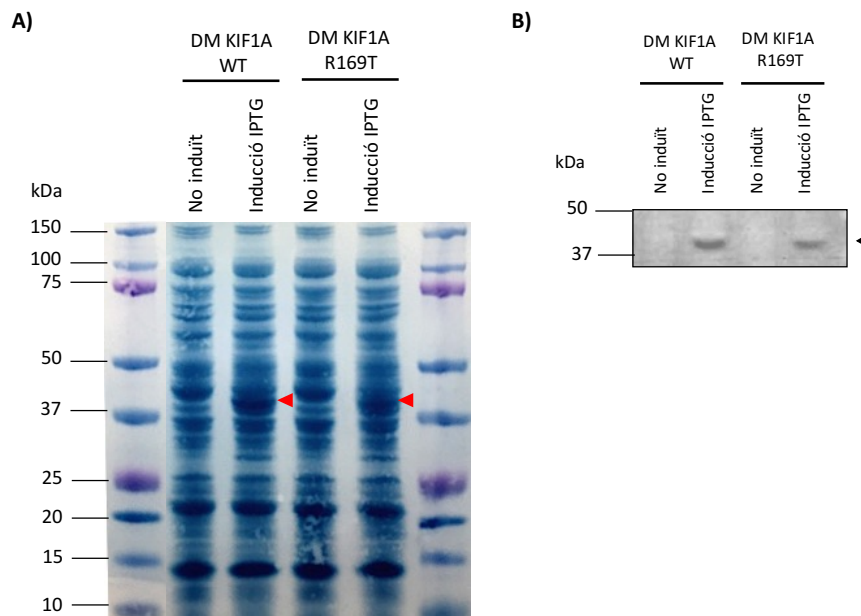


Figura 35 | Anàlisi d'expressió de la proteïna recombinant. **A)** Tinció amb *Naphtol Blue* d'extractes proteics abans i després de la inducció de l'expressió dels dominis motors WT i R169T. La fletxa vermella indica l'expressió de la proteïna. **B)** Immunodetecció de la proteïna recombinant que conté una cua de 6 histidines. Es pot observar com després de la inducció amb IPTG del cultiu hi ha expressió del domini motor de KIF1A WT i amb la variant R169T.

A continuació, s'ha realitzat la purificació de la proteïna i per comprovar que ha funcionat correctament s'han carregat les mostres dels diferents passos de la purificació en un altre gel, s'ha fet la transferència a una membrana de PVDF i una tinció amb *Naphtol Blue*. En els dos gels, es pot observar una banda de 43 KDa que correspon al domini motor de KIF1A WT i R169T (Figura 36).

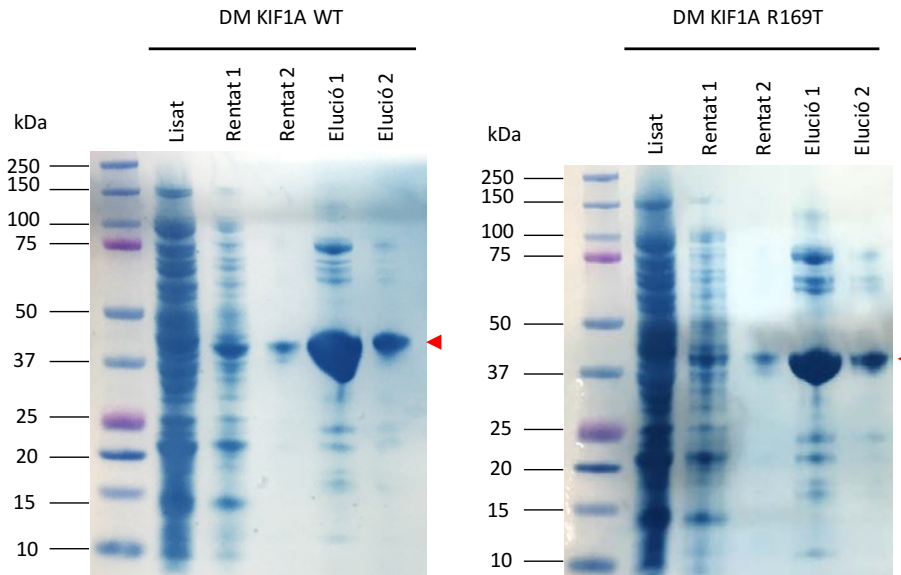


Figura 36 | Tinció amb *Naphtol blue* de les membranes corresponents als diferents passos de la purificació de la proteïna. En les fraccions de l'elució es pot observar una banda d'uns 43KDa que correspon al domini motor de KIF1A WT i R169T.

Els resultats obtinguts de l'assaig ATPasa mostren que a mesura que augmenta la quantitat de proteïna corresponent al domini motor de KIF1A WT en el rang de 0,078 μg a 0,24 μg , hi ha un augment exponencial del Pi alliberat (Figura 37). Quan la quantitat de proteïna és de 0,24 μg s'arriba a la màxima activitat ATPasa, que correspon a 1,76 nmols de Pi alliberat, el que suposa una activitat de 1466,6 nmols d'ATP hidrolitzat/min/mg. En el cas del domini motor amb la variant R169T el que s'observa és que a baixes quantitats de proteïna no hi ha activitat ATPasa i en el màxim valor de proteïna, 0,24 μg , la quantitat de Pi alliberat és de 0,010 nmols (Figura 37), el que suposa una activitat de 8,3 nmols d'ATP hidrolitzat/min/mg.

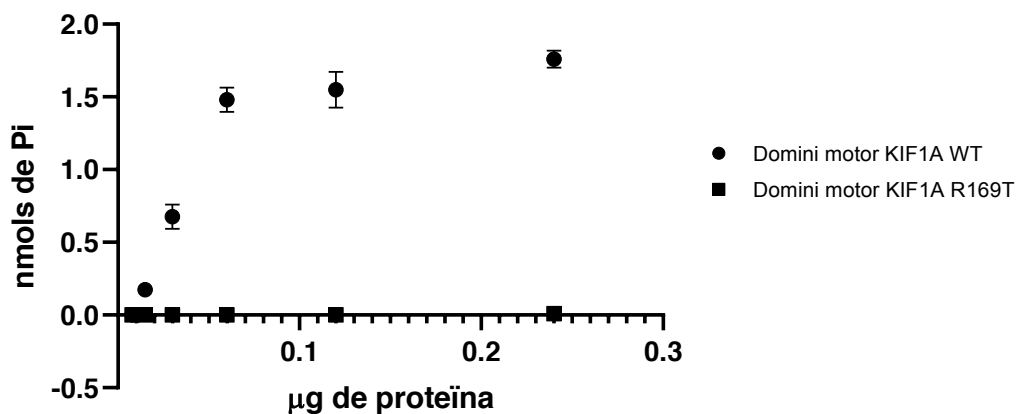


Figura 37 | Resultats de l'assaig ATPasa. A mesura que augmenta la quantitat de proteïna del domini motor de KIF1A WT es produeix un increment exponencial en l'alliberament de Pi. Per contra, s'observa que la variant R169T afecta a l'activitat ATPasa del domini motor de KIF1A, ja que no hi ha pràcticament alliberament de Pi a mesura que augmenta la quantitat de proteïna (T-Student, p-valor 0,0095).

4. RESULTATS

Per demostrar que es partia de la mateixa concentració de proteïna del domini motor WT i R169T i que les diferències observades en l'activitat ATPasa són degudes a què la variant R169T afecta a l'activitat ATPasa de la proteïna, s'ha carregat un gel amb diferents dilucions de la proteïna WT i R169T (3 µg, 1 µg i 0,3 µg). Les proteïnes s'han transferit a una membrana de PVDF i s'ha fet una tinció amb *Naphtol blue*. Els resultats confirmen que es parteix de la mateixa concentració de proteïna (Figura 38) i per tant, les diferències en l'activitat són degudes a l'impacte de la variant R169T sobre la funció del domini motor de KIF1A.

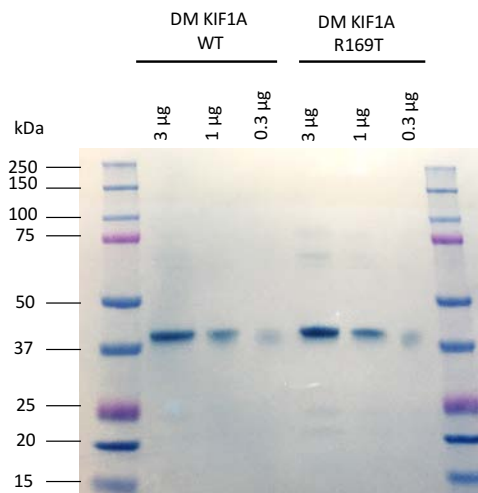


Figura 38 | Tinció de la membrana on s'han carregat diferents dilucions del domini motor WT i amb la variant R169T, que demostra que es parteix de la mateixa concentració de proteïna.

En resum, s'ha demostrat que el domini motor amb la variant R169T presenta una activitat ATPasa significativament inferior a la del domini motor WT (T-Student, p-valor 0,0095), indicant un efecte deleteri de la variant sobre la motilitat de la quinesina KIF1A. L'absència de motilitat de KIF1A afectaria el transport de proteïnes i vesícules a través dels axons neuronals, sent la responsable del fenotip present en el pacient 1.

Capítol 7. Xarxa d'interacció dels nous gens identificats responsables de la SA-like, els gens del diagnòstic diferencial de la SA i el gen *UBE3A*

S'ha construït una xarxa d'interacció que inclou els nous gens identificats en la nostra cohort com a responsables del fenotip SA-like, els gens del diagnòstic diferencial de la SA prèviament descrits i el gen *UBE3A*, amb l'objectiu d'analitzar si hi havia interaccions directes proteïna-proteïna, i les funcions i vies moleculars enriquides en aquesta xarxa. S'han analitzat les interaccions proteïna-proteïna amb l'eina STRING, seleccionant les interaccions amb un *score* de confiança del 0,7 i s'han avaluat els resultats activant i inactivant l'opció de *text mining*.

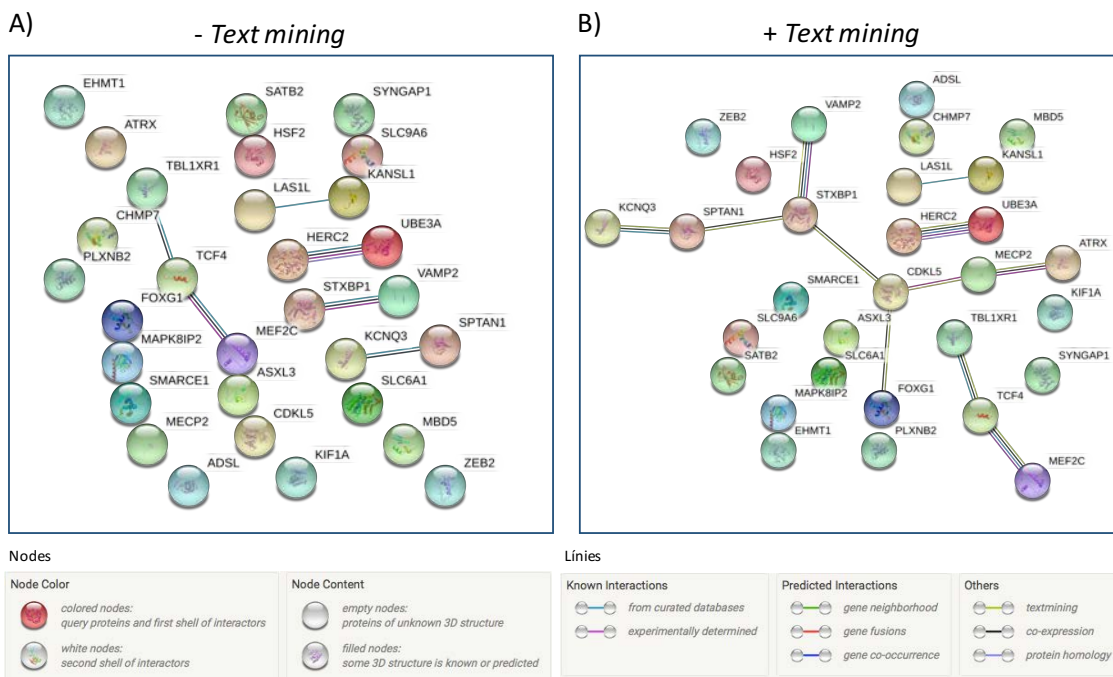


Figura 39 | Resultats obtinguts de l'anàlisi de la xarxa d'interacció dels nous gens identificats, els gens del diagnòstic diferencial de la SA i el gen *UBE3A* amb STRING. **A)** Resultats obtinguts inactivant la opció *text mining* **B)** Resultats obtinguts activant la opció *text mining*, on es pot observar que hi ha més interaccions.

Quan es comparen els resultats obtinguts activant i desactivant l'opció *text mining* s'observa un augment d'interaccions entre els gens identificats en la nostra cohort i els gens del diagnòstic diferencial de la SA quan està activada (SPTAN1-STXBP1, STXBP1-CDKL5, CDKL5-FOXG1, MECP2-ATRX i CDKL5-MECP2) (Figura 39). L'opció *text mining* fa un anàlisi estadístic de textos científics que citen dues proteïnes a la vegada, això vol dir que pot associar dues proteïnes que poden no estar interaccionant ni estar relacionades funcionalment. Per tant, per fer un anàlisi més acurat de les interaccions s'ha inactivat aquesta opció. Els resultats de l'anàlisi mostren que els gens identificats en la nostra cohort no interaccionen amb el gen *UBE3A*, però si que en tres casos (TBL1XR1-TCF4, LAS1L-KANSL1 i STXBP1-VAMP2) interaccionen amb els gens del diagnòstic

4. RESULTATS

diferencial de la SA i en un cas interaccionen dos dels gens identificats en la nostra cohort (KCNQ3-SPTAN1).

Taula 16 | Enriquiment funcional de la xarxa dels gens identificats en la nostra cohort, els del diagnòstic diferencial de la SA i el gen *UBE3A* obtingut amb el programa STRING. Es mostren els principals termes GO amb un FDR <0,05.

Biological Process (TOP 20 GO terms)			
<i>GO term</i>	<i>Description</i>	<i>Count in gene set</i>	<i>False discovery rate (FDR)</i>
GO:0007399	<i>Nervous system development</i>	18 of 2206	3,26e-07
GO:0022008	<i>Neurogenesis</i>	14 of 1519	1,06e-05
GO:0048699	<i>Generation of neurons</i>	13 of 1422	3,21e-05
GO:0016043	<i>Cellular component organization</i>	21 of 5163	0,00030
GO:0030182	<i>Neuron differentiation</i>	10 of 940	0,00030
GO:0048731	<i>System development</i>	19 of 4144	0,00030
GO:0048666	<i>Neuron development</i>	9 of 758	0,00032
GO:0007268	<i>Chemical synaptic transmission</i>	7 of 402	0,00038
GO:0031175	<i>Neuron projection development</i>	8 of 616	0,00038
GO:0048667	<i>Cell morphogenesis involved in neuron differentiation</i>	7 of 400	0,00038
GO:0007420	<i>Brain development</i>	8 of 650	0,00050
GO:0048167	<i>Regulation of synaptic plasticity</i>	5 of 164	0,00053
GO:0006325	<i>Chromatin organization</i>	8 of 683	0,00059
GO:0048468	<i>Cell development</i>	11 of 1493	0,00061
GO:0050804	<i>Modulation of chemical synaptic transmission</i>	6 of 316	0,00061
GO:0051962	<i>Positive regulation of nervous system development</i>	7 of 488	0,00061
GO:0030154	<i>Cell differentiation</i>	16 of 3457	0,00068
GO:0045666	<i>Positive regulation of neuron differentiation</i>	6 of 337	0,00070
GO:0016358	<i>Dendrite development</i>	4 of 90	0,00073
GO:0051276	<i>Chromosome organization</i>	9 of 999	0,00074
Molecular Function (GO terms)			
<i>GO term</i>	<i>Description</i>	<i>Count in gene set</i>	<i>False discovery rate (FDR)</i>
GO:0047485	<i>Protein N-terminus binding</i>	4 of 109	0,0063
GO:0003682	<i>Chromatin binding</i>	6 of 401	0,0126
GO:0140110	<i>Transcription regulator activity</i>	11 of 2069	0,0126
GO:0000981	<i>DNA-binding transcription factor activity, RNA polymerase II-specific</i>	9 of 1631	0,0232
GO:0001228	<i>DNA-binding transcription activator activity, RNA polymerase II-specific</i>	5 of 408	0,0232
GO:0003677	<i>DNA binding</i>	11 of 2457	0,0232
GO:0017075	<i>Syntaxin-1 binding</i>	2 of 20	0,0232
GO:0019904	<i>Protein domain specific binding</i>	6 of 706	0,0232
GO:0043565	<i>Sequence-specific DNA binding</i>	7 of 1047	0,0232
GO:0003712	<i>Transcription coregulator activity</i>	5 of 534	0,0258
Cellular Component (Top 20 GO terms)			
<i>GO term</i>	<i>Description</i>	<i>Count in gene set</i>	<i>False discovery rate (FDR)</i>
GO:0045202	<i>Synapse</i>	9 of 849	0,00089
GO:0044456	<i>Synapse part</i>	8 of 705	0,0011
GO:0097458	<i>Neuron part</i>	10 of 1449	0,0032
GO:0098794	<i>Postsynapse</i>	6 of 435	0,0032
GO:0005694	<i>Chromosome</i>	8 of 950	0,0036
GO:0043005	<i>Neuron projection</i>	8 of 1142	0,0085
GO:0044427	<i>Chromosomal part</i>	7 of 819	0,0085
GO:0044428	<i>Nuclear part</i>	16 of 4539	0,0085
GO:0031981	<i>Nuclear lumen</i>	15 of 4030	0,0086
GO:0043232	<i>Intracellular non-membrane-bounded organelle</i>	15 of 4005	0,0086
GO:0005634	<i>Nucleus</i>	20 of 6892	0,0096
GO:0000785	<i>Chromatin</i>	5 of 489	0,0164
GO:0005654	<i>Nucleoplasm</i>	13 of 3446	0,0170
GO:0030424	<i>Axon</i>	5 of 530	0,0196
GO:0071339	<i>MLL1 complex</i>	2 of 31	0,0196
GO:0000790	<i>Nuclear chromatin</i>	4 of 333	0,0217
GO:0033267	<i>Axon part</i>	4 of 341	0,0217
GO:0070013	<i>Intracellular organelle lumen</i>	16 of 5162	0,0217
GO:0098793	<i>Presynapse</i>	4 of 354	0,0227
GO:0044464	<i>Cell part</i>	30 of 16244	0,0399

A més, el programa STRING també realitza un anàlisi de les funcions enriquides en la xarxa, fent servir termes GO (Taula 16). Els termes GO enriquits en la nostra xarxa estan relacionats amb el sistema nerviós (*Biological Process*), unió a l'extrem N-terminal de proteïnes (*Molecular Function*) i sinapsis (*Cellular Component*). Per tant, tot i no haver-hi interaccions directes amb *UBE3A* i haver-hi poques interaccions entre els gens identificats en la nostra cohort i els gens del diagnòstic diferencial de la SA, estarien implicats en les mateixes funcions o vies moleculars a nivell biològic.

D'altra banda, també s'ha utilitzat el programa GeneMANIA per estudiar les interaccions entre els gens identificats en la nostra cohort, els del diagnòstic diferencial de la SA i el gen *UBE3A*. Els resultats mostren que en la xarxa hi ha un 33,62% de gens que es co-expressen, un 29,72% d'interaccions físiques, un 17,37% de co-localització, un 8,58% d'interaccions predites, un 8,32% d'interaccions en la mateixa via molecular, 1,80% de dominis compartits i 0,63% d'interaccions a nivell genètic (Figura 40). A més, GeneMANIA també realitza un anàlisi de les funcions enriquides en la xarxa, en aquest cas la principal funció enriquida és la regulació de la transmissió sinàptica, tot i que també hi ha altres funcions no estadísticament significatives relacionades amb neurogènesi, sistema nerviós i regulació de la transcripció (Taula 17).

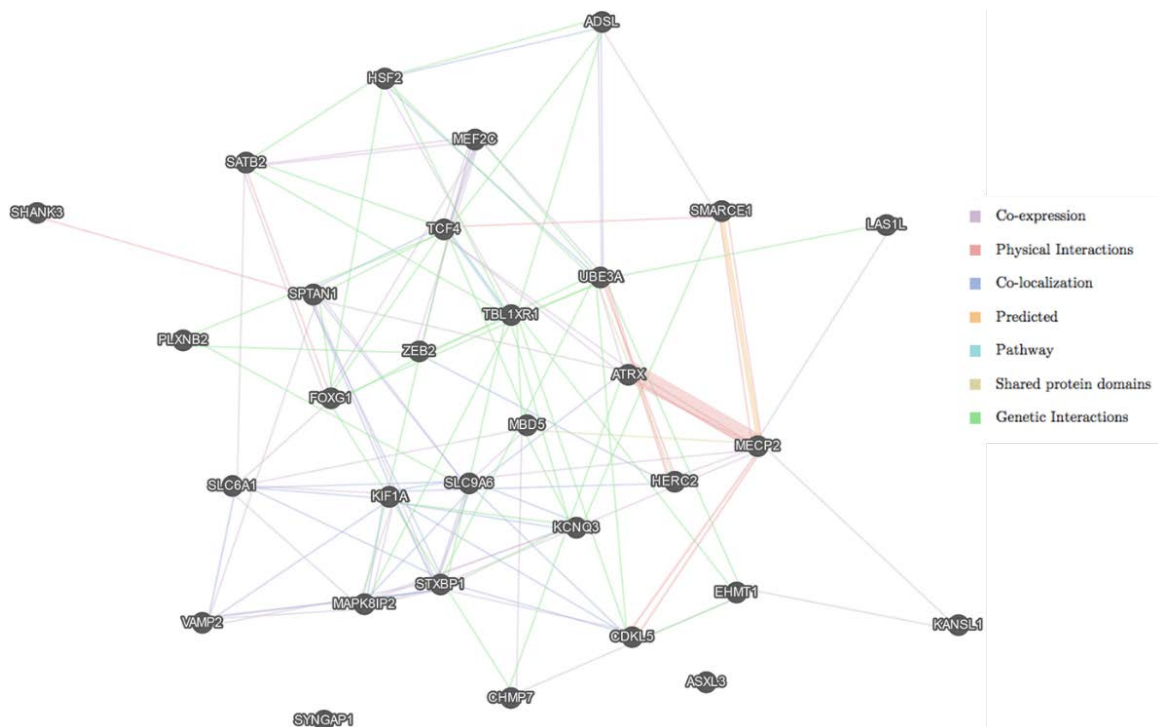


Figura 40 | Xarxa d'interacció dels nous gens identificats, els gens del diagnòstic diferencial de la SA i el gen *UBE3A* obtinguda amb el programa GeneMANIA.

Comparant els resultats obtinguts amb els programes STRING i GeneMANIA, s'observa que els gens identificats en la nostra cohort, els del diagnòstic diferencial de la SA i el gen *UBE3A* no interaccionarien directament, però si que serien gens que es trobarien co-expressats i que portarien a terme funcions comuns implicades en el neurodesenvolupament com la sinapsis, remodelació de la cromatina i regulació de la transcripció.

Taula 17 | Enriquiment funcional de la xarxa dels gens identificats en la nostra cohort, els del diagnòstic diferencial de la SA i el gen *UBE3A* obtingut amb el programa GeneMANIA. Es mostren les funcions ordenades segons el valor del *False Discovery Rate* (FDR).

Function	FDR	Genes in network	Genes in genome
<i>Regulation of synaptic transmission</i>	0,00168	5	71
<i>Single-organism behavior</i>	0,12392	4	93
<i>Neuron-neuron synaptic transmission</i>	0,12392	3	33
<i>Regulation of neurogenesis</i>	0,13314	5	228
<i>Regulation of nervous system development</i>	0,18516	5	256
<i>Learning or memory</i>	0,20592	3	57
<i>RNA polymerase II regulatory region DNA binding</i>	0,20592	3	60
<i>Neuron part</i>	0,20592	5	298
<i>Dendrite morphogenesis</i>	0,20592	3	52
<i>Neurotransmitter secretion</i>	0,20592	3	60

Finalment, utilitzant la base de dades Reactome s'ha analitzat les vies moleculars enriquides. Els resultats obtinguts mostren que entre les 25 vies moleculars estadísticament significatives amb un p-valor <0,05, estan incloses, vies implicades en sinapsis "*GABA synthesis, release, reuptake and degradation*" (p-valor 1,27e-09) i "*Neurotransmitter release cycle*" (p-valor 0,003); transcripció "*Generic Transcription pathway*" (p-valor 1,32e-04), "*RNA polymerase II Transcription*" (p-valor 3,17e-04) i vies implicades en el sistema nerviós "*Neuronal System*" (p-valor 0,002)(Taula 18).

Per tant, els gens identificats en la nostra cohort com a responsables de la SA-like estarien implicats en les mateixes vies moleculars que els gens ja descrits del diagnòstic diferencial de la SA i el gen *UBE3A*. Aquest fet explicaria perquè variants patogèniques/probablement patogèniques en algun d'aquests gens donen lloc a fenotips amb característiques clíniques solapants amb la SA.

4. RESULTATS

Taula 18 | Resultats obtinguts de l'anàlisi de les vies moleculars enriquides dels gens identificats en la nostra cohort, els del diagnòstic diferencial de la SA i el gen *UBE3A* amb la base de dades Reactome. La taula mostra les 25 vies moleculars estadísticament significatives, ordenades segons el p-valor. * *False discovery Rate*.

Pathway name	Entities			Reactions		
	Found	Ratio	p-value	FDR*	Found	Ratio
<i>Transcriptional Regulation by MECP2</i>	9/100	0,007	7,24e-12	2,99e-09	76/77	0,006
<i>Regulation of MECP2 expression and activity</i>	7/39	0,003	1,64e-11	2,26e-09	13/14	0,001
<i>MECP2 regulates transcription factors</i>	4/10	7,02e-04	1,96e-08	1,80e-06	8/8	6,50e-04
<i>Loss of MECP2 binding ability to the NCoR/SMRT complex</i>	3/8	5,61e-04	1,59e-06	1,10e-04	1/1	8,12e-05
<i>Loss of function of MECP2 in Rett syndrome</i>	3/16	0,001	1,26e-05	5,67e-04	5/5	4,06e-04
<i>Prevasive developmental disorders</i>	3/16	0,001	1,26e-05	5,67e-04	1/1	8,12e-05
<i>Loss of MECP2 binding ability to 5hmc-DNA</i>	2/2	1,40e-04	1,46e-05	5,67e-04	1/1	8,12e-05
<i>Loss of MECP2 binding ability to 5mc-DNA</i>	2/5	3,51e-04	9,05e-05	0,003	2/2	1,62e-04
<i>GABA synthesis, release, reuptake and degradation</i>	3/35	0,002	1,27e-04	0,003	3/10	8,12e-04
<i>MECP2 regulates transcription of genes involved in GABA signaling</i>	2/6	4,21e-04	1,30e-04	0,003	4/4	3,25e-04
<i>Generic Transcription Pathway</i>	13/1525	0,107	1,32e-04	0,003	105/811	0,066
<i>Loss of phosphorylation of MECP2 at T308</i>	2/7	4,91e-04	1,77e-04	0,004	1/1	8,12e-05
<i>RNA Polymerase II Transcription</i>	13/1664	0,117	3,17e-04	0,007	105/872	0,071
<i>MECP2 regulates transcription of neuronal ligands</i>	2/13	9,12e-04	6,03e-04	0,011	8/8	6,50e-04
<i>Gene expression (Transcription)</i>	13/1822	0,128	7,68e-04	0,014	107/983	0,08
<i>Serotonin Neurotransmitter Release Cycle</i>	2/23	0,002	0,002	0,032	2/4	3,25e-04
<i>Neuronal System</i>	6/499	0,035	0,002	0,035	21/213	0,017
<i>Acetylcholine Neurotransmitter Release Cycle</i>	2/26	0,002	0,002	0,035	3/6	4,87e-04
<i>Neurotransmitter release cycle</i>	3/99	0,007	0,003	0,035	15/37	0,003
<i>Dopamine Neurotransmitter Release Cycle</i>	2/28	0,002	0,003	0,035	2/5	4,06e-04
<i>Norepinephrine Neurotransmitter Release Cycle</i>	2/30	0,002	0,003	0,039	2/6	4,87e-04
<i>MECP2 regulates neuronal receptors and channels</i>	2/32	0,002	0,004	0,039	26/26	0,002
<i>Glutamate Neurotransmitter Release Cycle</i>	2/32	0,002	0,004	0,039	3/8	6,50e-04
<i>Myogenesis</i>	2/32	0,002	0,004	0,039	3/14	0,001
<i>Interaction between L1 and Ankyrins</i>	2/33	0,002	0,004	0,041	2/14	3,25e-04

Annex II. Resum de les variants identificades en els pacients SA-like

Taula 19| Variants patogèniques/probablement patogèniques responsables del fenotip SA-like.

Pacient	Gen	Isoforma	Canvi de nucleòtid	Canvi d'aminoàcid	Tipus de variant	Herència	Classificació Varsome	Descrita anteriorment
1	KIF1A	NM_001244008.1	c.506G>C	p.Arg169Thr	Missense	De novo	Patogènica	No
2	VAMP2	NM_014232.2	c.128_130delTTGG	p.Val43del	In-frame	De novo	Patogènica	Salprietto et al., 2019
3	SYNGAP1	NM_006772.2	c.1861C>T	p.Arg621*	Nonsense	De novo	Patogènica	No
4	TBL1XR1	NM_024665.5	c.1000T>C	p.Cys334Arg	Missense	De novo	Probablement patogènica	No
5	TBL1XR1	NM_024665.5	c.1043A>G	p.His348Arg	Missense	De novo	Probablement patogènica	No
6	SATB2	NM_001172509.1	c.1826delA	p.Asp609Alafs*15	Frameshift	De novo	Patogènica	No
7	KCNQ3	NM_004519.3	c.688C>T	p.Arg230Cys	Missense	De novo	Patogènica	Decipher; Miceli F et al., 2015; Sands TT et al., 2019
8	SMARCE1	NM_003079.4	c.237+1G>T	p.Ala53_Lys79del	Splicing	De novo	Probablement patogènica	No
9	SPTAN1	NM_001130438.2	c.6592_6597dupCTGCAG	p.Leu2198_Gln2199dup	In-frame	De novo	Probablement patogènica	No
10	ASXL3	NM_030632.2	c.3106C>T	p.Arg1036*	Nonsense	De novo	Patogènica	Kuechler A et al., 2016
11	LAS1L	NM_031206.4	c.1237G>A	p.Gly413Arg	Missense	Lligada al cr.X	Probablement patogènica	No
16	SLC6A1	NM_003042.3	c.889G>A	p.Gly297Arg	Missense	De novo	Patogènica	Carvill GL et al., 2015

Taula 20] Variants de significat incert candidates o modificadores del fenotip SA-like.

Pacient	Gen	Isoforma	Canvi de nucleòtid	Canvi d'aminoàcid	Tipus de variant	Programes de predicció <i>in silico</i>						Funció del gen
						PV	MA	MP	PSNP	SNP&GO	PP2	
12	COX7B	NM_001866.3	c.220C>A	p.Pro74Thr	Missense	D (-5,23)	M (2,64)	0,67	D (76%)	D	Codifica per la isoforma β de la tubulina, implicat en la guia dels axons neuronals	
		TUBB2B	c.751C>A	p.Arg251Ser	Missense							B (0,31)
13	HSF2	NM_004506.3	c.456_459delTTGAG	p.Ser152Argfs*40	Frameshift	NA	NA	NA	NA	NA	Pertany a la família dels factors <i>heat shock</i> que estan implicats en la resposta cel·lular a l'estrés	
		NM_018188.5	c.1301G>A	p.Arg434Gln	Missense	PD (0,95)	M (1,98)	0,66	N (65%)	N	Proteïna de la membrana mitocondrial implicada en el metabolisme mitocondrial, replicació de l'ADN mitocondrial i en el creixement cel·lular	
	ATAD3A	chr1:1454063-1454505	p.Val222Serfs*2	CNV	NA	NA	NA	NA	NA	NA		
14	RMND1	NM_017909.4	c.1286A>T	p.Glu429Val	Missense	PD (1)	M (3,3)	0,91	D (87%)	N	Es localitza en la membrana mitocondrial i es troba implicat en la traducció mitocondrial	
			c.1049T>C	p.Met350Thr	Missense	PD (0,96)	M (2,4)	0,73	N (60%)	N		
15	CSTF2	NM_001306206.1	c.32T>G	p.Val11Gly	Missense	B (0,18)	L (0,16)	0,43	N (63%)	N	Intervé en la polidenciació i decapació del pre-ARNm	
			chr2:228243643-228244178	p.Ala611Ilefs*12	CNV	NA	NA	NA	NA	NA	NA	Proteïna transmembrana del RE regulada per ceràmides i implicada en la translocació de proteïnes durant la traducció
17	CHMP7	NM_152272.5	c.1165C>T	p.Gln389*	Nonsense	NA	NA	NA	NA	NA	Pertany a la família dels complexos ESCRT* implicada en senyalització cel·lular, autofàgia, migració i transport d'ARNm	

PP2, PolyPhen2; PV, Provean; MA, Mutation Assessor; MP, MutPred; PSNP, PredictSNP; PsD, possibly damaging; PD, probably damaging; D, deleterious; N, neutral; M, Medium; NA, no aplica; RE, reticle endoplasmàtic. En negreta es troben marcadetes les variants candidates i en gris les variants modificadores.



Discussió

La SA és un trastorn neurogenètic que es caracteritza per DI greu, absència de llenguatge, atàxia, microcefàlia, epilèpsia amb un patró específic en l'EEG i un comportament característic que inclou aparença feliç, estereotípies, hiperactivitat i dèficit d'atenció. En el 10% dels casos es desconeix la causa a nivell molecular i aquests pacients es classifiquen com a *SA-like* (36,39,53). L'objectiu principal d'aquesta tesi doctoral ha estat la identificació de nous gens implicats en la *SA-like* en 17 pacients amb un fenotip SA i sense diagnòstic molecular a través de la seqüenciació de l'exoma complet (WES) en trios.

La introducció de les noves tecnologies de seqüenciació, en concret, el WES ha facilitat la identificació de nous gens associats a malalties monogèniques (261,262). Un bon exemple és el projecte DDD (<https://www.ddduk.org/>), en el que l'aplicació de WES en trios en una cohort d'individus amb trastorns del desenvolupament (TD) ha permès diagnosticar aproximadament un 42% dels individus, així com identificar 85 nous gens associats a aquests trastorns (140–142). Un altre exemple que il·lustra l'impacte de la implementació de WES, és la base de dades SysID (<https://sysid.cmbi.umcn.nl/>) que ha passat de tenir 777 gens associats a DI i 389 gens candidats al desembre de l'any 2015 a tenir-ne 1291 gens associats a DI i 1140 gens candidats al desembre de l'any 2019. Segons l'estudi portat a terme per Co et al., 2018 on es va realitzar un meta-anàlisi de les variants *de novo* identificades per exoma en pacients amb TEA, DI i/o retard del desenvolupament, s'estima que la identificació de nous gens implicats en TND amb un excés de variants *de novo* truncants i *missense* que actuen per un mecanisme de pèrdua de funció, arribarà properament a una fase de *plateau*. De tota manera, es considera que les variants *missense* que actuen per mecanismes diferents del de pèrdua de funció representen el reservori més prometedor per al futur descobriment de nous gens, fet que requerirà un major volum de dades obtingudes a través de l'aplicació de WES i WGS (263).

Des del punt de vista de la pràctica clínica, l'aplicació de WES ha demostrat ser una eina efectiva pel diagnòstic de trastorns mendelians (138), en concret, el WES en trios pacient-progenitors és particularment eficaç en els casos en els quals els progenitors són sans i es sospita d'una malaltia amb una herència *de novo* o autosòmica recessiva. En el cas dels TND, on hi ha una gran heterogeneïtat genètica, es recomana realitzar WES com a primera eina diagnòstica, en lloc d'un panell de gens (264,265). Pel que fa al cost normalment s'assumeix que els panells de gens tenen un cost menor que el WES, però s'ha de tenir en compte que els panells de gens requereixen actualitzacions, per incorporar els nous gens que es van identificant associats a malaltia. A més, el re-anàlisi de panells de gens en pacients sense diagnòstic té un ús limitat. En canvi, el re-anàlisi de l'exoma reportat en diferents estudis de pacients amb sospita de malalties monogèniques i en

el projecte DDD, ha demostrat que la taxa diagnòstica pot arribar a augmentar fins a un 10% (127–129,141). Per tant, molt probablement el millor mètode cost-efectiu és el WES.

En la nostra cohort, l'aplicació de WES ha permès la identificació de 12 variants patogèniques/probablement patogèniques responsables del fenotip *SA-like* localitzades en 11 gens prèviament relacionats amb TND (*KIF1A*, *VAMP2*, *SYNGAP1*, *TBL1XR1*, *SATB2*, *KCNQ3*, *SMARCE1*, *SPTAN1*, *ASXL3*, *SLC6A1* i *LAS1L*) (145,212,238,266–272) (Resultats, Annex II, Taula 17). La taxa de diagnòstic en la nostra cohort ha estat del 70,5% (12/17). Aquesta taxa de diagnòstic és superior a la reportada en la literatura (24–68%) per DI, retard del desenvolupament, TEA i anomalies congènites (265). El motiu és que la DI greu es produeix majoritàriament per variants rares (tant puntuals com estructurals) *de novo*, trobant-se la causa molecular en més del 50% dels casos (138,181), mentre que en altres entitats com la DI lleu o el TEA, que tenen una base poligènica causada per la suma de variants genètiques (tant puntuals com estructurals) amb penetrància i expressivitat variables, la taxa diagnòstica és menor (273,274).

El 91,6% (11/12) de les variants identificades en la nostra cohort són autosòmiques dominants *de novo* i en un dels casos (8,3%, 1/12) la variant és lligada al cromosoma X (gen *LAS1L*) heretada de la mare, que és portadora en mosaic (20%). Aquests resultats estan en consonància amb els reportats en la literatura, en concret, el projecte DDD estima que un 80% de les variants identificades en individus amb TD són *de novo* (142).

Entre les variants identificades en la nostra cohort es troben: dues *nonsense*, una *frameshift*, una de *splicing*, dues *in-frame* i sis *missense*. Cal destacar, que només el 25% (3/12) de les variants trunquen la proteïna i que el 50% (6/12) de les variants patogèniques/probablement patogèniques són *missense*. En altres cohorts també s'ha reportat aquest fet, així en el projecte DDD aproximadament un 50% de les variants puntuals són *missense* i un 50% són truncants (140). Un altre exemple, és la cohort d'individus amb DI greu descrita per Rauch et al., 2012 on es va aplicar WES. En aquest cas, un 60% de les variants identificades com a responsables del fenotip són *missense* i un 20% són truncants (275).

Mentre que la majoria de variants truncants actuen per un mecanisme de pèrdua de funció, les variants *missense* poden actuar tant per pèrdua de funció, guany de funció o com a dominant negatiu. Les variants *missense de novo* quan actuen per mecanismes de guany de funció o dominant negatiu, es tendeixen a localitzar en clústers de determinats dominis funcionals de la proteïna (142,163,276,277). En la nostra cohort, la variant *missense* en el gen *KCNQ3* actuaria per un mecanisme de guany de funció, essent una variant recurrent identificada en altres pacients i localitzada en un residu important per a la funció de la proteïna.

El WES en trios també ha permès identificar dues variants truncants en dos nous gens candidats al fenotip *SA-like*, *HSF2* i *CHMP7*, no associats prèviament a TND. Ambdues variants en els gens *HSF2* i *CHMP7* activen el mecanisme NMD, el que comportaria una pèrdua de funció pels dos gens. Tant *HSF2* com *CHMP7* són gens altament intolerants a les variants truncants ($pLI > 0,86$), suggerint una funció conservada al llarg de l'evolució per aquests dos gens. El gen *HSF2* s'expressa a alts nivells en cervell i pertany a la família dels factors *heat shock* que estan implicats en la resposta cel·lular a l'estrès. Estudis realitzats en ratolins *knockout* (*hsf2*^{-/-}) han mostrat que presenten defectes en l'espermatogènesis i anàlisis histològics del cervell mostren defectes en el desenvolupament del sistema nerviós central (278,279).

D'altra banda, el gen *CHMP7* pertany a la família dels complexos ESCRT (*Endosomal Sorting Complexes Required for Transport*) implicada en la senyalització cel·lular, autofàgia, migració i transport d'ARNm (280). A nivell neuronal és essencial en la formació de sinapsis, morfogènesis, supervivència i creixement cel·lular (281). La pèrdua o el mal funcionament d'aquests complexos afecta a la neurogenèsi, al neurodesenvolupament i provoca neurodegeneració en animals d'experimentació adults. En concret, la pèrdua del complex ESCRT-II s'ha observat que afecta al neurodesenvolupament (282).

A més a més, s'ha identificat una variant *missense de novo* en el gen *CSTF2*, que recentment s'ha descrit en una família associada a DI lleu no sindròmica (247). Les variants deletèries en gens lligats en el cromosoma X tenen un efecte altament penetrant en homes mentre que en dones donen lloc a un fenotip més lleu o a una absència de fenotip depenent de la inactivació del cromosoma X (283,284). En contraposició, alguns casos de noies portadores de variants de pèrdua de funció *de novo* presenten un fenotip més greu del que s'esperaria per una herència lligada al cromosoma X (p.e, gen *IQSEC2*). En aquests casos, en les noies les variants són sempre *de novo* mentre que els nois presenten variants *missense* heretades de mares portadores sense fenotip. En el cas de la variant p.Val11Gly en el gen *CSTF2*, no es pot descartar que es tracti d'una variant amb un efecte deleteri i que aquesta variant sigui la causa del seu fenotip.

La identificació d'altres pacients portadors de variants en els gens *HSF2*, *CHMP7* i *CSTF2* amb un fenotip similar, reforçaria la implicació d'aquests gens com a responsables del fenotip *SA-like* en els nostres pacients. En aquest sentit, les tres variants s'han introduït en l'eina *Genematcher*, que permet intercanviar dades genotípiques i fenotípiques entre diferents grups mèdics o científics amb l'objectiu d'identificar nous gens causants de malalties humanes (285,286). Fins al moment no tenim coneixement d'altres individus amb variants en els gens *HSF2* i *CHMP7* que presentin un fenotip similar. No obstant, el fet que en les dues variants s'hagi demostrat que activen el

mecanisme NMD indica que produirien una pèrdua de funció i afegeix evidència al seu paper com a responsables del fenotip SA.

Pel que fa a la variant Val11Gly en el gen *CSTF2* s'ha contactat amb un altre investigador que té un pacient amb una variant en aquest gen amb un fenotip de DI greu i absència de parla. Es considera que la variant p.Val11Gly en el gen *CSTF2* és una bona candidata malgrat que en l'actualitat no hi hagin estudis funcionals ni suficients pacients que ens confirmin la implicació de *CSTF2* en malaltia.

De manera paral·lela a l'anàlisi del model d'herència *de novo* per identificar variants patogèniques/probablement patogèniques responsables del fenotip SA-like, també s'ha analitzat la càrrega genètica de variants *de novo* en aquests pacients. Els resultats mostren que un 50% (8/16) dels pacients amb fenotip SA-like presenten més d'una variant amb un efecte deleteri sobre la proteïna. Aquest fet també s'ha observat en altres cohorts amb TND, com per exemple, en la síndrome de Rett-like on un 52,6% dels pacients presentaven més d'una variant deletèria (244). En el cas de l'autisme la freqüència és menor, un 6,3% dels pacients tenien més d'una variant amb un efecte deleteri, suggerint un model poligènic com un dels mecanismes responsables d'aquest TND (246). En la nostra cohort, la majoria dels pacients amb diagnòstic molecular eren els que presentaven més d'una variant amb un efecte deleteri, les quals podrien estar contribuint o modulant el fenotip. Només en un cas (pacient 13) s'han identificat dues variants amb un efecte deleteri en dos gens que no han estat prèviament associats a TND (*METTL24* i *HSF2*).

5.1 Anàlisi de la patogenicitat de les variants identificades

En relació a les 12 variants identificades com a responsables del fenotip SA-like, set variants s'han classificat com a patogèniques seguint la guia ACMG (156). D'aquestes, tres variants són truncats, en gens (*SYNGAP1*, *SATB2* i *ASXL3*) on s'ha reportat que variants de pèrdua de funció (haploinsuficiència) és un mecanisme causant de malaltia (287–289). Una de les variants, la identificada en el gen *ASXL3*, ja havia estat descrita amb anterioritat en quatre pacients (289–291). Pel que fa a les variants *missense* i *in-frame*, les quatre variants localitzades en els gens *SLC6A1*, *KCNQ3*, *KIF1A* i *VAMP2* s'han classificat com a patogèniques, considerant que: i) les mateixes variants han estat identificades en altres pacients amb un fenotip similar (269,292,293) i ii) en les variants identificades en els gens *KIF1A* (Resultats, Capítol 6) i *KCNQ3* (269,294), s'han realitzat estudis funcionals que confirmen el seu efecte deleteri sobre la proteïna.

Per la resta de variants *missense* i *in-frame* és difícil de predir la seva patogenicitat. Malgrat que els predictors *in silico* poden ajudar a la seva classificació, és necessari realitzar estudis funcionals que demostrin l'efecte deleteri sobre la funció de la proteïna. Per tant, les tres variants *missense* (les dues variants en el gen *TBL1XR1* i la variant en el gen *LAS1L*) i les dues *in-frame* (gens *SPTAN1* i *SMARCE1*) responsables del fenotip *SA-like* restants, s'han classificat com a probablement patogèniques.

Finalment, les dues variants truncants identificades en els gens candidats al fenotip *SA-like* (*HSF2* i *CHMP7*) s'han classificat com a variants en gens de significat incert. No hi ha suficients evidències que demostrin un paper causal d'aquests dos gens en malaltia (126). La variant *missense* en el gen *CSTF2* s'ha classificat com una variant de significat incert atès que només hi ha una variant associada a DI lleu no sindròmica en una família amb una herència lligada al cromosoma X que afecta només als homes i en un cas amb DI greu.

Variants truncants patogèniques

Per les variants truncants en els gens *SYNGAP1*, *SATB2* i *ASXL3*, s'ha volgut demostrar que produeixen una haploinsuficiència. En els tres casos, s'ha descrit que el principal mecanisme de patogenicitat és la pèrdua de funció, ja que la majoria de les variants identificades en aquests dos gens són truncants (287,288). Els estudis realitzats en ARNm per determinar si els transcrits portadors d'aquestes variants es degraden pel mecanisme NMD mostren que *SYNGAP1* i *SATB2* escapen aquest mecanisme. Podria explicar-se per un mecanisme de NMD incomplet en limfòcits com ha descrit Magyar et al., 2009 (295), suggerint que seria convenient analitzar un altre teixit, com fibroblasts per determinar si hi ha degradació de l'al·lel amb la variant truncant (295). D'altra banda, estudis realitzats de dues variants *missense* (W362R i P562L) i de la variant truncant R579X en el gen *SYNGAP1* en línies cel·lulars HEK293T i N7, han demostrat que no hi ha expressió de proteïna mutant, suggerint que variants a *SYNGAP1* donarien lloc a proteïnes inestables (287). D'aquestes tres variants es van realitzar també estudis funcionals en neurones piramidals corticals de cultius organotípics que van demostrar un augment de la fosforilació de ERK en les neurones transfectades amb les variants mutants comparat amb la proteïna salvatge (degut a la falta d'activitat GTPasa de *SYNGAP1*). Per tant, es suggereix que les variants *missense* i truncants a *SYNGAP1* actuarien per pèrdua de funció bé perquè afecten l'estabilitat de la proteïna o bé perquè interfereixen en la seva activitat GTPasa.

En el cas de la variant p.Arg621* en el gen *SYNGAP1*, identificada en el nostre pacient, es produiria una proteïna truncada de 620 aminoàcids (la proteïna salvatge té 1343 aminoàcids) que conservaria els dominis PH i C2 íntegres, mentre que el domini RasGAP, essencial pel correcte

funcionament de la proteïna, quedaria truncat (Figura 41). Per tant, basant-nos en els estudis funcionals descrits anteriorment es proposa que el mecanisme de patogenicitat pel qual actuaria la variant p.Arg621* és de pèrdua de funció bé per falta de proteïna (proteïna inestable) o per l'afectació de la seva activitat GTPasa.

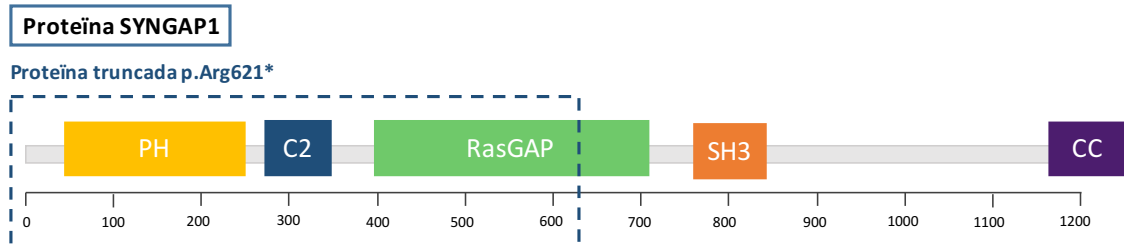


Figura 41 | Estructura de la proteïna SYNGAP1 on es troben indicats els diferents dominis de la proteïna. El requadre amb la línia discontinua correspon a la proteïna truncada que es produeix degut a la introducció d'un codó stop prematur per la variant p.Arg621*. PH, *pleckstrin homology*, C2 *domain essential for the stimulation of GTPase reaction*; RasGAP, *catalytic domain RasGAP*; SH3, *SRC Homology 3 Domain*; CC, *coiled-coil*.

Respecte al gen *SATB2*, s'ha descrit que variants localitzades en els dos últims exons d'aquest gen escapen al mecanisme NMD, donant lloc a proteïnes més curtes, però que retenen algunes funcions (288). Estudis funcionals amb assajos luciferasa de variants que escapen al mecanisme NMD, com la variant c.715C>T situada en l'exó 8, que manté el domini de dimerització (regió UDL), però li falten els dominis CUT1, CUT2 i HOX (dominis d'unió a l'ADN), mostren que la proteïna truncada afecta a la funció de la proteïna salvatge, suggerint que podria tenir un efecte dominant negatiu (296). La variant c.1826delA identificada en la nostra cohort, resulta en una proteïna truncada (p.Asp609Alafs*15) més curta (622 aminoàcids), però que retindria els dominis CUT1 i CUT2 intactes i només estaria afectat el domini d'unió a l'ADN HOX (Figura 42). La variant c.1826delA, no actuaria per un mecanisme de pèrdua de funció, sinó que el mecanisme de patogenicitat en aquest cas seria dominant negatiu, d'acord amb els resultats dels estudis funcionals obtinguts d'altres variants que escapen el mecanisme NMD.

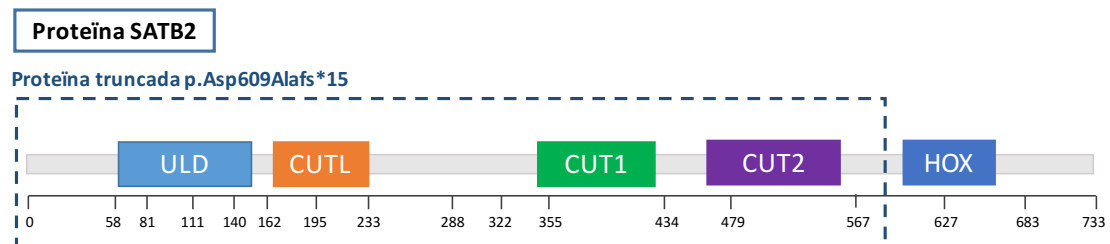


Figura 42 | Estructura de la proteïna SATB2 on es troben indicats els diferents dominis de la proteïna. El requadre amb la línia discontinua correspon a la proteïna truncada que es produeix degut a la introducció d'un codó stop prematur per la variant p.Asp609Alafs*15. ULD, *Ubiquitin-like domain*; CUTL, *Cut-like domain*; HOX, *Homeobox domain*. El domini ULD està implicat en l'oligomerització i CUTL, CUT1, CUT2 i HOX són dominis d'unió a l'ADN.

Finalment, la variant c.3106C>T, p.Arg1036* identificada en *ASXL3* està localitzada en l'últim exó i per tant no activaria NMD. Malauradament, en aquest cas no es va poder determinar la degradació de l'ARNm perquè aquest gen s'expressa principalment en cervell i molt poc o no s'expressa en sang. Es necessitaria mostra d'un altre teixit que derivi del neuroectoderm per comprovar si la variant c.3106C>T escapa al mecanisme NMD (Figura 43).

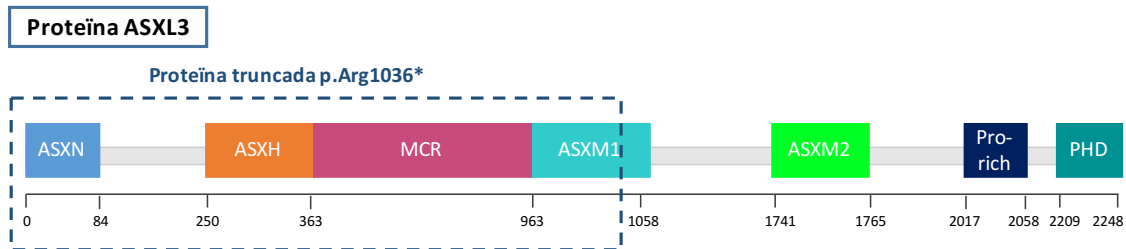


Figura 43 | Estructura de la proteïna ASXL3 on es troben indicats els diferents dominis de la proteïna. El requadre amb la línia discontinua correspon a la proteïna truncada que es produeix degut a la introducció d'un codó stop prematur per la variant p.Arg1036*. ASXN, DNA binding domain; ASXH, protein-protein interaction domain; MCR, mutation cluster region; ASXM2 i ASXM1, protein-protein interaction domains; Pro-rich, Proline-rich domain; PHD, Plant Homeodomain.

En el cas de què es produís proteïna, la proteïna truncada resultant tindria 1035 aminoàcids i conservaria intactes el domini ASXN d'unió a l'ADN, el domini ASXH que interacciona a nivell proteïna-proteïna amb altres reguladors epigenètics, la regió MCR (*Mutation Cluster Region*) i part del domini ASXM1, que juntament amb ASXM2, formen un mòdul implicat en interaccions proteïna-proteïna que s'associa amb receptors nuclears d'hormones i els seus cofactors. Mancarien els dominis ASXM2 i PHD que intervenen en la modificació post-transcripcional d'histones (297). Una possible hipòtesi és que la proteïna truncada p.Arg1036* podria actuar com a dominant negatiu, ja què s'ha suggerit que variants truncants situades en la regió dels dominis ASXN i ASXH en l'extrem N-terminal i els dominis ASXM2 i PHD en l'extrem C-terminal actuarien per aquest mecanisme. En aquestes regions es troben llocs de fosforilació i motius rics en serina altament conservats. S'especula que l'expressió de la proteïna ASXL3 amb aquests dominis i regions truncades afectaria a les interaccions i a la funció de remodelador de la cromatina de la proteïna salvatge, resultant en una desregulació de la modificació post-traducciona d'histones i en l'activació constitutiva dels gens regulats per ASXL3 (298).

L'únic estudi funcional realitzat en aquest gen de la variant *frameshift* c.1448dupT en fibroblasts, indica que no hi ha expressió de l'al·lel amb la variant. Per tant, el transcrit amb la variant és subjecte de degradació pel mecanisme NMD, sent compatible amb una pèrdua de funció. Com a conseqüència, s'observa un augment dels nivells de la histona H2A mono-ubiquitinada i canvis en la transcripció de 564 gens (299).

Al no conèixer si la variant c.3106C>T es degrada per NMD no podem determinar si aquesta actua com a dominant negatiu o per pèrdua de funció. En ambdós casos, alteraria l'expressió de gens regulats per ASXL3.

Variants missense i in-frame patològiques

De les tres variants *missense* i la variant *in-frame* classificades com a patològiques només s'han realitzat estudis funcionals en els gens *KIF1A* i *KCNQ3*, mentre que de les variants localitzades en els gens *VAMP2* i *SLC6A1* s'han portat a terme anàlisis *in silico* que recolzarien la seva patogenicitat.

Variants en el gen *KCNQ3* que actuen per un mecanisme de guany de funció, com seria la variant p.Arg230Cys, s'associen a autisme i TND (269). La variant p.Arg230Cys es troba situada en el segment transmembrana S4, concretament, en el residu arginina R2 de la subunitat codificada per *KCNQ3* del canal potassi dependent de voltatge tipus M. Estudis funcionals d'electrofisiologia han demostrat que l'arginina 230, amb càrrega positiva, estableix la forma oberta del canal de potassi, provocant una hiperpolarització de les neurones (Figura 44) (269,294). Cal destacar, que variants truncants que actuen per un mecanisme de pèrdua de funció en aquest mateix gen són responsables de l'epilèpsia neonatal familiar benigna (BFNE, MIM # 121201) i l'epilèpsia infantil familiar benigna (BFIE), mostrant que diferents tipus de variants amb efectes oposats sobre la proteïna (guany i pèrdua de funció) donen lloc a dos fenotips molt diferents (300,301).

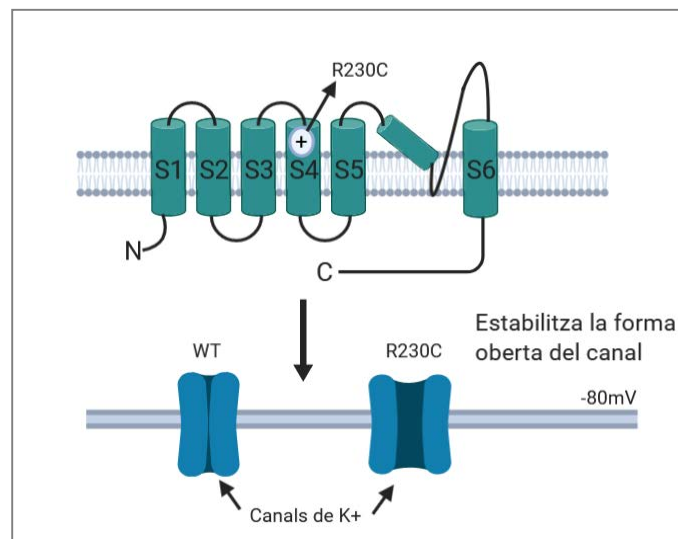


Figura 44 | Estructura de la subunitat del canal de potassi codificada per *KCNQ3* i com la variant R230C estableix la forma oberta del canal a voltatges negatius.

La variant p.Arg169Thr (R169T) en el gen *KIF1A* no ha estat descrita prèviament i es troba localitzada en el domini motor. Inicialment es va classificar com a probablement patològica, ja

què no es troba en població control, els programes *in silico* prediuen un efecte deleteri i altres variants *missense de novo* localitzades en el domini motor causen un efecte dominant negatiu afectant a la motilitat de la proteïna *wild type* (64). No obstant, per classificar aquesta variant com a patogènica és necessari portar a terme estudis funcionals que demostrin el seu efecte deleteri. L'anàlisi de l'activitat ATPasa del domini motor de KIF1A, ha demostrat que la variant R169T afecta aquesta activitat. L'activitat ATPasa és una mesura indirecta de la motilitat de la proteïna (la proteïna necessita de la hidròlisi d'ATP com a font d'energia per poder moure's) per tant la proteïna amb la variant R169T tindria una motilitat reduïda. Aquesta falta d'activitat ATPasa podria explicar-se per: i) una afectació directa de l'activitat ATPasa del domini motor o ii) una manca d'unió del domini motor als microtúbuls (la unió a microtúbuls indueix l'activitat ATPasa).

Estudis estructurals mostren que el residu Arg169 es troba situat en el *loop* L8, concretament en la fulla 5β, que forma part d'un dels dominis d'unió a microtúbuls (MB1), juntament amb els *loops* 11 i 12 que formen els altres dos dominis d'unió a microtúbuls (MB2 i K-*loop*) (302,303). La mutagènesi de diferents residus dels *loops* L11, L12 i L8 per Ala, entre ells el residu Arg169, provoca una menor afinitat pels microtúbuls quan s'uneix un anàleg no hidrolitzable de l'ATP (AMP-PNP) o quan es troba unit l'ADP (304).

El modelatge estructural de la variant R169T ha estat realitzat en col·laboració amb el grup de *Structural Biology of Cell Division, The University of Edinburgh* del Dr. A. Jeyaprakash Arulanandam. L'anàlisi de l'estructura mostra que la variant R169T té com a conseqüència la pèrdua d'interaccions electrostàtiques (ponts salins) amb els aminoàcids E417 i E420 de la subunitat β de la tubulina (Figura 45). Probablement, interferint negativament en la interacció de KIF1A amb els microtúbuls. Així doncs, la variant R169T en el gen *KIF1A* s'ha classificat finalment com a patogènica, doncs afectaria la funció de la proteïna resultant, com a conseqüència de la seva menor afinitat en la unió als microtúbuls el que provocaria una disminució de la seva motilitat, donant lloc al fenotip present en el pacient. No hem pogut demostrar si existeix un efecte dominant negatiu encara que sembla probable, ja que KIF1A actua com a dímer (198,206).

Convé destacar que *KIF1A* és un altre exemple en què diferents tipus de variants poden causar diferents fenotips. Variants *frameshift* en homozigosi o heterozigosi composta s'associen a la neuropatia sensorial hereditària tipus IIC (HSPN IIC, MIM # 614213) amb una herència autosòmica recessiva. Variants *missense* en el domini motor s'associen a paraplegia espàstica (SPG30, MIM # 610357) tant amb una herència autosòmica recessiva, com amb una herència autosòmica dominant. A més, variants *missense* al domini motor s'associen a una síndrome de DI greu amb

epilèpsia, atrofia del cerebel i del nervi òptic, hipotonia, paraparèsia espàstica, hiperreflèxia, microcefàlia post-natal i neuropatia perifèrica (212,213).



Figura 45 | Modelatge estructural de la variant R169T. En l'ampliació de la imatge esquerra, es pot observar la interacció electrostàtica de R169 amb els aminoàcids E417 i E420 de la subunitat β de la tubulina. En l'ampliació de la dreta es pot observar com la variant R169T provoca la pèrdua de les interaccions electrostàtiques. Imatge generada amb el programa PyMOL utilitzant l'estructura cristal·logràfica de KIF1A (PDB ID: 2HXF).

El fet que variants *missense* dins del domini motor donin lloc a diferents fenotips es podria explicar perquè aquestes afecten de manera diferent a la capacitat de KIF1A de moure's a través dels microtúbuls i per tant, de transportar proteïnes sinàptiques i vesícules. Inicialment, s'havia suggerit que variants que afectessin lleument a la motilitat de la proteïna s'associarien als fenotips de paraplegia espàstica (305). No obstant, estudis funcionals *in vitro* realitzats de les variants V8M, A255V i R350G (associades a paraplegia espàstica hereditària) amb la proteïna KIF1A completa, han demostrat que hi ha un augment de l'activitat motora respecte KIF1A WT, com a conseqüència de la pèrdua del mecanisme autoinhibitori que regula la unió a microtúbuls (198,214). La introducció d'aquestes mateixes variants en l'ortòleg *unc-104* en *Caenorhabditis elegans* mostra un augment en el transport de vesícules amb precursors sinàptics al final dels axons neuronals suggerint que el guany de funció de KIF1A estaria implicat en el fenotip de paraplegia espàstica sense DI (214). Una altra opció, que podria explicar les diferències entre els

fenotips és una possible redundància en l'expressió de diferents quinesines de la mateixa família (306).

Respecte al gen *SLC6A1* que codifica per un transportador de GABA (GAT-1), la variant p.Gly297Arg (G297R) es troba situada en la butxaca on s'uneix GABA. L'aminoàcid Gly297, juntament amb Ala61, Leu300 i Trp400 formen la butxaca d'unió de GABA (307). Variants que introdueixen aminoàcids de cadena llarga amb càrrega positiva, com la variant G297R, és probable que produeixin un canvi en la conformació de la butxaca, no permetent la unió de GABA i causant una pèrdua de funció (Figura 46) (293).

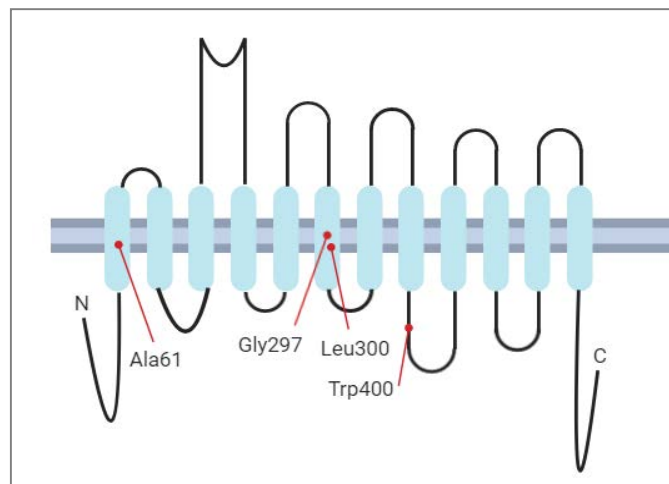


Figura 46 | Estructura del transportador GAT-1 codificat pel gen *SLC6A1*. Els aminoàcids que formen la butxaca on s'uneix el neurotransmissor GABA estan marcats amb una fletxa en vermell.

Estudis funcionals de les variants G94E, F270S, I270del, Y445C, W235R, W496Ter i G550R identificades en el gen *SLC6A1* en una cohort d'individus amb epilèpsia mostren que aquestes variants produeixen una disminució del transport de GABA. Les variants G94E, W496T i G550R aboleixen completament el transport de GABA, mentre que la resta de variants tenen una activitat residual del 2-27%. Un menor transport de GABA té com a conseqüència un augment en els nivells de GABA en les sinapsis, produint una sobreestimulació dels receptors de GABA ($GABA_A$ i $GABA_B$) que són responsables de les respostes inhibidores (tòniques) lentes i sostingudes (308). En dos estudis independents, l'anàlisi funcional de les variants P361T i G243S, mostren que hi ha una reducció de l'expressió total de la proteïna i que per tant, aquestes variants afecten a l'estabilitat de la proteïna. En ambdós casos, també hi ha una reducció en l'activitat de transport de GABA (309,310). En conjunt, aquests estudis recolzen que les variants en el gen *SLC6A1* associades a epilèpsia actuen per un mecanisme de pèrdua de funció.

En últim lloc, el gen *VAMP2* es troba implicat en la transmissió sinàptica, intervenint en la fusió i endocitosi de vesícules i alliberament de neurotransmissors (311–313). Recentment, el gen *VAMP2* ha estat associat a TND. S'han identificat cinc variants *de novo* (S75P, E78A, F77S, V43del i I45del) que es troben localitzades pròximes a l'extrem C-terminal del domini SNARE. L'anàlisi funcional de les variants *missense de novo* S75P i E78A mostren que afecten a la formació del complex SNARE, en concret, en el cas de la variant S75P s'ha demostrat que té un efecte dominant negatiu. El modelatge estructural d'aquestes dues variants (S75P i E78A) mostra que es perden ponts d'hidrogen amb aminoàcids de la proteïna STX1A. Les dues variants *in-frame* identificades en aquest gen V43del (pacient 2 de la nostra cohort) i I45del no han estat estudiades a nivell funcional, però anàlisis *in silico* mostren que els aminoàcids Val43 i Ile45 es troben altament conservats. Els individus amb variants *in-frame* comparats amb els que presenten variants *missense* mostren un fenotip menys greu a nivell neurològic, són capaços de caminar i d'utilitzar algunes paraules (292).

Variants probablement patogèniques

De les variants classificades com a probablement patogèniques no s'han realitzat estudis funcionals, però les prediccions *in silico* i estudis funcionals portats a terme d'altres variants del mateix tipus (*missense* o *in-frame*) en els mateixos gens, indiquen que aquestes actuarien per un mecanisme dominant negatiu o de guany de funció. Variants *missense de novo* en el gen *TBL1XR1* s'han associat a diferents TND (267,314). La variant p.Phe10Leu identificada en un pacient amb esquizofrènia, es van realitzar estudis funcionals i es va comprovar que aquesta variant alterava la via de senyalització de Wnt/ β -catenina, disminuint la interacció de TBL1XR1 amb el complex N-CoR i augmentant la unió de TBL1XR1 amb la β -catenina, resultant en un augment de l'activitat transcripcional de la via (315). S'ha descrit que variants probablement patogèniques en el domini WD40 β -propeller afecten a la rigidesa estructural, compromentent l'estabilitat o les característiques fisicoquímiques de la cara superior i en conseqüència afectant a la unió de la β -catenina (267). Els aminoàcids Cys334 i His348 que són els afectats en les dues variants identificades en la nostra cohort (p.Cys334Arg i p.His348Arg) es troben en aquest domini (Figura 47), per la qual cosa es prediu que podrien tenir un efecte disminuint l'activitat transcripcional de la via de senyalització de Wnt/ β -catenina.

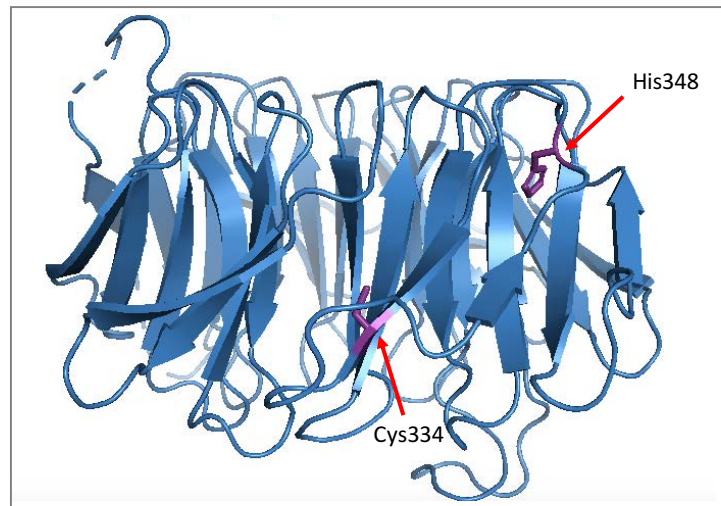


Figura 47 | Estructura de la proteïna codificada pel gen *TBL1XR1*. Marcats amb una fletxa en vermell es troben els aminoàcids Cys334 i His348, que es troben mutats en les pacients 4 i 5 de la nostra cohort. Imatge realitzada amb el programa PyMOL.

La variant *missense* identificada en el gen *LAS1L*, p.Gly314Arg, no ha estat descrita prèviament i segons els predictors *in silico* té un efecte deleteri. Aquest gen està implicat en la biosíntesis del ribosoma, en concret, és necessari per a la síntesi de la subunitat 60S del ribosoma i la maduració de l'ARN ribosòmic 28S. A més, intervé en la regulació de la transcripció i forma part del complex CoREST1/HDAC1 (71–73). Variants *missense* en el gen *LAS1L* s'associen a la síndrome de Wilson-Turner (MIM # 309585). Fins al moment, només s'han identificat dues variants *missense* (p.Ala269Gly i p.Arg415Trp) en el gen *LAS1L* en dues famílies Amish associades a la síndrome de Wilson-Turner (26). Es desconeix quina de les funcions de *LAS1L* es troba afectada per les variants *missense* identificades fins al moment i si actuen per un mecanisme de pèrdua de funció o bé com a dominant negatiu o guany de funció. Les dones portadores no estan afectades, aquest és el cas de la mare que és portadora en mosaic (20%) i no presenta trets clínics de la síndrome de Wilson-Turner. Hi ha descrit una altra variant *missense* p.Ser477Asn en el gen *LAS1L* identificada en un pacient amb una malaltia de la neurona progressiva que curiosament no presentava les característiques clíniques de la síndrome de Wilson-Turner. Estudis funcionals d'aquesta variant realitzats en peix zebra demostren que la pèrdua de funció d'aquest gen dóna lloc a una letalitat primerenca dels embrions i un increment en l'apoptosi en cervell, medulla espinal i musculatura. El fenotip es rescatava parcialment quan s'introduïa la forma salvatge del gen *LAS1L*, mentre que la introducció de la forma mutada p.Ser477Asn no rescatava el fenotip, suggerint que aquesta variant provocaria una pèrdua de funció del gen (74).

Pel que fa a les variants *in-frame*, s'ha descrit que delecions i duplicacions *in-frame* en els dos últims *repeats* (19 i 20) del gen *SPTAN1* tenen un efecte dominant negatiu, afectant la heterodimerització de l' α -espectrina i la β -espectrina i com a conseqüència donen lloc a la formació d'agregats (270,316). La variant p.Leu2198_Gln2199dup es troba localitzada en el *repeat* 19, per aquesta raó, es preveu que tindria un efecte dominant negatiu.

La variant de *splicing de novo* c.237+1G>T, identificada en el gen *SMARCE1* dona lloc a la delecio de l'exó 5 *in-frame*, produint-se una proteïna sense la part terminal del domini ric en prolina i l'inici del domini HMG, essent essencial pel correcte funcionament de la proteïna, ja que permet la seva unió a l'ADN (Resultats, Capítol 1, Manuscrit II) (317). Variants *missense* patogèniques en el gen *SMARCE1* que actuarien per un mecanisme de guany de funció o dominant negatiu s'associen a la síndrome de Coffin-Siris (CSS), per contra, variants truncants en aquest mateix gen causen meningiomes cranials i espinals (21,318). El pacient presenta un fenotip compatible amb la síndrome de Coffin-Siris suggerint que variants que donen lloc a delecions *in-frame* actuarien de la mateixa manera que variants *missense*, és a dir, amb un efecte dominant negatiu o de guany de funció. Cal remarcar que és la primera vegada que es descriu una variant de *splicing* (c.237+1G>T) en el gen *SMARCE1* que donaria lloc a una proteïna delecionada, causant la CSS. S'ha reportat una variant de *splicing* en la posició adjacent c.237+2T>C en dos membres d'una família amb múltiples meningiomes espinals, però que no presentaven característiques clíniques de la síndrome de Coffin-Siris. L'anàlisi de l'ARN dels individus afectats mostrava que la variant c.237+2T>C genera dos transcrits alternatius, un d'ells, el més abundant, dona lloc a un codó stop prematur i un segon transcrit menys abundant que produeix la delecio de l'exó 5 *in-frame*. Les diferències en la presentació clínica s'expliquen perquè el pacient 8 presenta un únic transcrit alternatiu que dona lloc a la delecio *in-frame* de l'exó 5, mentre que en el pacient descrit per Smith et al., 2013 el transcrit principal generat per la variant c.237+2T>C produeix una proteïna truncada, causant meningiomes cranials i espinals.

5.2 Anàlisis complementaris en els pacients sense diagnòstic molecular

En els pacients sense diagnòstic molecular (12, 13, 14, 15 i 17) l'anàlisi d'altres models d'herència (autosòmic recessiu i lligat al cromosoma X), de CNVs a partir de les dades de WES, de variants en mosaic i de la càrrega genètica de variants *de novo*, ha permès la identificació de variants candidates o de variants que podrien contribuir al fenotip *SA-like* (Resultats, Annex II, Taula 18).

En la pacient 12 l'anàlisi del model d'herència lligat al cromosoma X ha permès la identificació en el gen *COX7B* de la variant p.Pro74Thr. Variants truncants en el gen *COX7B* s'associen a la síndrome microftàlmia amb lesions linears a la pell amb una herència lligada al cromosoma X dominant (MIM # 300887). En la pacient no estan presents les característiques clíniques de la microftàlmia amb lesions linears a la pell. S'ha descartat que la pacient i la mare presentin una inactivació preferencial del cromosoma X. És per aquest motiu, que una possibilitat és que la variant *missense* p.Pro74Thr no actuï per un mecanisme de pèrdua de funció, sinó com a dominant negatiu o guany de funció, contribuint en la presentació del seu fenotip.

En aquesta mateixa pacient, l'anàlisi de variants en mosaic *de novo* ha identificat la variant de significat incert p.Arg251Ser (R251S) en el gen *TUBB2B*, que es troba en molt baix percentatge (12,5%). Aquesta variant anotada en la base de dades de ClinVar (VCV000432230.2) classificada com una variant de significat incert. Variants *missense* en el gen *TUBB2B* han estat associades a un trastorn del desenvolupament caracteritzat per anomalies estructurals en el cervell que inclouen displàsia cortical, paquigiria i polimicrogiria (319). S'ha descrit un cas amb la variant R380P en el gen *TUBB2B* en mosaic (23%) que mostrava paquigiria, polimicrogiria, displàsia del cerebel, hipoplàsia de l'àrea de pons i hipoplàsia dels nervi òptic, trets tots consistents amb el fenotip associat a *TUBB2B* (209). Considerant que el fenotip present en la pacient 12 no es correlaciona al no presentar malformacions corticals, la variant R251S s'exclouria com a responsable del fenotip *SA-like*, però si podria contribuir en la presentació del fenotip o actuar com a modulador. Mosaïcismes de molt baix nivell poden tenir un impacte en el fenotip, un exemple, és el cas de l'esclerosi tuberosa (MIM # 191100, MIM # 613254). Variants en mosaic en baix percentatge (15-25%) en el gen *TSC2*, es detecten en pacients amb el fenotip típic d'esclerosi tuberosa, suggerint que el percentatge al qual es troba una variant pot no ser predictiu de la gravetat del fenotip (320).

En el pacient 13, l'anàlisi de CNVs i variants en mosaic ha permès la identificació d'una duplicació heretada del pare i una variant *missense* en mosaic *de novo* (p.Arg434Gln) en el gen *ATAD3A*. Variants en *ATAD3A* s'associen a un ampli espectre de fenotips entre els que s'inclou la hipoplàsia pontocerebelar, hipotonia i insuficiència respiratòria (MIM # 618810); paraplegia espàstica hereditària i la síndrome de Harel Yoon (MIM # 617183). Els diferents fenotips s'atribueixen a diversos tipus de variants entre ells, variants hipomòrfiques i delecions bial·lèliques, i variants *missense* monoal·lèliques que actuen per un mecanisme dominant negatiu. Recentment, duplicacions intergèniques en el clúster de gens *ATAD3* s'han associat a opacitat corneal, cardiomiopatia, hipotonia i canvis en la matèria blanca del cervell (321). La duplicació detectada

en el pacient no solapa amb les reportades i el pare és portador sa. No s'han descrit variants *missense* en mosaic associades a clínica i la variant p.Arg434Gln no ha estat descrita prèviament associada a malaltia. A més, el pacient no presenta les característiques clíniques associades a la forma recessiva de la síndrome de Harel-Yoon (255), per aquest motiu s'han classificat com variants de significat incert. L'anàlisi de la càrrega genètica *de novo* ha identificat la variant c.32G>T (p.Cys11Phe) en el gen *METTL24*, que s'ha classificat com una variant de significat incert ja que es prediu que té un efecte neutre i és un gen no associat prèviament amb TND. En aquest pacient, també s'ha identificat una variant truncant *de novo* (p.Ser152Argfs*40) en el gen *HSF2* que es considera la principal candidata a ser la responsable del fenotip *SA-like*.

En la pacient 14 l'anàlisi del model d'herència autosòmic recessiu ha permès la identificació de dues variants en el gen *RMND1* (p.Glu429Val, p.Met350Thr). Variants patogèniques en el gen *RMND1* s'associen a un trastorn mitocondrial (MIM # 614922) amb herència autosòmica recessiva, que es caracteritza per hipotonia neonatal, acidosi làctica, encefalopatia, sordesa neurosensorial, insuficiència renal i retard global del desenvolupament. També s'ha descrit variabilitat en la presentació clínica, amb fenotip més lleu (322,323). El fenotip present en la pacient 14 no es correlaciona amb el del trastorn mitocondrial associat al gen *RMND1*, a excepció del retard del desenvolupament. Podria explicar-se perquè les dues variants *missense* identificades en aquesta pacient tinguin un impacte més lleu sobre la funció de la proteïna. La variant p.Met350Thr es troba localitzada en el domini altament conservat DUF155 del qual es desconeix la funció que porta a terme i que és comú en la família de proteïnes *sif2* (324). D'altra banda, la variant p.Glu429Val al localitzar-se en el domini transmembrana (TM), fet que podria afectar l'ancoratge de la proteïna a la membrana mitocondrial (Figura 48). Cap de les dues variants *missense* han estat reportades anteriorment i no es troben en les bases de dades d'ExAC i gnomAD. Les dues variants s'han classificat com de significat incert seguint la guia ACMG. L'efecte que puguin tenir aquestes variants en el fenotip present en la pacient es desconeix.

En aquesta mateixa pacient s'ha identificat una variant de significat incert en el gen *CSTF2* (p.Val11Gly) que podria ser una bona candidata perquè s'ha demostrat que és un gen implicat en TND i recentment s'ha associat a DI lleu no sindròmica en homes (247).

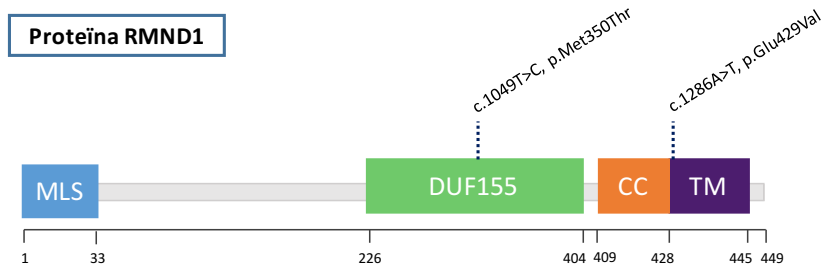


Figura 48 | Estructura de la proteïna codificada pel gen *RMND1*. La línia discontinua indica on es troben localitzades les variants identificades en la pacient 14. MLS, *Mitochondrial localization signal*; DUF155, *domain of unknown function 155*; CC, *coiled-coil*; TM, *transmembrane domain*.

En el pacient 15 l'anàlisi de CNVs ha identificat una duplicació que inclou la regió 5'UTR, l'exó 1 i part de l'intró 1 del gen *TM4SF20* associat a trastorn específic de llenguatge (256). La duplicació ha estat heretada de la mare que és sana, per tant, s'ha descartat com a responsable del fenotip SA-like, però no es pot excloure que estigui contribuint o modulant el fenotip. La resta d'anàlisis realitzats no han permès la identificació de variants candidates al fenotip SA-like en aquest pacient.

Finalment, en el pacient 17 l'anàlisi de les dades de WES ha estat dificultosa perquè no es disposa de les dades genòmiques del pare, tot i això, s'ha identificat una variant truncant en el gen *CHMP7* (p.Gln389*) que considerem que és una bona candidata al fenotip SA-like.

En 4 dels 5 pacients (12, 13, 14 i 15) sense diagnòstic s'ha portat a terme el WGS, els resultats preliminars no han permès la identificació de variants no detectades prèviament per exoma que puguin ser les responsables del fenotip SA-like.

El WGS té diversos avantatges respecte el WES, entre ells: i) cobreix el 98% del genoma, mentre que l'exoma només cobreix les regions codificants que corresponen al 1-2% del genoma ii) permet la detecció d'indels, variants estructurals i de CNVs de manera més eficient, iii) hi ha una major uniformitat en la cobertura perquè no hi ha fase de captura i enriquiment per PCR i iv) s'ha reportat un 7% addicional de diagnòstic a l'aplicar WGS respecte al re-anàlisi de les dades de WES. Entre les desavantatges es troben: i) l'anàlisi de les dades és més complex, ii) es necessita una gran capacitat d'emmagatzematge de dades, iii) la dificultat d'interpretar variants intròniques o en regions intergèniques i iv) un cost més elevat que el WES (134–136,325).

El punt més destacable del WGS és que permet analitzar els elements reguladors, s'ha descrit que variants localitzades en aquests elements poden provocar la pèrdua o el guany d'expressió del gen que regulen. Així mateix, variants estructurals com delecions, inversions o duplicacions que afectin als TADs (*Topologically Associating Domains*) poden alterar potencialment el plegament 3D de la cromatina, resultant en defectes en el control de l'expressió gènica (326,327). El projecte DDD, ha analitzat en pacients amb resultat negatiu d'exoma les variants *de novo* en elements reguladors (elements no codificants conservats, *enhancers* experimentalment validats i *enhancers* del cor). Els resultats indiquen que entre 1-3% dels individus amb TND i sense diagnòstic presenten variants *de novo* patogèniques en elements reguladors que es troben actius en el cervell fetal. (143). És per aquest motiu, que caldrà prestar una especial atenció a aquest tipus de variants en els pacients sense diagnòstic per WES.

5.3 Xarxa d'interacció dels gens identificats com a responsables del fenotip SA-like i UBE3A

La nostra hipòtesi inicial és que els gens identificats en la nostra cohort com a responsables del fenotip SA-like haurien d'interaccionar amb el gen *UBE3A* o intervenir en les vies de senyalització regulades per *UBE3A*. Per a la seva comprovació, s'ha construït una xarxa d'interacció que inclou els gens identificats en la nostra cohort, els gens del diagnòstic diferencial de la SA i el gen *UBE3A*. Com a resultat, cap dels gens identificats en la nostra cohort interaccionen directament amb *UBE3A*. El fet que no s'hagin trobat interaccions directes, no vol dir que els gens identificats no interaccionin de manera indirecta amb *UBE3A* a través d'altres proteïnes, que acabin confluint en les mateixes vies moleculars.

En aquest sentit, hi ha tres interaccions proteïna-proteïna identificades entre els gens SA-like de la nostra cohort i els gens del diagnòstic diferencial de la SA (*TBL1XR1-TCF4*, *STXBP1-VAMP2* i *LAS1L-KANSL1*) i una altra entre dos dels gens identificats en la nostra cohort (*SPTAN1-KCNQ3*). *TBL1XR1* i *TCF4* es troben regulant la transcripció dependent de Wnt. El *knockdown* de *TBL1XR1*, impedeix la unió de la β -catenina amb *TCF4*, impedit la transcripció de gens regulats per Wnt- β -catenina (328). D'altra banda, *VAMP2* forma part del complex SNARE implicat en la fusió de vesícules sinàptiques en els terminals nerviosos i *STXBP1* codifica per *Munc18-1* que intervé en la correcta formació del complex SNARE (292,329). *LAS1L* i *KANSL1* formen part del complex MLL1, que catalitza la metilació de His3Lys4, activant la transcripció (241,252). Finalment, *SPTAN1* és una proteïna del citoesquelet d'espectrina implicada en l'estabilització de la membrana plasmàtica, s'ha descrit que el citoesquelet d'espectrina influencia l'expressió i activació de

canals, com és el cas de *KCNQ3* que codifica per una subunitat d'un canal de potassi dependent de voltatge (330).

L'anàlisi de les funcions i vies moleculars enriquides en la xarxa formada pels gens identificats associats al fenotip SA, mostra que hi ha un enriquiment en vies moleculars implicades en el desenvolupament del sistema nerviós, sinapsis, remodelació de la cromatina i regulació de la transcripció, d'acord amb les funcions que porten a terme els gens identificats en la nostra cohort. Els gens *VAMP2*, *SYNGAP1*, *SLC6A1*, *KCNQ3*, *KIF1A* i *CHMP7* es troben implicats en sinapsis, mentre que els gens *TBL1XR1*, *SATB2*, *SMARCE1*, *ASXL3*, *LAS1L* i *HSF2* actuen com a remodeladors de la cromatina i/o reguladors de la transcripció i el gen *SPTAN1* intervé en l'organització del citoesquelet. Aquests resultats també concorden amb els d'altres xarxes moleculars de gens que participen en el fenotip de TEA, implicats principalment en quatre vies moleculars: remodelació de la cromatina, proliferació cel·lular, migració neuronal i sinapsis (9,331). En la síndrome de Rett-like també s'ha observat un enriquiment de gens implicats en la remodelació de la cromatina i sinapsis (332). Per tant, tot i que els gens identificats majoritàriament no interaccionen directament amb els del diagnòstic diferencial de la SA ni tampoc amb el gen *UBE3A*, si que intervindrien en les mateixes vies moleculars el que explicaria perquè pacients amb variants patogèniques/probablement patogèniques en gens diferents presenten un fenotip solapant.

5.4 Re-avaluació clínica dels pacients amb diagnòstic molecular

La re-avaluació clínica dels pacients ens ha permès determinar si presenten les característiques clíniques associades al gens identificats com a responsables del fenotip SA-like.

Variants patogèniques/probablement patogèniques en els gens *SMARCE1*, *SATB2*, *SYNGAP1*, *KIF1A*, *ASXL3* i *SPTAN1* s'associen a síndromes que presenten trets diferencials no presents en la SA. Per exemple, l'atròfia del cerebel és característica dels gens *KIF1A* i *SPTAN1* (212,270) i està present en els pacients 1 i 9 (Figura 49).

També variants en *SMARCE1* s'associen al fenotip d'hipoplàsia de la 5^a ungla del dit del peu o de la mà, hipertricosi, creixement del cabell a clapes i comportament agressiu, trets tots presents en el pacient 8 (317). Però també succeeix que els pacients no presentin totes les característiques típiques associades als gens identificats, així la pacient 3 no té hipotonia ni marxa atàxica, trets descrits en la majoria de pacients amb variants patogèniques en *SYNGAP1* (266). La pacient 6 no presenta sialorrea ni problemes d'alimentació, trets freqüents, en més del 70% dels individus amb variants patogèniques en *SATB2* (288). Finalment, la marxa atàxica, les estereotípies i la hipotonia

es troben presents en tots els pacients descrits amb la variant p.Arg230Cys en el gen *KCNQ3* (269), però aquestes característiques clíniques no es troben presents en la pacient 7 de la nostra cohort.

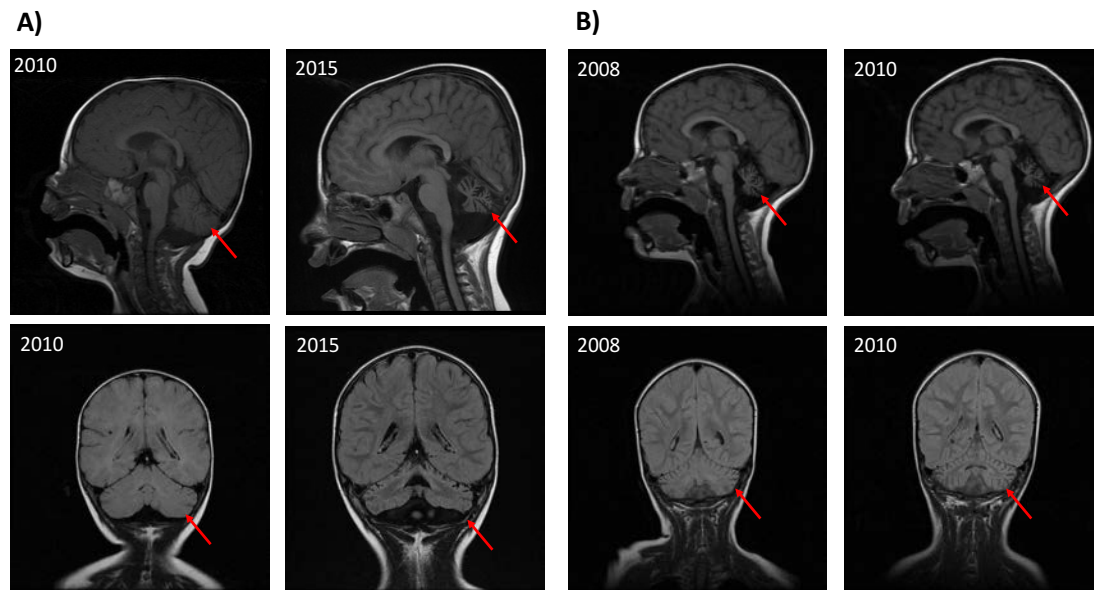


Figura 49 | Ressonàncies magnètiques dels pacients 1 i 9. **A)** Les imatges pertanyen al pacient 1 amb la variant c.506G>C, p.Arg169Thr en el gen *KIF1A*. Es pot observar com hi ha una atròfia progressiva del cerebel. **B)** Escàner cerebral que pertany a la pacient 9 amb la variant c.6592_6597dupCTGCAG, p.Leu2198_2199dup. Es pot observar com la pacient presenta una atròfia del cerebel progressiva. Les fletxes en vermell indiquen el cerebel.

En els TND existeix una gran variabilitat fenotípica, un exemple és el TEA, on s'ha descrit que en un 10-20% dels casos dels pacients amb variants en un mateix gen (*MECP2*, *NLGN4X* i *SHANK2*, entre d'altres) presenten diferències en la presentació clínica (333). Aquesta variabilitat es deu a factors genètics i ambientals que modularien el fenotip, entre ells: i) interacció amb altres variants tant rares com freqüents localitzades en altres gens, ii) variants en el mateix gen que actuen com a modificadors i donen lloc a una presentació diferent de la malaltia i iii) factors ambientals (155,334).

Un cop revisades les característiques clíniques dels nostres pacients podem afirmar que tots són SA-like. En alguns casos s'observen característiques clíniques no associades a la SA i que podrien suggerir un diagnòstic diferencial. Per tant, es proposa que els gens identificats en la nostra cohort s'inclouin en el diagnòstic diferencial de la SA. Exceptuant el gen *SYNGAP1* (335), que anteriorment havia estat proposat, la resta de gens no han estat prèviament relacionats amb la SA.

En resum, els resultats d'aquest projecte posen de manifest que la major part (70,5%) del 10% dels pacients amb un fenotip SA i sense diagnòstic molecular (*SA-like*) presenten alteracions genètiques en altres gens que no són *UBE3A*. Els nous gens identificats en l'etiologia del fenotip *SA-like* estan implicats en els TND, molts d'ells solapants amb la SA. Existeix una gran heterogeneïtat genètica subjacent entre els pacients amb un fenotip *SA-like*. L'aplicació de WES en aquests pacients ha estat una eina útil per a la identificació de nous gens implicats en aquest fenotip, així com pel descobriment de nous gens candidats. A més, tenint en compte l'alta taxa de diagnòstic obtinguda, considerem que el WES s'hauria d'incloure dins de l'algorisme diagnòstic de la SA (Figura 50). Identificar la causa a nivell genètic i conèixer les vies moleculars afectades en aquests pacients portarà a identificar possibles dianes terapèutiques. El diagnòstic genètic en aquests pacients és molt important pel correcte assessorament genètic a la família, un millor seguiment de la malaltia, un major coneixement sobre el seu pronòstic i possibles complicacions, i disposar de noves dianes terapèutiques.

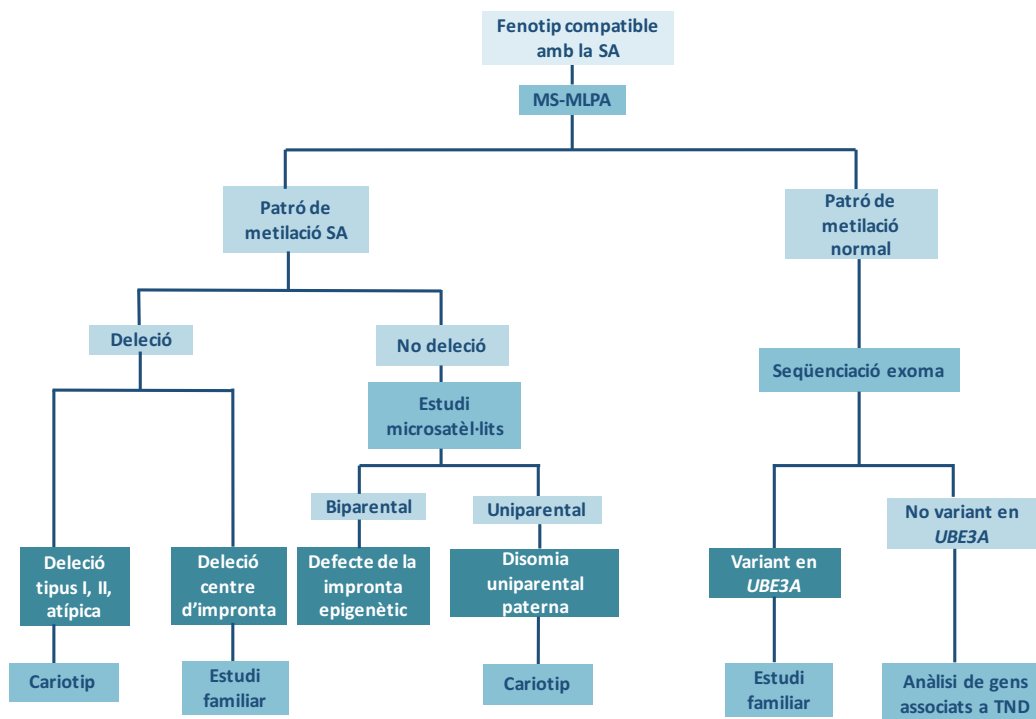


Figura 50| Algorisme diagnòstic proposat per als pacients amb un fenotip compatible amb la SA. TND, trastorns del neurodesenvolupament.



Conclusions

6. CONCLUSIONS

1. El 70,5% dels pacients *SA-like* presenten variants patogèniques/probablement patogèniques en gens prèviament associats a trastorns del neurodesenvolupament. Aquesta taxa de diagnòstic recolza el WES en trios com a eina principal pel diagnòstic de la *SA-like*.
2. El 91,6% de les variants són *de novo*.
3. Els pacients *SA-like* presenten una gran heterogeneïtat genètica, només 2 dels 12 pacients amb diagnòstic molecular presenten variants en el mateix gen (*TBL1XR1*).
4. En dos pacients s'ha identificat una variant truncant *de novo* en dos nous gens candidats a la *SA-like* (*HSF2* i *CHMP7*) no associats prèviament amb trastorns del neurodesenvolupament.
5. Les variants truncants en els gens *SYNGAP1* i *SATB2* no presenten activació del mecanisme de *Nonsense-mediated mRNA Decay* (NMD) i per tant, no es pot afirmar que actuïn per un mecanisme de pèrdua de funció. Les variants truncants en els gens *HSF2* i *CHMP7* activen el mecanisme NMD i per tant actuen per un mecanisme d'haploinsuficiència.
6. L'estudi funcional de la variant R169T localitzada en el domini motor de *KIF1A*, demostra que es tracta d'una variant deletèria que afecta l'activitat ATPasa del domini motor i de manera indirecta la motilitat de la proteïna. Aquest resultat ens permet classificar la variant com a patogènica i directament responsable del fenotip *SA-like* del pacient.
7. Els gens identificats com a responsables del fenotip *SA-like* i els nous gens candidats no interaccionen directament amb *UBE3A*. Tot i això, aquests gens intervenen en vies moleculars comuns relacionades amb neurogènesi, sinapsi, remodelació de la cromatina i regulació de la transcripció.
8. Els gens identificats en la nostra cohort responsables del fenotip *SA-like* s'haurien d'incloure en el diagnòstic diferencial de la SA.



Bibliografia

1. Stiles J, Jernigan TL. The basics of brain development. *Neuropsychol Rev*. 2010 Dec;20(4):327–48.
2. Tau GZ, Peterson BS. Normal development of brain circuits. *Neuropsychopharmacology*. 2010 Jan;35(1):147–68.
3. Moreno-De-Luca A, Myers SM, Challman TD, Moreno-De-Luca D, Evans DW, Ledbetter DH. Developmental brain dysfunction: revival and expansion of old concepts based on new genetic evidence. *Lancet Neurol*. 2013 Apr;12(4):406–14.
4. Cardoso AR, Lopes-Marques M, Silva RM, Serrano C, Amorim A, Prata MJ, et al. Essential genetic findings in neurodevelopmental disorders. *Hum Genomics*. 2019 Jul;13(1):31.
5. Mullin AP, Gokhale A, Moreno-De-Luca A, Sanyal S, Waddington JL, Faundez V. Neurodevelopmental disorders: mechanisms and boundary definitions from genomes, interactomes and proteomes. *Transl Psychiatry*. 2013 Dec;3:e329.
6. Niemi MEK, Martin HC, Rice DL, Gallone G, Gordon S, Kelemen M, et al. Common genetic variants contribute to risk of rare severe neurodevelopmental disorders. *Nature*. 2018 Oct;562(7726):268–71.
7. Owen MJ, O'Donovan MC. Schizophrenia and the neurodevelopmental continuum:evidence from genomics. *World Psychiatry*. 2017 Oct;16(3):227–35.
8. Diagnostic and statistical manual of mental disorders: DSM-5™, 5th ed. Diagnostic and statistical manual of mental disorders: DSM-5™, 5th ed. Arlington, VA, US: American Psychiatric Publishing, Inc.; 2013. xlv, 947–xlv, 947.
9. Hormozdiari F, Penn O, Borenstein E, Eichler EE. The discovery of integrated gene networks for autism and related disorders. *Genome Res*. 2015 Jan;25(1):142–54.
10. Ilyas M, Mir A, Efthymiou S, Houlden H. The genetics of intellectual disability: advancing technology and gene editing. *F1000Research*. 2020;9.
11. Maulik PK, Mascarenhas MN, Mathers CD, Dua T, Saxena S. Prevalence of intellectual disability: a meta-analysis of population-based studies. *Res Dev Disabil*. 2011;32(2):419–36.
12. Patel DR, Cabral MD, Ho A, Merrick J. A clinical primer on intellectual disability. *Transl Pediatr*. 2020 Feb;9(Suppl 1):S23–35.
13. King BH. Intellectual Disability: Understanding Its Development, Causes, Classification, Evaluation, and Treatment. *JAMA*. 2008 Mar 12;299(10):1194.
14. Pipan M. Intellectual Disability: Definition, Classification, and Systems of Supports 11th ed. *J Dev Behav Pediatr*. 2012;33(5):386.
15. Huang J, Zhu T, Qu Y, Mu D. Prenatal, Perinatal and Neonatal Risk Factors for Intellectual Disability: A Systemic Review and Meta-Analysis. *PLoS One*. 2016;11(4):e0153655.
16. Chiurazzi P, Pirozzi F. Advances in understanding - genetic basis of intellectual disability. *F1000Research*. 2016;5.
17. Ronan JL, Wu W, Crabtree GR. From neural development to cognition: unexpected roles for chromatin. *Nat Rev Genet*. 2013 May;14(5):347–59.
18. Narayanan R, Pirouz M, Kerimoglu C, Pham L, Wagener RJ, Kiszka KA, et al. Loss of BAF (mSWI/SNF) Complexes Causes Global Transcriptional and Chromatin State Changes in Forebrain Development. *Cell Rep*. 2015 Dec;13(9):1842–54.
19. Bachmann C, Nguyen H, Rosenbusch J, Pham L, Rabe T, Patwa M, et al. mSWI/SNF (BAF) Complexes Are Indispensable for the Neurogenesis and Development of Embryonic Olfactory Epithelium. *PLoS Genet*. 2016 Sep;12(9):e1006274.
20. Ooi L, Belyaev ND, Miyake K, Wood IC, Buckley NJ. BRG1 chromatin remodeling activity is required for efficient chromatin binding by repressor element 1-silencing transcription factor (REST) and facilitates REST-mediated repression. *J Biol Chem*. 2006 Dec;281(51):38974–80.
21. Koshio T, Okamoto N. Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. *Am J Med Genet C Semin Med Genet*. 2014 Sep;166C(3):262–75.

22. Van Houdt JKJ, Nowakowska BA, Sousa SB, van Schaik BDC, Seuntjens E, Avonce N, et al. Heterozygous missense mutations in SMARCA2 cause Nicolaides-Baraitser syndrome. *Nat Genet.* 2012 Feb;44(4):445–9, S1.
23. Cross NCP. Histone modification defects in developmental disorders and cancer. Vol. 3, *Oncotarget.* United States; 2012. p. 3–4.
24. Kramer JM, van Bokhoven H. Genetic and epigenetic defects in mental retardation. *Int J Biochem Cell Biol.* 2009 Jan;41(1):96–107.
25. Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. *Cell.* 2013 Mar;152(6):1237–51.
26. Forrest MP, Hill MJ, Kavanagh DH, Tansey KE, Waite AJ, Blake DJ. The Psychiatric Risk Gene Transcription Factor 4 (TCF4) Regulates Neurodevelopmental Pathways Associated With Schizophrenia, Autism, and Intellectual Disability. *Schizophr Bull.* 2018 Aug;44(5):1100–10.
27. Quednow BB, Brzozka MM, Rossner MJ. Transcription factor 4 (TCF4) and schizophrenia: integrating the animal and the human perspective. *Cell Mol Life Sci.* 2014 Aug;71(15):2815–35.
28. Bengani H, Handley M, Alvi M, Ibitoye R, Lees M, Lynch SA, et al. Clinical and molecular consequences of disease-associated de novo mutations in SATB2. *Genet Med.* 2017 Aug;19(8):900–8.
29. Coyle JE, Nikolov DB. GABARAP: lessons for synaptogenesis. *Neurosci a Rev J bringing Neurobiol Neurol psychiatry.* 2003 Jun;9(3):205–16.
30. Hsueh Y-P. Synaptic Formation, Neural Circuits and Neurodevelopmental Disorders Controlled by Signaling, Translation, and Epigenetic Regulation. Vol. 79, *Developmental neurobiology.* United States; 2019. p. 2–7.
31. Takesian AE, Hensch TK. Balancing plasticity/stability across brain development. *Prog Brain Res.* 2013;207:3–34.
32. Chen J, Yu S, Fu Y, Li X. Synaptic proteins and receptors defects in autism spectrum disorders. *Front Cell Neurosci.* 2014;8:276.
33. Giovedi S, Corradi A, Fassio A, Benfenati F. Involvement of synaptic genes in the pathogenesis of autism spectrum disorders: the case of synapsins. *Front Pediatr.* 2014;2:94.
34. Williams CA, Zori RT, Hendrickson J, Stalker H, Marum T, Whidden E, et al. Angelman syndrome. *Curr Probl Pediatr.* 1995 Aug;25(7):216–31.
35. Williams CA, Angelman H, Clayton-Smith J, Driscoll DJ, Hendrickson JE, Knoll JH, et al. Angelman syndrome: consensus for diagnostic criteria. Angelman Syndrome Foundation. *Am J Med Genet.* 1995 Mar;56(2):237–8.
36. Williams C, Beaudet AL, Clayton-Smith J, Knoll JH, Kyllerman M, Laan LA, et al. Angelman Syndrome 2005: Updated Consensus for Diagnostic Criteria Charles. *Am J Med Genet Part A.* 2006;140(5):413–8.
37. Kyllerman M. On the prevalence of Angelman syndrome. Vol. 59, *American journal of medical genetics.* United States; 1995. p. 405; author reply 403-4.
38. Buckley RH, Dinno N, Weber P. Angelman syndrome: are the estimates too low? *Am J Med Genet.* 1998 Dec;80(4):385–90.
39. Williams C, Driscoll DJ, Dagi AI. Clinical and genetic aspects of Angelman syndrome. *Genet Med.* 2010;12(7):385–95.
40. Beckung E, Steffenburg S, Kyllerman M. Motor impairments, neurological signs, and developmental level in individuals with Angelman syndrome. *Dev Med Child Neurol.* 2004 Apr;46(4):239–43.
41. Pearson E, Wilde L, Heald M, Royston R, Oliver C. Communication in Angelman syndrome: a scoping review. *Dev Med Child Neurol.* 2019 May;61(11):1266–74.
42. Barry RJ, Leitner RP, Clarke AR, Einfeld SL. Behavioral aspects of Angelman syndrome: a case control study. *Am J Med Genet A.* 2005 Jan;132A(1):8–12.

43. Paprocka J, Jamroz E, Szwed-Bialozyt B, Jezela-Stanek A, Kopyta I, Marszal E. Angelman syndrome revisited. *Neurologist*. 2007 Sep;13(5):305–12.
44. Tan W-H, Bacino CA, Skinner SA, Anselm I, Barbieri-Welge R, Bauer-Carlin A, et al. Angelman syndrome: Mutations influence features in early childhood. *Am J Med Genet A*. 2011 Jan;155A(1):81–90.
45. Wheeler AC, Sacco P, Cabo R. Unmet clinical needs and burden in Angelman syndrome: a review of the literature. *Orphanet J Rare Dis*. 2017 Oct;12(1):164.
46. Clayton-Smith J, Laan L. Angelman syndrome: a review of the clinical and genetic aspects. *J Med Genet*. 2003 Feb;40(2):87–95.
47. Buoni S, Grosso S, Pucci L, Fois A. Diagnosis of Angelman syndrome: clinical and EEG criteria. *Brain Dev*. 1999 Jul;21(5):296–302.
48. Boyd SG, Harden A, Patton MA. The EEG in early diagnosis of the Angelman (happy puppet) syndrome. *Eur J Pediatr*. 1988 Jun;147(5):508–13.
49. Buiting K. Prader-Willi syndrome and Angelman syndrome. *Am J Med Genet Part C Semin Med Genet*. 2010;154(3):365–76.
50. Bird LM. Angelman syndrome: Review of clinical and molecular aspects. *Appl Clin Genet*. 2014;7:93–104.
51. Dindot S V, Antalffy BA, Bhattacharjee MB, Beaudet AL. The Angelman syndrome ubiquitin ligase localizes to the synapse and nucleus, and maternal deficiency results in abnormal dendritic spine morphology. *Hum Mol Genet*. 2008 Jan;17(1):111–8.
52. Meng L, Person RE, Huang W, Zhu PJ, Costa-Mattioli M, Beaudet AL. Truncation of Ube3a-ATS unsilences paternal Ube3a and ameliorates behavioral defects in the Angelman syndrome mouse model. *PLoS Genet*. 2013;9(12):e1004039.
53. Ramsden SC, Clayton-Smith J, Birch R, Buiting K. Practice guidelines for the molecular analysis of Prader-Willi and Angelman syndromes. *BMC Med Genet*. 2010 May;11:70.
54. Beygo J, Buiting K, Ramsden SC, Ellis R, Clayton-Smith J, Kanber D. Update of the EMQN/ACGS best practice guidelines for molecular analysis of Prader-Willi and Angelman syndromes. *Eur J Hum Genet*. 2019 Sep;27(9):1326–40.
55. Amos-Landgraf JM, Ji Y, Gottlieb W, Depinet T, Wandstrat AE, Cassidy SB, et al. Chromosome breakage in the Prader-Willi and Angelman syndromes involves recombination between large, transcribed repeats at proximal and distal breakpoints. *Am J Hum Genet*. 1999 Aug;65(2):370–86.
56. Sadikovic B, Fernandes P, Zhang VW, Ward PA, Miloslavskaya I, Rhead W, et al. Mutation Update for UBE3A variants in Angelman syndrome. *Hum Mutat*. 2014 Dec;35(12):1407–17.
57. Tzagkaraki E, Sofocleous C, Fryssira-Kanioura H, Dinopoulos A, Goulielmos G, Mavrou A, et al. Screening of UBE3A gene in patients referred for Angelman Syndrome. *Eur J Paediatr Neurol*. 2013 Jul;17(4):366–73.
58. Aguilera C, Vinas-Jornet M, Baena N, Gabau E, Fernandez C, Capdevila N, et al. Novel intragenic deletions within the UBE3A gene in two unrelated patients with Angelman syndrome: case report and review of the literature. *BMC Med Genet*. 2017 Nov;18(1):137.
59. Robinson WP, Kuchinka BD, Bernasconi F, Petersen MB, Schulze A, Brondum-Nielsen K, et al. Maternal meiosis I non-disjunction of chromosome 15: dependence of the maternal age effect on level of recombination. *Hum Mol Genet*. 1998 Jun;7(6):1011–9.
60. Robinson WP, Christian SL, Kuchinka BD, Penaherrera MS, Das S, Schuffenhauer S, et al. Somatic segregation errors predominantly contribute to the gain or loss of a paternal chromosome leading to uniparental disomy for chromosome 15. *Clin Genet*. 2000 May;57(5):349–58.
61. Buiting K, Gross S, Lich C, Gillissen-Kaesbach G, el-Maarri O, Horsthemke B. Epimutations in Prader-Willi and Angelman syndromes: a molecular study of 136 patients with an imprinting defect. *Am J Hum Genet*. 2003 Mar;72(3):571–7.
62. Sell GL, Margolis SS. From UBE3A to Angelman syndrome: a substrate perspective. *Front*

- Neurosci. 2015;9:322.
63. Kee Y, Huibregtse JM. Regulation of catalytic activities of HECT ubiquitin ligases. *Biochem Biophys Res Commun.* 2007 Mar;354(2):329–33.
 64. Lee SY, Ramirez J, Franco M, Lectez B, Gonzalez M, Barrio R, et al. Ube3a, the E3 ubiquitin ligase causing Angelman syndrome and linked to autism, regulates protein homeostasis through the proteasomal shuttle Rpn10. *Cell Mol Life Sci.* 2014 Jul;71(14):2747–58.
 65. Tomaic V, Banks L. Angelman syndrome-associated ubiquitin ligase UBE3A/E6AP mutants interfere with the proteolytic activity of the proteasome. *Cell Death Dis.* 2015 Jan;6:e1625.
 66. Ramirez J, Lectez B, Osinalde N, Sivá M, Elu N, Aloria K, et al. Quantitative proteomics reveals neuronal ubiquitination of Rngo/Ddi1 and several proteasomal subunits by Ube3a, accounting for the complexity of Angelman syndrome. *Hum Mol Genet.* 2018 Jun;27(11):1955–71.
 67. Mabb AM, Judson MC, Zylka MJ, Philpot BD. Angelman syndrome: insights into genomic imprinting and neurodevelopmental phenotypes. *Trends Neurosci.* 2011 Jun;34(6):293–303.
 68. Gossan NC, Zhang F, Guo B, Jin D, Yoshitane H, Yao A, et al. The E3 ubiquitin ligase UBE3A is an integral component of the molecular circadian clock through regulating the BMAL1 transcription factor. *Nucleic Acids Res.* 2014 May;42(9):5765–75.
 69. Egawa K, Kitagawa K, Inoue K, Takayama M, Takayama C, Saitoh S, et al. Decreased tonic inhibition in cerebellar granule cells causes motor dysfunction in a mouse model of Angelman syndrome. *Sci Transl Med.* 2012 Dec;4(163):163ra157.
 70. Martinez-Noel G, Luck K, Kuhnle S, Desbuleux A, Szajner P, Galligan JT, et al. Network Analysis of UBE3A/E6AP-Associated Proteins Provides Connections to Several Distinct Cellular Processes. *J Mol Biol.* 2018 Mar;430(7):1024–50.
 71. Martínez-Noël G, Galligan JT, Sowa ME, Arndt V, Overton TM, Harper JW, et al. Identification and proteomic analysis of distinct UBE3A/E6AP protein complexes. *Mol Cell Biol.* 2012 Aug;32(15):3095–106.
 72. Nawaz Z, Lonard DM, Smith CL, Lev-Lehman E, Tsai SY, Tsai MJ, et al. The Angelman syndrome-associated protein, E6-AP, is a coactivator for the nuclear hormone receptor superfamily. *Mol Cell Biol.* 1999 Feb;19(2):1182–9.
 73. Greer PL, Hanayama R, Bloodgood BL, Mardinly AR, Lipton DM, Flavell SW, et al. The Angelman Syndrome protein Ube3A regulates synapse development by ubiquitinating arc. *Cell.* 2010 Mar;140(5):704–16.
 74. Kühnle S, Mothes B, Matentzoglou K, Scheffner M. Role of the ubiquitin ligase E6AP/UBE3A in controlling levels of the synaptic protein Arc. *Proc Natl Acad Sci U S A.* 2013 May;110(22):8888–93.
 75. Godavarthi SK, Dey P, Maheshwari M, Jana NR. Defective glucocorticoid hormone receptor signaling leads to increased stress and anxiety in a mouse model of Angelman syndrome. *Hum Mol Genet.* 2012 Apr;21(8):1824–34.
 76. Bochman ML, Schwacha A. The Mcm2-7 complex has in vitro helicase activity. *Mol Cell.* 2008 Jul;31(2):287–93.
 77. Jiang YH, Armstrong D, Albrecht U, Atkins CM, Noebels JL, Eichele G, et al. Mutation of the Angelman ubiquitin ligase in mice causes increased cytoplasmic p53 and deficits of contextual learning and long-term potentiation. *Neuron.* 1998 Oct;21(4):799–811.
 78. Jana NR. Understanding the pathogenesis of Angelman syndrome through animal models. *Neural Plast.* 2012;2012:710943.
 79. Weeber EJ, Jiang Y-H, Elgersma Y, Varga AW, Carrasquillo Y, Brown SE, et al. Derangements of hippocampal calcium/calmodulin-dependent protein kinase II in a mouse model for Angelman mental retardation syndrome. *J Neurosci.* 2003 Apr;23(7):2634–44.
 80. van Woerden GM, Harris KD, Hojjati MR, Gustin RM, Qiu S, de Avila Freire R, et al. Rescue of neurological deficits in a mouse model for Angelman syndrome by reduction of alphaCaMKII inhibitory phosphorylation. *Nat Neurosci.* 2007 Mar;10(3):280–2.

81. Kaphzan H, Hernandez P, Jung JI, Cowansage KK, Deinhardt K, Chao M V, et al. Reversal of impaired hippocampal long-term potentiation and contextual fear memory deficits in Angelman syndrome model mice by ErbB inhibitors. *Biol Psychiatry*. 2012 Aug;72(3):182–90.
82. Tonazzini I, Van Woerden GM, Masciullo C, Mientjes EJ, Elgersma Y, Cecchini M. The role of ubiquitin ligase E3A in polarized contact guidance and rescue strategies in UBE3A-deficient hippocampal neurons. *Mol Autism*. 2019;10:41.
83. Sonzogni M, Hakonen J, Bernabe Kleijn M, Silva-Santos S, Judson MC, Philpot BD, et al. Delayed loss of UBE3A reduces the expression of Angelman syndrome-associated phenotypes. *Mol Autism*. 2019;10:23.
84. Silva-Santos S, van Woerden GM, Bruinsma CF, Mientjes E, Jolfaei MA, Distel B, et al. Ube3a reinstatement identifies distinct developmental windows in a murine Angelman syndrome model. *J Clin Invest*. 2015 May;125(5):2069–76.
85. Stockwell J, Chen Z, Niazi M, Nosib S, Cayabyab FS. Protein phosphatase role in adenosine A1 receptor-induced AMPA receptor trafficking and rat hippocampal neuronal damage in hypoxia/reperfusion injury. *Neuropharmacology*. 2016 Mar;102:254–65.
86. Yu UY, Yoo BC, Ahn J-H. Regulatory B Subunits of Protein Phosphatase 2A Are Involved in Site-specific Regulation of Tau Protein Phosphorylation. *Korean J Physiol Pharmacol Off J Korean Physiol Soc Korean Soc Pharmacol*. 2014 Apr;18(2):155–61.
87. Wang J, Lou S-S, Wang T, Wu R-J, Li G, Zhao M, et al. UBE3A-mediated PTPA ubiquitination and degradation regulate PP2A activity and dendritic spine morphology. *Proc Natl Acad Sci U S A*. 2019 Jun;116(25):12500–5.
88. Kim S, Chahrour M, Ben-Shachar S, Lim J. Ube3a/E6AP is involved in a subset of MeCP2 functions. *Biochem Biophys Res Commun*. 2013 Jul;437(1):67–73.
89. Kuhnle S, Kogel U, Glockzin S, Marquardt A, Ciechanover A, Matentzoglou K, et al. Physical and functional interaction of the HECT ubiquitin-protein ligases E6AP and HERC2. *J Biol Chem*. 2011 Jun;286(22):19410–6.
90. Forrest MP, Waite AJ, Martin-Rendon E, Blake DJ. Knockdown of human TCF4 affects multiple signaling pathways involved in cell survival, epithelial to mesenchymal transition and neuronal differentiation. *PLoS One*. 2013;8(8):e73169.
91. Straub J, Gregor A, Sauerer T, Fliedner A, Distel L, Suchy C, et al. Genetic interaction screen for severe neurodevelopmental disorders reveals a functional link between Ube3a and Mef2 in *Drosophila melanogaster*. *Sci Rep*. 2020 Jan;10(1):1204.
92. Thibert RL, Larson AM, Hsieh DT, Raby AR, Thiele EA. Neurologic Manifestations of Angelman Syndrome. *Pediatr Neurol*. 2013;48(4):271–9.
93. Varela MC, Kok F, Otto PA, Koiffmann CP. Phenotypic variability in Angelman syndrome: comparison among different deletion classes and between deletion and UPD subjects. *Eur J Hum Genet*. 2004 Dec;12(12):987–92.
94. Sahoo T, Bacino CA, German JR, Shaw CA, Bird LM, Kimonis V, et al. Identification of novel deletions of 15q11q13 in Angelman syndrome by array-CGH: molecular characterization and genotype-phenotype correlations. *Eur J Hum Genet*. 2007 Sep;15(9):943–9.
95. Valente KD, Varela MC, Koiffmann CP, Andrade JQ, Grossmann R, Kok F, et al. Angelman syndrome caused by deletion: a genotype-phenotype correlation determined by breakpoint. *Epilepsy Res*. 2013 Jul;105(1–2):234–9.
96. Sahoo T, Peters SU, Madduri NS, Glaze DG, German JR, Bird LM, et al. Microarray based comparative genomic hybridization testing in deletion bearing patients with Angelman syndrome: genotype-phenotype correlations. *J Med Genet*. 2006 Jun;43(6):512–6.
97. Mertz LGB, Thaulov P, Trillingsgaard A, Christensen R, Vogel I, Hertz JM, et al. Neurodevelopmental outcome in Angelman syndrome: genotype-phenotype correlations. *Res Dev Disabil*. 2014 Jul;35(7):1742–7.
98. Bonati MT, Russo S, Finelli P, Valsecchi MR, Cogliati F, Cavalleri F, et al. Evaluation of autism traits in Angelman syndrome: a resource to unfold autism genes. *Neurogenetics*. 2007

- Aug;8(3):169–78.
99. Peters SU, Horowitz L, Barbieri-Welge R, Taylor JL, Hundley RJ. Longitudinal follow-up of autism spectrum features and sensory behaviors in Angelman syndrome by deletion class. *J Child Psychol Psychiatry*. 2012 Feb;53(2):152–9.
 100. Trillingsgaard A, OStergaard JR. Autism in Angelman syndrome: an exploration of comorbidity. *Autism*. 2004 Jun;8(2):163–74.
 101. Moncla A, Malzac P, Voelckel MA, Auquier P, Girardot L, Mattei MG, et al. Phenotype-genotype correlation in 20 deletion and 20 non-deletion Angelman syndrome patients. *Eur J Hum Genet*. 1999;7(2):131–9.
 102. Lossie AC, Whitney MM, Amidon D, Dong HJ, Chen P, Theriaque D, et al. Distinct phenotypes distinguish the molecular classes of Angelman syndrome. *J Med Genet*. 2001 Dec;38(12):834–45.
 103. Nazlican H, Zeschnick M, Claussen U, Michel S, Boehringer S, Gillissen-Kaesbach G, et al. Somatic mosaicism in patients with Angelman syndrome and an imprinting defect. *Hum Mol Genet*. 2004 Nov;13(21):2547–55.
 104. Le Fevre A, Beygo J, Silveira C, Kamien B, Clayton-Smith J, Colley A, et al. Atypical Angelman syndrome due to a mosaic imprinting defect: Case reports and review of the literature. *Am J Med Genet A*. 2017 Mar;173(3):753–7.
 105. Aypar U, Hoppman NL, Thorland EC, Dawson DB. Patients with mosaic methylation patterns of the Prader-Willi/Angelman Syndrome critical region exhibit AS-like phenotypes with some PWS features. *Mol Cytogenet*. 2016;9:26.
 106. Carson RP, Bird L, Childers AK, Wheeler F, Duis J. Preserved expressive language as a phenotypic determinant of Mosaic Angelman Syndrome. *Mol Genet genomic Med*. 2019 Sep;7(9):e837.
 107. LaSalle JM, Reiter LT, Chamberlain SJ. Epigenetic regulation of UBE3A and roles in human neurodevelopmental disorders. *Epigenomics*. 2015 Oct;7(7):1213–28.
 108. Finucane BM, Lusk L, Arkilo D, Chamberlain S, Devinsky O, Dindot S, et al. 15q Duplication Syndrome and Related Disorders [Internet]. *GeneReviews*®. 1993 [cited 2020 Jan 26]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27308687>
 109. Scoles HA, Urraca N, Chadwick SW, Reiter LT, Lasalle JM. Increased copy number for methylated maternal 15q duplications leads to changes in gene and protein expression in human cortical samples. *Mol Autism*. 2011 Dec;2(1):19.
 110. Noor A, Dupuis L, Mittal K, Lionel AC, Marshall CR, Scherer SW, et al. 15q11.2 Duplication Encompassing Only the UBE3A Gene Is Associated with Developmental Delay and Neuropsychiatric Phenotypes. *Hum Mutat*. 2015 Jul;36(7):689–93.
 111. Yi JJ, Berrios J, Newbern JM, Snider WD, Philpot BD, Hahn KM, et al. An Autism-Linked Mutation Disables Phosphorylation Control of UBE3A. *Cell*. 2015 Aug;162(4):795–807.
 112. Angelman Syndrome - *GeneReviews*® - NCBI Bookshelf [Internet]. [cited 2020 May 22]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK1144/?report=reader>
 113. Tan WH, Bird LM, Thibert RL, Williams CA. If not Angelman, what is it? A review of Angelman-like syndromes. *Am J Med Genet Part A*. 2014;164(4):975–92.
 114. Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, Walters L. New goals for the U.S. Human Genome Project: 1998-2003. *Science*. 1998 Oct;282(5389):682–9.
 115. Understanding Human Genetic Variation - NIH Curriculum Supplement Series - NCBI Bookshelf [Internet]. [cited 2020 May 9]. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK20363/>
 116. Lupski JR. Structural variation mutagenesis of the human genome: Impact on disease and evolution. *Environ Mol Mutagen*. 2015 Jun;56(5):419–36.
 117. Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, et al. Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science*. 2007 Feb;315(5813):848–53.
 118. Moraes F, Goes A. A decade of human genome project conclusion: Scientific diffusion

- about our genome knowledge. *Biochem Mol Biol Educ*. 2016 May;44(3):215–23.
119. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al. A global reference for human genetic variation. *Nature*. 2015 Oct;526(7571):68–74.
 120. Turnbull C, Scott RH, Thomas E, Jones L, Murugaesu N, Pretty FB, et al. The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. *BMJ*. 2018 Apr;361:k1687.
 121. Scott RH, Fowler TA, Caulfield M. Genomic medicine: time for health-care transformation. *Lancet (London, England)*. 2019 Aug;394(10197):454–6.
 122. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A*. 1977 Dec;74(12):5463–7.
 123. Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet*. 2016 May;17(6):333–51.
 124. Mantere T, Kersten S, Hoischen A. Long-Read Sequencing Emerging in Medical Genetics. *Front Genet*. 2019;10:426.
 125. Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, Gouil Q. Opportunities and challenges in long-read sequencing data analysis. *Genome Biol*. 2020 Feb;21(1):30.
 126. Bean LJH, Funke B, Carlston CM, Gannon JL, Kantarci S, Krock BL, et al. Diagnostic gene sequencing panels: from design to report-a technical standard of the American College of Medical Genetics and Genomics (ACMG). *Genet Med*. 2020 Mar;22(3):453–61.
 127. Dillon OJ, Lunke S, Stark Z, Yeung A, Thorne N, Gaff C, et al. Exome sequencing has higher diagnostic yield compared to simulated disease-specific panels in children with suspected monogenic disorders. *Eur J Hum Genet*. 2018 May;26(5):644–51.
 128. Wenger AM, Guturu H, Bernstein JA, Bejerano G. Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. *Genet Med*. 2017 Feb;19(2):209–14.
 129. Ewans LJ, Schofield D, Shrestha R, Zhu Y, Gayevskiy V, Ying K, et al. Whole-exome sequencing reanalysis at 12 months boosts diagnosis and is cost-effective when applied early in Mendelian disorders. *Genet Med*. 2018 Dec;20(12):1564–74.
 130. Lawrence L, Sincan M, Markello T, Adams DR, Gill F, Godfrey R, et al. The implications of familial incidental findings from exome sequencing: the NIH Undiagnosed Diseases Program experience. *Genet Med*. 2014 Oct;16(10):741–50.
 131. Yang Y, Muzny DM, Reid JG, Bainbridge MN, Willis A, Ward PA, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med*. 2013 Oct;369(16):1502–11.
 132. Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med*. 2016 Jul;18(7):696–704.
 133. Xue Y, Ankala A, Wilcox WR, Hegde MR. Solving the molecular diagnostic testing conundrum for Mendelian disorders in the era of next-generation sequencing: single-gene, gene panel, or exome/genome sequencing. *Genet Med*. 2015 Jun;17(6):444–51.
 134. Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, et al. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci U S A*. 2015 Apr;112(17):5473–8.
 135. Alfares A, Aloraini T, Subaie L Al, Alissa A, Qudsi A Al, Alahmad A, et al. Whole-genome sequencing offers additional but limited clinical utility compared with reanalysis of whole-exome sequencing. *Genet Med*. 2018 Nov;20(11):1328–33.
 136. Barbitoff YA, Polev DE, Glotov AS, Serebryakova EA, Shcherbakova I V, Kiselev AM, et al. Systematic dissection of biases in whole-exome and whole-genome sequencing reveals major determinants of coding sequence coverage. *Sci Rep*. 2020 Feb;10(1):2057.
 137. Monroe GR, Frederix GW, Savelberg SMC, de Vries TI, Duran KJ, van der Smagt JJ, et al. Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genet Med*. 2016 Sep;18(9):949–56.
 138. Srivastava S, Love-Nichols JA, Dies KA, Ledbetter DH, Martin CL, Chung WK, et al. Meta-

- analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genet Med*. 2019 Nov;21(11):2413–21.
139. Thevenon J, Duffourd Y, Masurel-Paulet A, Lefebvre M, Feillet F, El Chehadeh-Djebbar S, et al. Diagnostic odyssey in severe neurodevelopmental disorders: toward clinical whole-exome sequencing as a first-line diagnostic test. *Clin Genet*. 2016 Jun;89(6):700–7.
 140. Wright CF, Fitzgerald TW, Jones WD, Clayton S, McRae JF, Van Kogelenberg M, et al. Genetic diagnosis of developmental disorders in the DDD study: A scalable analysis of genome-wide research data. *Lancet*. 2015;385(9975):1305–14.
 141. Wright CF, McRae JF, Clayton S, Gallone G, Aitken S, FitzGerald TW, et al. Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. *Genet Med*. 2018 Oct;20(10):1216–23.
 142. McRae J, Clayton S, Fitzgerald T, Kaplanis J, Prigmore E, Rajan D, et al. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017 Feb;542(7642):433–8.
 143. Short PJ, McRae JF, Gallone G, Sifrim A, Won H, Geschwind DH, et al. De novo mutations in regulatory elements in neurodevelopmental disorders. *Nature*. 2018 Mar;555(7698):611–6.
 144. Martin HC, Jones WD, McIntyre R, Sanchez-Andrade G, Sanderson M, Stephenson JD, et al. Quantifying the contribution of recessive coding variation to developmental disorders. *Science*. 2018 Dec;362(6419):1161–4.
 145. Zarate YA, Bhoj E, Kaylor J, Li D, Tsurusaki Y, Miyake N, et al. SMARCE1, a rare cause of Coffin–Siris Syndrome: Clinical description of three additional cases. *Am J Med Genet Part A*. 2016;170(8):1967–73.
 146. Smith MJ, O’Sullivan J, Bhaskar SS, Hadfield KD, Poke G, Caird J, et al. Loss-of-function mutations in SMARCE1 cause an inherited disorder of multiple spinal meningiomas. *Nat Genet*. 2013;45(3):295–8.
 147. Banka S, Sayer R, Breen C, Barton S, Pavaine J, Sheppard SE, et al. Genotype-phenotype specificity in Menke-Hennekam syndrome caused by missense variants in exon 30 or 31 of CREBBP. *Am J Med Genet A*. 2019 Jun;179(6):1058–62.
 148. Kleefstra T, Kramer JM, Neveling K, Willemsen MH, Koemans TS, Vissers LELM, et al. Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am J Hum Genet*. 2012;91(1):73–82.
 149. Koemans TS, Kleefstra T, Chubak MC, Stone MH, Reijnders MRF, de Munnik S, et al. Functional convergence of histone methyltransferases EHMT1 and KMT2C involved in intellectual disability and autism spectrum disorder. *PLoS Genet*. 2017 Oct;13(10):e1006864.
 150. Lopes F, Barbosa M, Ameur A, Soares G, de Sa J, Dias AI, et al. Identification of novel genetic causes of Rett syndrome-like phenotypes. *J Med Genet*. 2016 Mar;53(3):190–9.
 151. Vidal S, Brandi N, Pacheco P, Maynou J, Fernandez G, Xiol C, et al. The most recurrent monogenic disorders that overlap with the phenotype of Rett syndrome. *Eur J Paediatr Neurol*. 2019 Jul;23(4):609–620.
 152. Wang J, Zhang Q, Chen Y, Yu S, Wu X, Bao X. Rett and Rett-like syndrome: Expanding the genetic spectrum to KIF1A and GRIN1 gene. *Mol Genet genomic Med*. 2019 Nov;7(11):e968.
 153. Berger SI, Ciccone C, Simon KL, Malicdan MC, Vilboux T, Billington C, et al. Exome analysis of Smith–Magenis-like syndrome cohort identifies de novo likely pathogenic variants. *Hum Genet*. 2017;136(4):409–20.
 154. Pizzo L, Jensen M, Polyak A, Rosenfeld JA, Mannik K, Krishnan A, et al. Rare variants in the genetic background modulate cognitive and developmental phenotypes in individuals carrying disease-associated variants. *Genet Med*. 2019 Apr;21(4):816–25.
 155. Kousi M, Katsanis N. Genetic modifiers and oligogenic inheritance. *Cold Spring Harb*

- Perspect Med. 2015 Jun;5(6):a017145.
156. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–23.
 157. Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. *Hum Mutat*. 2008 Nov;29(11):1282–91.
 158. Li Q, Wang K. InterVar: Clinical Interpretation of Genetic Variants by the 2015 ACMG-AMP Guidelines. *Am J Hum Genet*. 2017;100(2):267–80.
 159. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Aguilera MA, Meyer R, et al. VarSome: the human genomic variant search engine. *Bioinformatics*. 2018;(October):1–3.
 160. Lek M, Karczewski KJ, Minikel E V, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016 Aug;536(7616):285–91.
 161. Hijikata A, Tsuji T, Shionyu M, Shirai T. Decoding disease-causing mechanisms of missense mutations from supramolecular structures. *Sci Rep*. 2017 Aug;7(1):8541.
 162. Veitia RA, Caburet S, Birchler JA. Mechanisms of Mendelian dominance. *Clin Genet*. 2018 Mar;93(3):419–28.
 163. Lecoquierre F, Duffourd Y, Vitobello A, Bruel A-L, Urteaga B, Coubes C, et al. Variant recurrence in neurodevelopmental disorders: the use of publicly available genomic data identifies clinically relevant pathogenic missense variants. *Genet Med*. 2019 Nov;21(11):2504-2511.
 164. Lykke-Andersen S, Jensen TH. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. *Nat Rev Mol Cell Biol*. 2015 Nov;16(11):665–77.
 165. Nickless A, Bailis JM, You Z. Control of gene expression through the nonsense-mediated RNA decay pathway. *Cell Biosci*. 2017;7:26.
 166. Le Hir H, Izaurralde E, Maquat LE, Moore MJ. The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J*. 2000 Dec;19(24):6860–9.
 167. Singh G, Jakob S, Kleedehn MG, Lykke-Andersen J. Communication with the exon-junction complex and activation of nonsense-mediated decay by human Upf proteins occur in the cytoplasm. *Mol Cell*. 2007 Sep;27(5):780–92.
 168. Okada-Katsuhata Y, Yamashita A, Kutsuzawa K, Izumi N, Hirahara F, Ohno S. N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5:SMG-7 during NMD. *Nucleic Acids Res*. 2012 Feb;40(3):1251–66.
 169. Lejeune F, Li X, Maquat LE. Nonsense-mediated mRNA decay in mammalian cells involves decapping, deadenylating, and exonucleolytic activities. *Mol Cell*. 2003 Sep;12(3):675–87.
 170. Nagy E, Maquat LE. A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem Sci*. 1998 Jun;23(6):198–9.
 171. Maquat LE, Li X. Mammalian heat shock p70 and histone H4 transcripts, which derive from naturally intronless genes, are immune to nonsense-mediated decay. *RNA*. 2001 Mar;7(3):445–56.
 172. Hoek TA, Khuperkar D, Lindeboom RGH, Sonneveld S, Verhagen BMP, Boersma S, et al. Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. *Mol Cell*. 2019 Jul;75(2):324-339.e11.
 173. Lindeboom RGH, Supek F, Lehner B. The rules and impact of nonsense-mediated mRNA decay in human cancers. *Nat Genet*. 2016 Oct;48(10):1112–8.
 174. Lindeboom RGH, Vermeulen M, Lehner B, Supek F. The impact of nonsense-mediated mRNA decay on genetic disease, gene editing and cancer immunotherapy. *Nat Genet*. 2019 Nov;51(11):1645–51.
 175. Nguyen LS, Wilkinson MF, Gez J. Nonsense-mediated mRNA decay: inter-individual variability and human disease. *Neurosci Biobehav Rev*. 2014 Oct;46 Pt 2:175–86.

176. Wang D, Zavadil J, Martin L, Parisi F, Friedman E, Levy D, et al. Inhibition of nonsense-mediated RNA decay by the tumor microenvironment promotes tumorigenesis. *Mol Cell Biol*. 2011 Sep;31(17):3670–80.
177. Gardner LB. Nonsense-mediated RNA decay regulation by cellular stress: implications for tumorigenesis. *Mol Cancer Res*. 2010 Mar;8(3):295–308.
178. Xu X, Zhang L, Tong P, Xun G, Su W, Xiong Z, et al. Exome sequencing identifies UPF3B as the causative gene for a Chinese non-syndrome mental retardation pedigree. *Clin Genet*. 2013 Jun;83(6):560–4.
179. Tarpey PS, Raymond FL, Nguyen LS, Rodriguez J, Hackett A, Vandeleur L, et al. Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. *Nat Genet*. 2007 Sep;39(9):1127–33.
180. Nguyen LS, Kim H-G, Rosenfeld JA, Shen Y, Gusella JF, Lacassie Y, et al. Contribution of copy number variants involving nonsense-mediated mRNA decay pathway genes to neurodevelopmental disorders. *Hum Mol Genet*. 2013 May;22(9):1816–25.
181. Tarlunganu DC, Novarino G. Genomics in neurodevelopmental disorders: an avenue to personalized medicine. *Exp Mol Med*. 2018 Aug;50(8):100.
182. Chamberlain SJ, Chen P-F, Ng KY, Bourgois-Rocha F, Lemtiri-Chlieh F, Levine ES, et al. Induced pluripotent stem cell models of the genomic imprinting disorders Angelman and Prader-Willi syndromes. *Proc Natl Acad Sci U S A*. 2010 Oct;107(41):17668–73.
183. Fatehullah A, Tan SH, Barker N. Organoids as an in vitro model of human development and disease. *Nat Cell Biol*. 2016 Mar;18(3):246–54.
184. Stefl S, Nishi H, Petukh M, Panchenko AR, Alexov E. Molecular mechanisms of disease-causing missense mutations. *J Mol Biol*. 2013 Nov;425(21):3919–36.
185. Zlotogora J. Penetrance and expressivity in the molecular age. *Genet Med*. 2003;5(5):347–52.
186. Raraigh KS, Han ST, Davis E, Evans TA, Pellicore MJ, McCague AF, et al. Functional Assays Are Essential for Interpretation of Missense Variants Associated with Variable Expressivity. *Am J Hum Genet*. 2018 Jun;102(6):1062–77.
187. Luo J, Zhou T, You X, Zi Y, Li X, Wu Y, et al. Assessing concordance among human, in silico predictions and functional assays on genetic variant classification. *Bioinformatics*. 2019 May;35(24):5163–70.
188. Seifi M, Walter MA. Accurate prediction of functional, structural, and stability changes in PITX2 mutations using in silico bioinformatics algorithms. *PLoS One*. 2018;13(4):e0195971.
189. Scarabelli G, Grant BJ. Mapping the structural and dynamical features of kinesin motor domains. *PLoS Comput Biol*. 2013;9(11):e1003329.
190. Wang W, Cao L, Wang C, Gigant B, Knossow M. Kinesin, 30 years later: Recent insights from structural studies. *Protein Sci*. 2015 Jul;24(7):1047–56.
191. Hirokawa N, Pfister KK, Yorifuji H, Wagner MC, Brady ST, Bloom GS. Submolecular domains of bovine brain kinesin identified by electron microscopy and monoclonal antibody decoration. *Cell*. 1989 Mar;56(5):867–78.
192. Hirokawa N, Takemura R. Molecular motors and mechanisms of directional transport in neurons. *Nat Rev Neurosci*. 2005 Mar;6(3):201–14.
193. Hirokawa N. Kinesin and dynein superfamily proteins and the mechanism of organelle transport. *Science*. 1998 Jan;279(5350):519–26.
194. Miki H, Okada Y, Hirokawa N. Analysis of the kinesin superfamily: insights into structure and function. *Trends Cell Biol*. 2005 Sep;15(9):467–76.
195. Hirokawa N, Noda Y. Intracellular transport and kinesin superfamily proteins, KIFs: structure, function, and dynamics. *Physiol Rev*. 2008 Jul;88(3):1089–118.
196. Kikkawa M, Sablin EP, Okada Y, Yajima H, Fletterick RJ, Hirokawa N. Switch-based mechanism of kinesin motors. *Nature*. 2001 May;411(6836):439–45.
197. Guo Y, Chen Y, Yang M, Xu X, Lin Z, Ma J, et al. A Rare KIF1A Missense Mutation Enhances

- Synaptic Function and Increases Seizure Activity. *Front Genet.* 2020;11:61.
198. Hammond JW, Cai D, Blasius TL, Li Z, Jiang Y, Jih GT, et al. Mammalian Kinesin-3 motors are dimeric in vivo and move by processive motility upon release of autoinhibition. *PLoS Biol.* 2009 Mar;7(3):e72.
 199. Lee J-R, Shin H, Ko J, Choi J, Lee H, Kim E. Characterization of the movement of the kinesin motor KIF1A in living cultured neurons. *J Biol Chem.* 2003 Jan;278(4):2624–9.
 200. Klopfenstein DR, Tomishige M, Stuurman N, Vale RD. Role of phosphatidylinositol(4,5)biphosphate organization in membrane transport by the Unc104 kinesin motor. *Cell.* 2002 May;109(3):347–58.
 201. Dietrich KA, Sindelar C V, Brewer PD, Downing KH, Cremona CR, Rice SE. The kinesin-1 motor protein is regulated by a direct interaction of its head and tail. *Proc Natl Acad Sci U S A.* 2008 Jul;105(26):8938–43.
 202. Imanishi M, Endres NF, Gennerich A, Vale RD. Autoinhibition regulates the motility of the *C. elegans* intraflagellar transport motor OSM-3. *J Cell Biol.* 2006 Sep;174(7):931–7.
 203. Espeut J, Gaussen A, Bieling P, Morin V, Prieto S, Fesquet D, et al. Phosphorylation relieves autoinhibition of the kinetochore motor Cenp-E. *Mol Cell.* 2008 Mar;29(5):637–43.
 204. Rashid DJ, Bononi J, Triplet BP, Hodges RS, Pierce DW. Monomeric and dimeric states exhibited by the kinesin-related motor protein KIF1A. *J Pept Res.* 2005 Jun;65(6):538–49.
 205. Okada Y, Yamazaki H, Sekine-Aizawa Y, Hirokawa N. The neuron-specific kinesin superfamily protein KIF1A is a unique monomeric motor for anterograde axonal transport of synaptic vesicle precursors. *Cell.* 1995 Jun;81(5):769–80.
 206. Soppina V, Norris SR, Dizaji AS, Kortus M, Veatch S, Peckham M, et al. Dimerization of mammalian kinesin-3 motors results in superprocessive motion. *Proc Natl Acad Sci U S A.* 2014 Apr;111(15):5562–7.
 207. McVicker DP, Awe AM, Richters KE, Wilson RL, Cowdrey DA, Hu X, et al. Transport of a kinesin-cargo pair along microtubules into dendritic spines undergoing synaptic plasticity. *Nat Commun.* 2016 Sep;7:12741.
 208. Kondo M, Takei Y, Hirokawa N. Motor protein KIF1A is essential for hippocampal synaptogenesis and learning enhancement in an enriched environment. *Neuron.* 2012 Feb;73(4):743–57.
 209. Jamuar SS, Walsh CA. Somatic mutations in cerebral cortical malformations. Vol. 371, *The New England journal of medicine.* United States; 2014. p. 2038.
 210. Carabona A, Hu DJ-K, Vallee RB. KIF1A inhibition immortalizes brain stem cells but blocks BDNF-mediated neuronal migration. *Nat Neurosci.* 2016 Feb;19(2):253–62.
 211. Tanaka Y, Niwa S, Dong M, Farkhondeh A, Wang L, Zhou R, et al. The Molecular Motor KIF1A Transports the TrkA Neurotrophin Receptor and Is Essential for Sensory Neuron Survival and Function. *Neuron.* 2016 Jun;90(6):1215–29.
 212. Esmaeeli Nieh S, Madou MRZ, Sirajuddin M, Fregeau B, Mcknight D, Lexa K, et al. De novo mutations in KIF1A cause progressive encephalopathy and brain atrophy. *Ann Clin Transl Neurol.* 2015;2(6):623–35.
 213. Lee JR, Srour M, Kim D, Hamdan FF, Lim SH, Brunel-Guitton C, et al. De novo mutations in the motor domain of KIF1A cause cognitive impairment, spastic paraparesis, axonal neuropathy, and cerebellar atrophy. *Hum Mutat.* 2015;36(1):69–78.
 214. Chiba K, Takahashi H, Chen M, Obinata H, Arai S, Hashimoto K, et al. Disease-associated mutations hyperactivate KIF1A motility and anterograde axonal transport of synaptic vesicle precursors. *Proc Natl Acad Sci U S A.* 2019 Sep;116(37):18429–34.
 215. Bell L, Wittkowski A, Hare DJ. Movement Disorders and Syndromic Autism: A Systematic Review. *J Autism Dev Disord.* 2019 Jan;49(1):54–67.
 216. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2010 Mar;26(5):589–95.
 217. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best

- practices pipeline. *Curr Protoc Bioinforma*. 2013;43:11.10.1-33.
218. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010 Sep;38(16):e164.
 219. Desmet F-O, Hamroun D, Lalande M, Collod-Bérout G, Claustres M, Bérout C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res*. 2009;37(9):e67–e67.
 220. ClinVar [Internet]. [cited 2019 Nov 17]. Available from: <https://www.ncbi.nlm.nih.gov/clinvar/>
 221. DECIPHER v9.30: Mapping the clinical genome [Internet]. [cited 2019 Nov 17]. Available from: <https://decipher.sanger.ac.uk/>
 222. OMIM - Online Mendelian Inheritance in Man [Internet]. [cited 2019 Nov 14]. Available from: <https://www.omim.org/>
 223. The Deciphering Developmental Disorders (DDD) Study - DECIPHER v9.30 [Internet]. [cited 2019 Nov 14]. Available from: <https://decipher.sanger.ac.uk/ddd#ddgenes>
 224. Kochinke K, Zweier C, Nijhof B, Fencikova M, Cizek P, Honti F, et al. Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am J Hum Genet*. 2016 Jan;98(1):149–64.
 225. Home - PubMed - NCBI [Internet]. [cited 2019 Nov 14]. Available from: <https://www.ncbi.nlm.nih.gov/pubmed/>
 226. Fromer M, Moran JL, Chambert K, Banks E, Bergen SE, Ruderfer DM, et al. Discovery and statistical genotyping of copy-number variation from whole-exome sequencing depth. *Am J Hum Genet*. 2012 Oct;91(4):597–607.
 227. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 2001 Dec;25(4):402–8.
 228. Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol*. 1990;185:60–89.
 229. Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol*. 1986 May;189(1):113–30.
 230. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019 Jan;47(D1):D607–13.
 231. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, et al. The STRING database in 2017: Quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res*. 2017;45(D1):D362–8.
 232. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res*. 2010;38(SUPPL. 2):214–20.
 233. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol*. 2008;9 Suppl 1(Suppl 1):S4–S4.
 234. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res*. 2018 Jan;46(D1):D649–55.
 235. Fabregat A, Sidiropoulos K, Viteri G, Forner O, Marin-Garcia P, Arnau V, et al. Reactome pathway analysis: a high-performance in-memory approach. *BMC Bioinformatics*. 2017 Mar;18(1):142.
 236. Mele M, Ferreira PG, Reverter F, DeLuca DS, Monlong J, Sammeth M, et al. Human genomics. The human transcriptome across tissues and individuals. *Science*. 2015 May;348(6235):660–5.
 237. Thul PJ, Akesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, et al. A subcellular map of the human proteome. *Science*. 2017 May;356(6340):eaal3321.
 238. Hu H, Haas S, Chelly J, Esch H Van, Raynaud M, Brouwer A de, et al. X-exome sequencing

- of 405 unresolved families identifies seven novel intellectual disability genes. *Mol Psychiatry*. 2015;21(1):133–48.
239. Ouyang J, Shi Y, Valin A, Xuan Y, Gill G. Direct binding of CoREST1 to SUMO-2/3 contributes to gene-specific repression by the LSD1/CoREST1/HDAC complex. *Mol Cell*. 2009 Apr;34(2):145–54.
 240. Castle CD, Cassimere EK, Lee J, Denicourt C. Las1L is a nucleolar protein required for cell proliferation and ribosome biogenesis. *Mol Cell Biol*. 2010 Sep;30(18):4404–14.
 241. Fanis P, Gillemans N, Aghajani-refah A, Pourfarzad F, Demmers J, Esteghamat F, et al. Five friends of methylated chromatin target of protein-arginine-methyltransferase[prmt]-1 (chttop), a complex linking arginine methylation to desumoylation. *Mol Cell Proteomics*. 2012 Nov;11(11):1263–73.
 242. Morleo M, Franco B. Microphthalmia with Linear Skin Defects Syndrome [Internet]. GeneReviews®. University of Washington, Seattle; 1993 [cited 2019 Oct 24]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20301767>
 243. Morleo M, Franco B. Dosage compensation of the mammalian X chromosome influences the phenotypic variability of X-linked dominant male-lethal disorders. *J Med Genet*. 2008 Jul;45(7):401–8.
 244. Sajan SA, Jhangiani SN, Muzny DM, Gibbs RA, Lupski JR, Glaze DG, et al. Enrichment of mutations in chromatin regulators in people with Rett syndrome lacking mutations in MECP2. *Genet Med*. 2017 Jan;19(1):13–9.
 245. Gonzaga-Jauregui C, Harel T, Gambin T, Kousi M, Griffin LB, Francescato L, et al. Exome Sequence Analysis Suggests that Genetic Burden Contributes to Phenotypic Variability and Complex Neuropathy. *Cell Rep*. 2015 Aug;12(7):1169–83.
 246. Du Y, Li Z, Liu Z, Zhang N, Wang R, Li F, et al. Nonrandom occurrence of multiple de novo coding variants in a proband indicates the existence of an oligogenic model in autism. *Genet Med*. 2019 Jul; 22(1):170–180.
 247. Grozdanov PN, Masoumzadeh E, Kalscheuer VM, Bienvenu T, Billuart P, Delrue M-A, et al. A Point Mutation in the RNA Recognition Motif of CSTF2 Associated with Intellectual Disability in Humans Causes Defects in 3' End Processing. *bioRxiv*. 2020 Jan 1;2020.01.02.893107.
 248. MacDonald CC. Tissue-specific mechanisms of alternative polyadenylation: Testis, brain, and beyond (2018 update). *Wiley Interdiscip Rev RNA*. 2019 Jul;10(4):e1526.
 249. Anthony K, Gallo J-M. Aberrant RNA processing events in neurological disorders. *Brain Res*. 2010 Jun;1338:67–77.
 250. Wang L, Yi R. 3'UTRs take a long shot in the brain. *Bioessays*. 2014 Jan;36(1):39–45.
 251. Harris JC, Martinez JM, Grozdanov PN, Bergeson SE, Grammas P, MacDonald CC. The Cstf2t Polyadenylation Gene Plays a Sex-Specific Role in Learning Behaviors in Mice. *PLoS One*. 2016;11(11):e0165976.
 252. The UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*. 2019 Jan;47(D1):D506–15.
 253. Colwill K, Graslund S. A roadmap to generate renewable protein binders to the human proteome. *Nat Methods*. 2011 May;8(7):551–8.
 254. Lund RJ, Osmala M, Malonzo M, Lukkarinen M, Leino A, Salmi J, et al. Atopic asthma after rhinovirus-induced wheezing is associated with DNA methylation change in the SMAD3 gene promoter. *Allergy*. 2018 Aug;73(8):1735–40.
 255. Harel T, Yoon WH, Garone C, Gu S, Coban-Akdemir Z, Eldomery MK, et al. Recurrent De Novo and Biallelic Variation of ATAD3A, Encoding a Mitochondrial Membrane Protein, Results in Distinct Neurological Syndromes. *Am J Hum Genet*. 2016 Oct;99(4):831–45.
 256. Wiszniewski W, Hunter J V, Hanchard NA, Willer JR, Shaw C, Tian Q, et al. TM4SF20 ancestral deletion and susceptibility to a pediatric disorder of early language delay and cerebral white matter hyperintensities. *Am J Hum Genet*. 2013 Aug;93(2):197–210.
 257. Stosser MB, Lindy AS, Butler E, Retterer K, Piccirillo-Stosser CM, Richard G, et al. High

- frequency of mosaic pathogenic variants in genes causing epilepsy-related neurodevelopmental disorders. *Genet Med*. 2018 Apr;20(4):403–10.
258. Myers CT, Hollingsworth G, Muir AM, Schneider AL, Thuesmann Z, Knupp A, et al. Parental Mosaicism in “De Novo” Epileptic Encephalopathies. Vol. 378, *The New England journal of medicine*. United States; 2018. p. 1646–8.
 259. Jimenez J, Herrera DA, Vargas SA, Montoya J, Castillo M. beta-Tubulinopathy caused by a mutation of the TUBB2B gene: magnetic resonance imaging findings of the brain. *Neuroradiol J*. 2019 Apr;32(2):148–50.
 260. Wang H, Li S, Li S, Jiang N, Guo J, Zhang W, et al. De Novo Mutated TUBB2B Associated Pachygyria Diagnosed by Medical Exome Sequencing and Long-Range PCR. *Fetal Pediatr Pathol*. 2019 Feb;38(1):63–71.
 261. Fitzgerald1 TW, Gerety SS, Jones WD, Kogelenberg M van, King DA, McRae J, et al. Large-scale discovery of novel genetic causes of developmental disorders. *Nature*. 2014;519(7542):223–8.
 262. Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet*. 2011 Sep 27;12(11):745–55.
 263. Coe BP, Stessman HAF, Sulovari A, Geisheker MR, Bakken TE, Lake AM, et al. Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. *Nat Genet*. 2018/12/17. 2019 Jan;51(1):106–16.
 264. Monroe GR, Frederix GW, Savelberg SMC, Vries TI de, Duran KJ, Smagt JJ van der, et al. Effectiveness of whole-exome sequencing and costs of the traditional diagnostic trajectory in children with intellectual disability. *Genet Med*. 2016;18(9):949–56.
 265. Clark MM, Stark Z, Farnaes L, Tan TY, White SM, Dimmock D, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *npj Genomic Med*. 2018;3(1):1–10.
 266. Mignot C, von Stulpnagel C, Nava C, Ville D, Sanlaville D, Lesca G, et al. Genetic and neurodevelopmental spectrum of SYNGAP1-associated intellectual disability and epilepsy. *J Med Genet*. 2016 Aug;53(8):511–22.
 267. Laskowski RA, Thornton JM, Tyagi N, Johnson D, McWilliam C, Kinning E, et al. Integrating population variation and protein structural analysis to improve clinical interpretation of missense variation: application to the WD40 domain. *Hum Mol Genet*. 2016;25(5):927–35.
 268. Bengani H, Handley M, Alvi M, Ibitoye R, Lees M, Lynch SA, et al. Clinical and molecular consequences of disease-associated de novo mutations in SATB2. *Genet Med*. 2017;19(8):900–8.
 269. Sands TT, Miceli F, Lesca G, Beck AE, Sadleir LG, Arrington DK, et al. Autism and developmental disability caused by KCNQ3 gain-of-function variants. *Ann Neurol*. 2019 Jun; 86(2):181-192.
 270. Syrbe S, Harms FL, Parrini E, Montomoli M, Mutze U, Helbig KL, et al. Delineating SPTAN1 associated phenotypes: from isolated epilepsy to encephalopathy with progressive brain atrophy. *Brain*. 2017 Sep;140(9):2322–36.
 271. Balasubramanian M, Willoughby J, Fry AE, Weber A, Firth H V, Deshpande C, et al. Delineating the phenotypic spectrum of Bainbridge-Ropers syndrome: 12 new patients with de novo , heterozygous, loss-of-function mutations in ASXL3 and review of published literature. *J Med Genet*. 2017;54(8):537–43.
 272. Johannesen KM, Gardella E, Linnankivi T, Courage C, Martin A de Saint, Lehesjoki A-E, et al. Defining the phenotypic spectrum of SLC6A1 mutations. *Epilepsia*. 2018;59(2):389–402.
 273. Kurki MI, Saarentaus E, Pietiläinen O, Gormley P, Lal D, Kerminen S, et al. Contribution of rare and common variants to intellectual disability in a sub-isolate of Northern Finland.

- Nat Commun. 2019 Jan;10(1):410.
274. Reichenberg A, Cederlöf M, McMillan A, Trzaskowski M, Kapra O, Fruchter E, et al. Discontinuity in the genetic and environmental causes of the intellectual disability spectrum. *Proc Natl Acad Sci U S A*. 2015/12/28. 2016 Jan;113(4):1098–103.
 275. Rauch A, Wieczorek D, Graf E, Wieland T, Ende S, Schwarzmayr T, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet (London, England)*. 2012 Nov;380(9854):1674–82.
 276. Geisheker MR, Heymann G, Wang T, Coe BP, Turner TN, Stessman HAF, et al. Hotspots of missense mutation identify neurodevelopmental disorder genes and functional domains. *Nat Neurosci*. 2017 Aug;20(8):1043–51.
 277. Lelieveld SH, Wiel L, Venselaar H, Pfundt R, Vriend G, Veltman JA, et al. Spatial Clustering of de Novo Missense Mutations Identifies Candidate Neurodevelopmental Disorder-Associated Genes. *Am J Hum Genet*. 2017 Sep;101(3):478–84.
 278. Wang G, Zhang J, Moskophidis D, Mivechi NF. Targeted disruption of the heat shock transcription factor (hsf)-2 gene results in increased embryonic lethality, neuronal defects, and reduced spermatogenesis. *Genesis*. 2003;36(1):48–61.
 279. Jin X, Eroglu B, Moskophidis D, Mivechi NF. Targeted deletion of Hsf1, 2, and 4 genes in mice. *Methods Mol Biol*. 2011;787:1–20.
 280. Rusten TE, Vaccari T, Stenmark H. Shaping development with ESCRTs. *Nat Cell Biol*. 2011 Dec;14(1):38–45.
 281. Lee J-A, Gao F-B. Neuronal Functions of ESCRTs. *Exp Neurobiol*. 2012 Mar;21(1):9–15.
 282. Kaul Z, Chakrabarti O. Endosomal sorting complexes required for ESCRTing cells toward death during neurogenesis, neurodevelopment and neurodegeneration. *Traffic*. 2018;19(7):485–95.
 283. Shoubridge C, Harvey RJ, Dudding-Byth T. IQSEC2 mutation update and review of the female-specific phenotype spectrum including intellectual disability and epilepsy. *Hum Mutat*. 2019 Jan;40(1):5–24.
 284. Balaton BP, Cotton AM, Brown CJ. Derivation of consensus inactivation status for X-linked genes from genome-wide studies. *Biol Sex Differ*. 2015;6:35.
 285. Bruel A-L, Vitobello A, Mau-Them FT, Nambot S, Duffourd Y, Quéré V, et al. 2.5 years' experience of GeneMatcher data-sharing: a powerful tool for identifying new genes responsible for rare diseases. *Genet Med*. 2018/12/19. 2019 Jul;21(7):1657–61.
 286. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat*. 2015 Oct;36(10):928–30.
 287. Berryer MH, Hamdan FF, Klitten LL, Moller RS, Carmant L, Schwartzentruber J, et al. Mutations in SYNGAP1 cause intellectual disability, autism, and a specific form of epilepsy by inducing haploinsufficiency. *Hum Mutat*. 2013 Feb;34(2):385–94.
 288. Zarate YA, Bosanko KA, Caffrey AR, Bernstein JA, Martin DM, Williams MS, et al. Mutation update for the SATB2 gene. *Hum Mutat*. 2019 Apr; 40(8):1013-1029.
 289. Kuechler A, Czeschik JC, Graf E, Grasshoff U, Huffmeier U, Busa T, et al. Bainbridge-Ropers syndrome caused by loss-of-function variants in ASXL3: a recognizable condition. *Eur J Hum Genet*. 2017 Feb;25(2):183–91.
 290. De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*. 2014 Nov;515(7526):209–15.
 291. Koboldt DC, Mihalic Mosher T, Kelly BJ, Sites E, Bartholomew D, Hickey SE, et al. A de novo nonsense mutation in ASXL3 shared by siblings with Bainbridge-Ropers syndrome. *Cold Spring Harb Mol case Stud*. 2018 Jun;4(3):a002410.
 292. Salpietro V, Malintan NT, Llano-Rivas I, Spaeth CG, Efthymiou S, Striano P, et al. Mutations in the Neuronal Vesicular SNARE VAMP2 Affect Synaptic Membrane Fusion and Impair Human Neurodevelopment. *Am J Hum Genet*. 2019 Apr;104(4):721–30.

293. Carvill GL, McMahon JM, Schneider A, Zemel M, Myers CT, Saykally J, et al. Mutations in the GABA Transporter SLC6A1 Cause Epilepsy with Myoclonic-Atonic Seizures. *Am J Hum Genet.* 2015 May;96(5):808–15.
294. Barro-Soria R. Epilepsy-associated mutations in the voltage sensor of KCNQ3 affect voltage dependence of channel opening. *J Gen Physiol.* 2018;151(2):247–57.
295. Magyar I, Colman D, Arnold E, Baumgartner D, Bottani A, Fokstuen S, et al. Quantitative sequence analysis of FBN1 premature termination codons provides evidence for incomplete NMD in leukocytes. *Hum Mutat.* 2009 Sep;30(9):1355–64.
296. Leoyklang P, Suphapeetiporn K, Srichomthong C, Tongkobpetch S, Fietze S, Dorward H, et al. Disorders with similar clinical phenotypes reveal underlying genetic interaction: SATB2 acts as an activator of the UPF3B gene. *Hum Genet.* 2013 Dec;132(12):1383–93.
297. Katoh M. Functional and cancer genomics of ASXL family members. *Br J Cancer.* 2013 Jul;109(2):299–306.
298. Bainbridge MN, Hu H, Muzny DM, Musante L, Lupski JR, Graham BH, et al. De novo truncating mutations in ASXL3 are associated with a novel clinical phenotype with similarities to Bohring-Opitz syndrome. *Genome Med.* 2013;5(2):11.
299. Srivastava A, Ritesh KC, Tsan Y-C, Liao R, Su F, Cao X, et al. De novo dominant ASXL3 mutations alter H2A deubiquitination and transcription in Bainbridge-Ropers syndrome. *Hum Mol Genet.* 2016 Feb;25(3):597–608.
300. Soldovieri MV, Boutry-Kryza N, Milh M, Doummar D, Heron B, Bourel E, et al. Novel KCNQ2 and KCNQ3 mutations in a large cohort of families with benign neonatal epilepsy: first evidence for an altered channel regulation by syntaxin-1A. *Hum Mutat.* 2014 Mar;35(3):356–67.
301. Miceli F, Soldovieri MV, Joshi N, Weckhuysen S, Cooper EC, Tagliatalata M. KCNQ3-Related Disorders [Internet]. *GeneReviews*[®]. University of Washington, Seattle; 1993 [cited 2019 Feb 22]. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/24851285>
302. Kikkawa M, Okada Y, Hirokawa N. 15 A resolution model of the monomeric kinesin motor, KIF1A. *Cell.* 2000 Jan;100(2):241–52.
303. Nitta R, Kikkawa M, Okada Y, Hirokawa N. KIF1A alternately uses two loops to bind microtubules. *Science.* 2004 Jul;305(5684):678–83.
304. Woehlke G, Ruby AK, Hart CL, Ly B, Hom-Booher N, Vale RD. Microtubule interaction site of the kinesin motor. *Cell.* 1997 Jul;90(2):207–16.
305. Pennings M, Schouten MI, van Gaalen J, Meijer RPP, de Bot ST, Kriek M, et al. KIF1A variants are a frequent cause of autosomal dominant hereditary spastic paraplegia. *Eur J Hum Genet.* 2020 Jan;28(1):40–9.
306. Hirokawa N, Tanaka Y. Kinesin superfamily proteins (KIFs): Various functions and their relevance for important phenomena in life and diseases. *Exp Cell Res.* 2015 May;334(1):16–25.
307. Wein T, Wanner KT. Generation of a 3D model for human GABA transporter hGAT-1 using molecular modeling and investigation of the binding of GABA. *J Mol Model.* 2010 Jan;16(1):155–61.
308. Mattison KA, Butler KM, Inglis GAS, Dayan O, Boussidan H, Bhambhani V, et al. SLC6A1 variants identified in epilepsy patients reduce γ -aminobutyric acid transport. *Epilepsia.* 2018 Sep;59(9):e135–41.
309. Wang J, Poliquin S, Mermer F, Eissman J, Delpire E, Wang J, et al. Endoplasmic reticulum retention and degradation of a mutation in SLC6A1 associated with epilepsy and autism. *Mol Brain.* 2020 May;13(1):76.
310. Cai K, Wang J, Eissman J, Wang J, Nwosu G, Shen W, et al. A missense mutation in SLC6A1 associated with Lennox-Gastaut syndrome impairs GABA transporter 1 protein trafficking and function. *Exp Neurol.* 2019 Oct;320:112973.
311. Hernandez JM, Stein A, Behrmann E, Riedel D, Cypionka A, Farsi Z, et al. Membrane fusion intermediates via directional and full assembly of the SNARE complex. *Science.* 2012

- Jun;336(6088):1581–4.
312. Shen C, Rathore SS, Yu H, Gulbranson DR, Hua R, Zhang C, et al. The trans-SNARE-regulating function of Munc18-1 is essential to synaptic exocytosis. *Nat Commun.* 2015 Nov;6:8852.
 313. Deak F, Schoch S, Liu X, Sudhof TC, Kavalali ET. Synaptobrevin is essential for fast synaptic-vesicle endocytosis. *Nat Cell Biol.* 2004 Nov;6(11):1102–8.
 314. Lemattre C, Thevenon J, Duffourd Y, Nambot S, Haquet E, Vuadelle B, et al. TBL1XR1 mutations in Pierpont syndrome are not restricted to the recurrent p.Tyr446Cys mutation. Vol. 176, *American journal of medical genetics. Part A. United States*; 2018. p. 2813–8.
 315. Nishi A, Numata S, Tajima A, Zhu X, Ito K, Saito A, et al. De novo non-synonymous TBL1XR1 mutation alters Wnt signaling activity. *Sci Rep.* 2017 Jun;7(1):2887.
 316. Saitsu H, Tohyama J, Kumada T, Egawa K, Hamada K, Okada I, et al. Dominant-negative mutations in alpha-II spectrin cause West syndrome with severe cerebral hypomyelination, spastic quadriplegia, and developmental delay. *Am J Hum Genet.* 2010 Jun;86(6):881–91.
 317. Aguilera C, Gabau E, Laurie S, Baena N, Derdak S, Capdevila N, et al. Identification of a de novo splicing variant in the Coffin–Siris gene, SMARCE1, in a patient with Angelman-like syndrome. *Mol Genet Genomic Med.* 2018;(July):1–6.
 318. Tsurusaki Y, Okamoto N, Ohashi H, Kosho T, Imai Y, Hibi-Ko Y, et al. Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat Genet.* 2012 Mar;44(4):376–8.
 319. Cushion TD, Dobyns WB, Mullins JGL, Stoodley N, Chung S-K, Fry AE, et al. Overlapping cortical malformations and mutations in TUBB2B and TUBA1A. *Brain.* 2013 Feb;136(Pt 2):536–48.
 320. Byers HM, Jensen DM, Glass IA, Bennett JT. Minimal mosaicism, maximal phenotype: Discordance between clinical and molecular findings in two patients with tuberous sclerosis. *Am J Med Genet C Semin Med Genet.* 2018 Sep;178(3):374–8.
 321. Gunning AC, Strucinska K, Muñoz Oreja M, Parrish A, Caswell R, Stals KL, et al. Recurrent De Novo NAHR Reciprocal Duplications in the ATAD3 Gene Cluster Cause a Neurogenetic Trait with Perturbed Cholesterol and Mitochondrial Metabolism. *Am J Hum Genet.* 2020;106(2):272–279.
 322. Janer A, van Karnebeek CD, Sasarman F, Antonicka H, Al Ghamdi M, Shyr C, et al. RMND1 deficiency associated with neonatal lactic acidosis, infantile onset renal failure, deafness, and multiorgan involvement. *Eur J Hum Genet.* 2015 Oct;23(10):1301–7.
 323. Ng YS, Alston CL, Diodato D, Morris AA, Ulrick N, Kmoch S, et al. The clinical, biochemical and genetic features associated with RMND1 -related mitochondrial disease. *J Med Genet.* 2016;53(11):768–75.
 324. Janer A, Antonicka H, Lalonde E, Nishimura T, Sasarman F, Brown GK, et al. An RMND1 Mutation causes encephalopathy associated with multiple oxidative phosphorylation complex deficiencies and a mitochondrial translation defect. *Am J Hum Genet.* 2012 Oct;91(4):737–43.
 325. Stavropoulos DJ, Merico D, Jobling R, Bowdin S, Monfared N, Thiruvahindrapuram B, et al. Whole Genome Sequencing Expands Diagnostic Utility and Improves Clinical Management in Pediatric Medicine. *NPJ genomic Med.* 2016 Jan; 13;1:15012.
 326. Mathelier A, Shi W, Wasserman WW. Identification of altered cis-regulatory elements in human disease. *Trends Genet.* 2015 Feb;31(2):67–76.
 327. Spielmann M, Mundlos S. Looking beyond the genes: the role of non-coding variants in human disease. *Hum Mol Genet.* 2016 Oct;25(R2):R157–65.
 328. Li JY, Daniels G, Wang J, Zhang X. TBL1XR1 in physiological and pathological states. *Am J Clin Exp Urol.* 2015;3(1):13–23.
 329. Jahn R, Fasshauer D. Molecular machines governing exocytosis of synaptic vesicles. *Nature.* 2012 Oct;490(7419):201–7.

330. Odell AF, Van Helden DF, Scott JL. The spectrin cytoskeleton influences the surface expression and activation of human transient receptor potential channel 4 channels. *J Biol Chem*. 2008 Feb;283(7):4395–407.
331. Wilfert AB, Sulovari A, Turner TN, Coe BP, Eichler EE. Recurrent de novo mutations in neurodevelopmental disorders: properties and clinical implications. *Genome Med*. 2017 Nov;9(1):101.
332. Sajan SA, Jhangiani SN, Muzny DM, Gibbs RA, Lupski JR, Glaze DG, et al. Enrichment of mutations in chromatin regulators in people with Rett syndrome lacking mutations in MECP2. *Genet Med*. 2017;19(1):13–9.
333. Rylaarsdam L, Guemez-Gamboa A. Genetic Causes and Modifiers of Autism Spectrum Disorder. *Front Cell Neurosci*. 2019;13:385.
334. Lehner B. Molecular mechanisms of epistasis within and between genes. *Trends Genet*. 2011 Aug;27(8):323–31.
335. Parker MJ, Magee AC, Maystadt I, Benoit V, study D, FitzPatrick DR, et al. De novo, heterozygous, loss-of-function mutations in SYNGAP1 cause a syndromic form of intellectual disability. *Am J Med Genet Part A*. 2015;167(10):2231–7.



Annex III

Novel intragenic deletions within the *UBE3A* gene in two unrelated patients with Angelman syndrome: case report and review of the literature

Cinthia Aguilera, Marina Viñas-Jornet, Neus Baena, Elisabeth Gabau, Concepción Fernández,
Nuria Capdevila, Sanja Cirkovic, Adrijan Sarajlija, Danijela Radivojevic, Anna Ruiz, Miriam Guitart.

BMC Medical Genetics 2017 18:137

DOI: [10.1186/s12881-017-0500-x](https://doi.org/10.1186/s12881-017-0500-x)


Les delecions intragèniques en el gen *UBE3A* són molt poc freqüents, només hi ha 9 casos descrits en la literatura. En aquest article es descriuen dues pacients que presenten un fenotip SA, a les quals se li va realitzar la tècnica de MS-MLPA (*Methylation Sensitive-Multiplex Ligation-dependent Amplification Probe*) com a primera eina diagnòstica. En tots dos casos el patró de metilació de la regió 15q11.2-q13 era normal, però l'anàlisi de dosi mostrava delecions intragèniques en el gen *UBE3A*. Per determinar si hi havia més exons delecionats es va realitzar la MLPA específica del gen *UBE3A*. En el pacient 1 es va confirmar que la delecio només incloïa l'exó 2 que conté la G del codó d'inici de traducció (ATG), per tant, no es produiria proteïna. En la pacient 12 la delecio incloïa els exons 9 i 10 que té com a conseqüència la delecio de 88 aminoàcids (p.Arg765_Leu865del88) del domini catalític ubiqüitina lligasa (HECT). En ambdós casos, l'estudi dels progenitors va indicar que es tractaven de delecions *de novo*. Es proposa la inclusió de la MLPA del gen *UBE3A* en l'algoritme diagnòstic de la SA i la seva realització en els casos en que presenten un fenotip SA, però el patró de metilació de la regió 15q11.2-q13 és normal i no s'han identificat variants patogèniques/probablement patogèniques en la seqüenciació del gen *UBE3A*.

CASE REPORT

Open Access



Novel intragenic deletions within the *UBE3A* gene in two unrelated patients with Angelman syndrome: case report and review of the literature

Cinthia Aguilera¹, Marina Viñas-Jornet¹, Neus Baena¹, Elisabeth Gabau², Concepción Fernández², Nuria Capdevila², Sanja Cirkovic³, Adrijan Sarajlija^{4,5}, Marijana Miskovic³, Danijela Radivojevic³, Anna Ruiz^{1*}  and Miriam Guitart^{1*}

Abstract

Background: Patients with Angelman syndrome (AS) are affected by severe intellectual disability with absence of speech, distinctive dysmorphic craniofacial features, ataxia and a characteristic behavioral phenotype. AS is caused by the lack of expression in neurons of the *UBE3A* gene, which is located in the 15q11.2-q13 imprinted region. Functional loss of *UBE3A* is due to 15q11.2-q13 deletion, mutations in the *UBE3A* gene, paternal uniparental disomy and genomic imprinting defects.

Case presentation: We report here two patients with clinical features of AS referred to our hospital for clinical follow-up and genetic diagnosis. Methylation Specific-Multiplex Ligation-Dependent Probe Amplification (MS-MLPA) of the 15q11.2-q13 region was carried out in our laboratory as the first diagnostic tool detecting two novel *UBE3A* intragenic deletions. Subsequently, the MLPA P336-A2 kit was used to confirm and determine the size of the *UBE3A* deletion in the two patients. A review of the clinical features of previously reported patients with whole *UBE3A* gene or partial intragenic deletions is presented here together with these two new patients.

Conclusion: Although rare, *UBE3A* intragenic deletions may represent a small fraction of AS patients without a genetic diagnosis. Testing for *UBE3A* intragenic exonic deletions should be performed in those AS patients with a normal methylation pattern and no mutations in the *UBE3A* gene.

Keywords: Angelman syndrome (AS), *UBE3A*, Intragenic deletions, MLPA

Background

Angelman syndrome (AS) is a neurogenetic disorder characterized by a severe intellectual disability with absence of speech, distinctive dysmorphic craniofacial features such as microcephaly with flat occiput and occipital groove, wide mouth, ataxia related neurological problems and/or tremor in the limbs and seizures with specific EEG pattern abnormalities that persist into adulthood. The behavioural phenotype is characterized by happy disposition, hyperactivity, attention deficit and frequent disruption of sleep cycles. Other clinical

problems are gastrointestinal difficulties with feeding problems, gastroesophageal reflux, constipation, scoliosis, and an increased sensitivity to heat. Its prevalence is about 1/15000 births [1–3].

The cause of AS is the loss of function in neurons of the ubiquitin protein ligase E6-AP (E6-Associated Protein) encoded by the *UBE3A* gene, which is located on chromosome 15q11.2-q13 region. This region includes a cluster of genes under imprinting control which show differential expression depending on the parental origin, in a tissue-specific manner. *UBE3A* shows a specific expression of the maternal allele in neurons while the paternal allele is silenced by the synthesis of an antisense transcript from the *SNURF-SNPRN* gene

* Correspondence: aruizn@tauli.cat; mguitart@tauli.cat

¹Genetics Laboratory, UDIAT-Centre Diagnòstic, Parc Taulí Hospital Universitari, Institut d'Investigació i Innovació Parc Taulí I3PT, Universitat Autònoma de Barcelona, Parc del Taulí 1, 08208 Barcelona, Sabadell, Spain
Full list of author information is available at the end of the article



© The Author(s). 2017 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

(*UBE3A-ATS*). The absence of the specific maternal expression of *UBE3A* is the main cause of AS [4].

The functional loss of *UBE3A* in the maternal allele is due to different genetic mechanisms: (i) deletion of the region 15q11.2-q13 in 70–75% of patients, (ii) mutations in the *UBE3A* gene in 10–15% of the patients, (iii) paternal uniparental disomy in 1–3% of patients, and (iv) genomic imprinting defects in 2–4% of the patients. The genetic cause remains unknown in approximately 10% of patients presenting clinical features characteristic of AS [4, 5]. There are also some reports that describe a proportion of AS patients who harbor microdeletions of the *UBE3A* gene [6–10].

The *UBE3A* gene spans 120 kb of genomic DNA and consists of 10 exons (NM_130838.1). The ubiquitin protein ligase E6-AP encoded by the *UBE3A* gene is necessary for the ubiquitination of proteins targeted for degradation. E6-AP belongs to the HECT (homologous to E6-AP COOH terminus) class of E3 enzymes that share a 40 kDa COOH-terminal catalytic domain. This domain is encoded by exons 3–10 [11]. Exons 1 to 5 encode a steroid co-activation region which has several LXXLL motifs that are known to be receptor interacting motifs [12].

Mutations in the *UBE3A* gene are widely distributed and have been detected throughout all regions of the gene [13, 14]. Rare intragenic deletions and duplications may represent a higher percentage of mutations than expected, because these alterations escape the routine screening [15].

Here, we report two patients with AS who present two novel intragenic deletions within the *UBE3A* gene.

Case presentation

Patient 1 is a 5 year-old girl born at 37 weeks of gestation. Birth weight was 2850 g. Neither perinatal problems nor hypotonia was reported. Initially, she didn't show feeding problems but when solid food was introduced into her diet, she had difficulties to swallow. Psychomotor development was delayed. She sat unsupported at 9 months, walked independently at 24 months and pronounced only two words which she lost later. She came to the clinical geneticist when she was 2-year-8-month old. Phenotypic features included relative microcephaly (47 cm, -1.75 SD), small and wide-spaced teeth and protruding tongue. Her weight was 12 Kg (25th percentile) and her height was 80 cm (50th percentile). Neurological examination detected severe intellectual disability, ataxia of gait, receptive and non-verbal communication skills higher than verbal ones, hyperreflexia of the lower extremities, tremulous movement of limbs and frequent drooling. She also had abnormal sleep-wake cycles and by the age of 20 months she developed seizures. EEG showed generalized slow wave

activity with paroxysmal activity. The behavioural phenotype included frequent laughter, happy demeanour, easily excitable personality, hyperactive behaviour, attention deficit and exploration of objects throughout the mouth. She also presented attraction to water and an increased sensitivity to heat (Table 1).

Methylation PCR and *UBE3A* sequencing were carried out in another centre with negative results. MS-MLPA analysis was performed in our laboratory using a specific kit, SALSA MS-MLPA ME028-B2 Prader Willi/Angelman, (MRC Holland, Amsterdam, The Netherlands). This kit contains 46 probes, 32 of which are specific for sequences in or close to the PWS/AS critical region on 15q11.2-q13 which can be used to detect copy number changes in this region. As a control for copy number changes, 14 probes outside the PWS/AS region are included. Among the PWS/AS specific probes, seven probes are methylation-sensitive and contain a *HhaI* restriction site. Five probes are located within the *UBE3A* gene corresponding to exons 1, 2, 3, 4 and 9 (NM_130838.1). In this paper, we use exon numbering according to NM_130838.1 sequence which is the standard reference sequence for the *UBE3A* gene where exons 1 to 10 correspond to exons 5b to 14 of the NM_000462.3 sequence. MLPA was performed according to manufacturer's instructions (MRC Holland). Amplification products were run on an ABI3130 analyser (Applied Biosystems, California, USA) and analyzed using the GeneMapper software (Applied Biosystems).

Methylation pattern of the 15q11.2-q13 region was normal. However, dosage analysis showed a reduction of 50% in the relative peak height of the probe corresponding to *UBE3A* exon 2 while probes corresponding to *UBE3A* exons 1, 3, 4 and 9 (NM_130838.1) showed a normal diploid dosage (Fig. 1a). To confirm the specific *UBE3A* exon 2 deletion the MLPA P336-A2 kit (MRC Holland) was used. This kit contains 42 MLPA probes, including at least one probe for each exon of the *UBE3A* gene.

A reduction of 50% in the relative peak height corresponding to the probe located in the exon 2 was detected while all other *UBE3A* exonic probes presented normal dosage confirming that the deletion only involved exon 2 (Fig. 1b). Parents testing showed that deletion was de novo (Fig. 1b), although a germline mosaic cannot be excluded. In order to investigate in which parental allele the deletion was originated, we used six SNP located in intron 2: rs71418040, rs17114442, rs28754450, rs57111857, rs757438 and rs61336305. All of them were not informative so the maternal origin of the deletion could not be confirmed.

Patient 2 is a 5 year-old girl born at term following an uneventful pregnancy. Her birth weight was 3000 g. No perinatal problems were reported but she was noted to be hypotonic. She presented developmental delay.

Table 1 Clinical data of patients with *UBE3A* intragenic deletions, following the clinical features of AS described by Williams et al., 2006

	Patient 1	Patient 2	Bürger, et al., 2002	Boyes, et al., 2006	Boyes, et al., 2006	Call, et al., 2010	Beleza-Meireles, et al., 2011	Piard, et al., 2011	Piard, et al., 2011
Genetics									
Intragenic deletion	Exon 2 ^a	Exons 9-10 ^a	Whole <i>UBE3A</i>	Exons 8-16 ^b	Exons 8-16 ^b	Exon 8 ^c	Exons 5-12 ^c	Exons 6-12 ^c	Exons 6-12 ^c
Predicted protein	p.0?	p.Arg765_852del88	p.0?	p.0?	p.0?	p.Leu517Valfs*27	p.0?	p.0?	p.0?
Inheritance	De novo	De novo	Maternal	Maternal	Maternal	Maternal	NA	Maternal	Maternal
Clinical data									
Sex	F	F	M	F	M	M	F	F	F
Weight at birth	2850 g	3000 g	2670 g	3200 g	3350 g	3370 g	3880 g	NA	NA
Perinatal problems	-	-	NA	NA	NA	-	-	NA	NA
Age of sedestation (months)	9	8	12	18	21	18	NA	9	NA
Age of walk (months)	24	24	18	39	39	24	24	24	24
Consistent clinical features of AS (100% of the affected individuals)									
Ataxia of gait	+	+	+	+	+	+	+	NA	+
Frequent laughter/smiling	+	+	+	+	+	+	+	+	+
Apparent happy demeanor	+	+	NA	+	+	NA	+	+	+
Easily excitable personality	+	+	NA	NA	NA	NA	NA	NA	NA
Hyperactive behaviour	+	+	NA	+	+	NA	+	NA	+
Attention deficit	+	+	NA	NA	NA	NA	NA	NA	NA
Hand flapping/stereotypies	-	+	NA	+	NA	NA	+	+	+
Development delay	+	+	+	+	+	+	+	+	+
Severe mental retardation	+	+	+	+	+	+	+	+	+
Speech impairment	+	+	+	+	+	+	+	+	+
Receptive and non-verbal communication skills higher than verbal ones	+	+	+	NA	NA	NA	NA	NA	NA
Frequent clinical features of AS (more than 80% of the affected individuals)									
Microcephaly	Relative	-	+	+	+	+	-	+	+
Seizures	+	+	NA	+	+	-	+	+	+
Abnormal EEG	+	+	-	+	+	+	+	+	NA
Associated clinical features of AS (20-80% of the affected individuals)									
Hypotonia	-	+	-	+	+	+	+	NA	NA
Feeding problems	-	-	-	NA	NA	NA	+	NA	NA
Prognathia	-	+	NA	NA	NA	NA	+	NA	NA
Flat Occiput	-	-	NA	NA	NA	NA	+	NA	NA
Occipital groove	-	-	NA	NA	NA	NA	NA	NA	NA
Hypopigmented skin, light hair and eye color	-	-	-	+	+	NA	NA	NA	NA
Strabismus	-	-	NA	NA	NA	NA	NA	+	+
Wide mouth	-	+	-	+	NA	NA	NA	+	NA
Wide-spaced teeth	+	+	-	NA	NA	NA	NA	+	+
Protruding tongue	+	-	-	+	NA	NA	NA	NA	NA
Small hands and feet	-	-	NA	NA	NA	NA	NA	NA	NA
Scoliosis	-	-	NA	NA	+	+	NA	NA	+
Uplifted flexed arm position, especially during deambulation	-	+	NA	NA	NA	NA	NA	NA	NA

Table 1 Clinical data of patients with *UBE3A* intragenic deletions, following the clinical features of AS described by Williams et al., 2006 (Continued)

	Patient 1	Patient 2	Bürger, et al., 2002	Boyes, et al., 2006	Boyes, et al., 2006	Call, et al., 2010	Beleza-Meireles, et al., 2011	Piard, et al., 2011	Piard, et al., 2011
Hyperreflexia of the lower extremities	+	–	NA	+	NA	+	NA	NA	NA
Tremulous movement of limbs	+	–	NA	NA	NA	NA	NA	NA	NA
Frequent drooling	+	+	NA	NA	NA	+	NA	NA	NA
Suck/swallowing disorders	+	–	NA	NA	NA	NA	NA	NA	+
Abnormal sleep-wake cycle	+	–	NA	+	+	NA	NA	+	+
Chewing/mouthing behaviour	+	+	NA	NA	NA	NA	NA	NA	NA
Attraction to/fascination with water	+	–	NA	NA	NA	NA	NA	NA	NA
Increased sensitivity to heat	+	–	NA	NA	NA	NA	NA	NA	NA

F female, M male, + present, – absent, NA not available

^aNomenclature exons 1–10 according to NM_130838.1

^bNomenclature exons according to Kishino and Wagstaff, 1998; Yamamoto, et al., 1997

^cNomenclature exons 1–14 according to NM_00462.3

Unsupported sedation was achieved at 8 months and independent walking at 24 months. Oral language was limited to around 10 words.

Physical features of the patient included prognathia, wide mouth and wide-spaced teeth. The neurological manifestations included severe intellectual disability, ataxia of gait, uplifted and flexed arm position during deambulation, speech impairment, frequent drooling and seizures. The EEG showed generalized slow wave activity with paroxysmal activity. She didn't present disturbances of the sleep-wake cycle.

The behavioral traits included a happy demeanor with smiling, laughing, hyperactive behavior, attention deficit, hand flapping and chewing/mouthing behaviors. At the age of the diagnosis, 5.6-year-old, her weight was 15 Kg (10th percentile) and the head circumference was 48 cm (–1.5 SD) (Table 1).

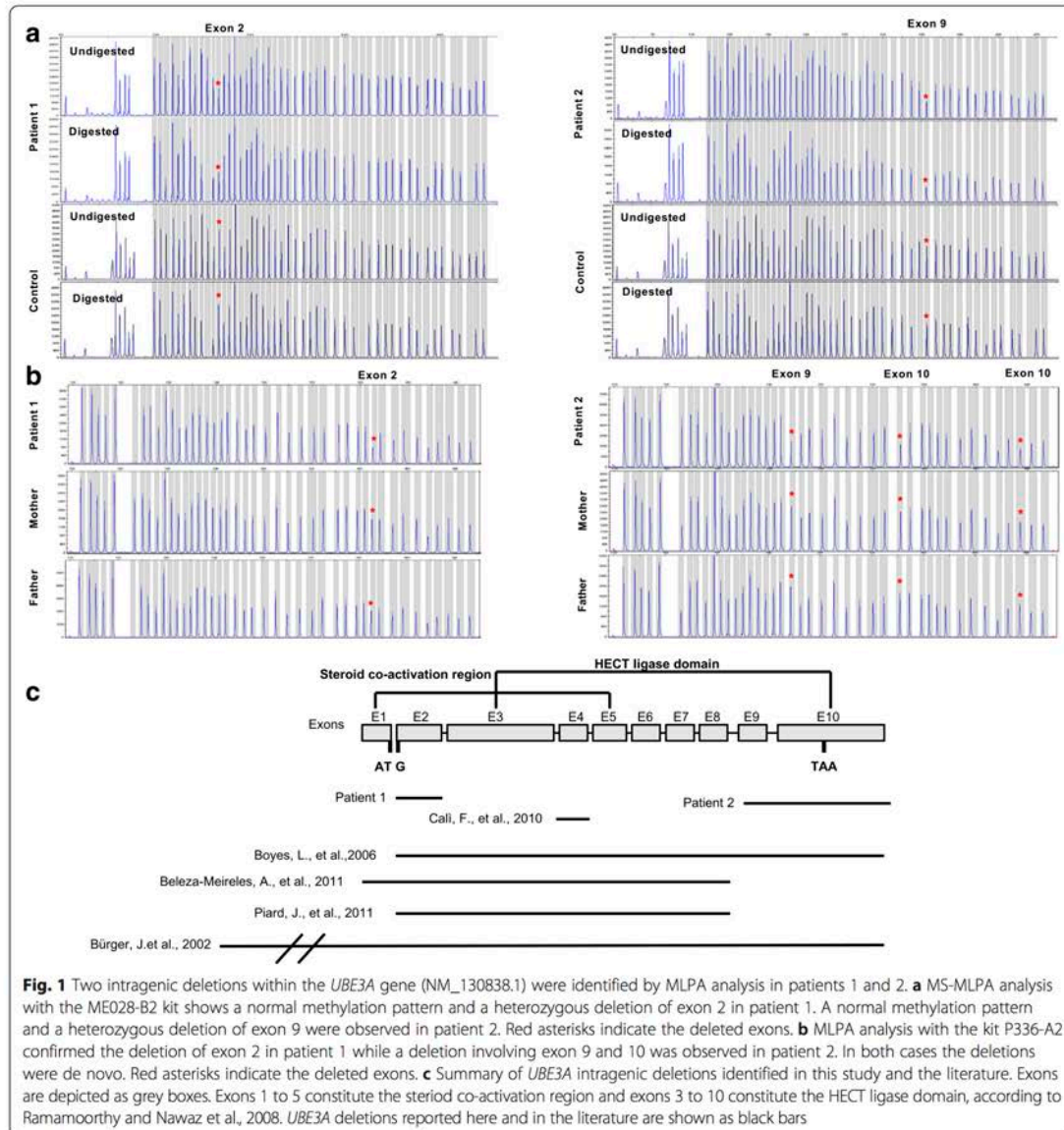
Initially, MLPA analysis was carried out using the SALSA MLPA P064 Mental Retardation probemix-1 in the Mother and Child Health Care Institute of Serbia. The results showed a deletion of the exon 10 (NM_130838.1) of the *UBE3A* gene (data not shown). Subsequently, a DNA sample of the patient was sent to our laboratory for further studies. MS-MLPA testing was performed and showed a normal methylation pattern. However, a reduction of 50% in the relative peak height of the probe corresponding to the *UBE3A* exon 9 was observed while probes corresponding to *UBE3A* exons 1, 2, 3, 4 (NM_130838.1) showed a normal diploid dosage (Fig. 1a). To confirm the *UBE3A* exon 9 and 10 deletion and to determine whether the deletion expanded to other exons, MLPA P336-A2 kit was used. MLPA analysis confirmed the deletion only involving exon 9 and 10 (Fig. 1b). The genetic analysis in the parents

showed that the deletion encompassing exons 9 and 10 was de novo (Fig. 1b). In patient 2, two SNPs in intron 9: rs573271880, rs17115485 and 1 SNP in exon 10: rs10047992 were analyzed to investigate the parental origin of the deletion. As in patient 1, we could not confirm the maternal origin of the deletion.

Discussion and conclusions

Angelman syndrome is caused by the lack of expression of the maternal copy in neurons of the *UBE3A* gene due to four different molecular etiologies: Deletion of the 15q11.2-q13 region on the maternal chromosome, mutations in the maternally-inherited copy of *UBE3A*, paternal uniparental disomy for chromosome 15 and an imprinting defect causing lack of expression of the maternal copy of *UBE3A* [16].

Genotype-phenotype correlations among molecular subclasses have shown that deletion patients show a more severe phenotype in all aspects of neurodevelopmental delay except for expressive language skills that are extremely poor regardless of their molecular subclass [17–21]. However, Mertz et al., 2014 report that children with AS due to an *UBE3A* mutation or paternal uniparental disomy present significantly better linguistic properties than deletion patients, in particular the expressive area, where 6 out of 9 children with *UBE3A* mutation or paternal uniparental disomy could use 2–7 words while it was just the case in 3 out of 30 children with a 15q11.2-q13 deletion [22]. In addition, the deletion class is the most severely affected regarding microcephaly, seizures, relative hypopigmentation, motor difficulties while paternal uniparental disomy and imprinting defect individuals have better physical growth, have less movement and ataxia abnormalities, and have a lower



prevalence (but not absence) of seizures [18]. In particular, individuals with deletions manifest higher rates of epilepsy (90%, compared to 75% among individuals without deletion), earlier onset of seizures (mean of age 1.9 years, compared with 6.3 years among individuals without deletion) and may exhibit a more severe electroencephalogram phenotype compared to the other etiologies [23].

It has been suggested that AS patients carrying *UBE3A* mutations show a phenotype somewhere in the middle

between deletion and paternal uniparental disomy carriers. They present higher incidences of seizures and microcephaly, similar to deletion patients, while their neurodevelopment delay is similar to paternal uniparental disomy and imprinting defect patient carriers [18, 19]. A revision of genotype-phenotype differences has been published recently by La Salle et al. [24].

The patients presented in this report carry intragenic *UBE3A* exonic deletions that affect the *UBE3A* open reading frame. Exon 2 deletion in patient 1 is predicted

to eliminate the start codon and consequently no protein will be produced (p.0?). Deletion of exons 9 and 10 in patient 2 is predicted to eliminate the last 88 aminoacids (p.Arg765_Leu852del88) disrupting the ubiquitin ligase catalytic domain located at the 3' end of the protein (Fig. 1c).

So far, seven patients with whole *UBE3A* gene or partial intragenic deletions have been reported (Fig. 1c). A summary of their clinical features is presented in Table 1. All deletions reported previously are also predicted to affect the *UBE3A* open reading frame and create a premature stop codon at the beginning of the protein (Table 1). Only patient 2 and the patient reported by Cali et al., 2010 present exon deletions that would alter the 3' end of the *UBE3A* disrupting the ubiquitin ligase activity of the protein.

All patients show the consistent clinical features of AS except for patient 1 who does not present stereotypes. This is surprising because stereotypes have been described as a clinical feature often present in all patients with AS since early in development. Microcephaly is present in almost all patients (7 of nine patients) while seizures are present in all except one (7 of eight patients available). The high incidence of seizures and microcephaly is in accordance to what has been reported before by Lossie et al., 2001 and Moncla et al., 1999. Associated clinical features show variability between patients. A comparison of AS patients with mutations and intragenic deletions has not been reported before and from our observations there are no differences between the two groups.

In our laboratory, array Comparative Genomic Hybridization (aCGH) is the first diagnostic tool in neurodevelopmental disorders. Methylation Specific-Multiplex Ligation-Dependent Probe Amplification (MS-MLPA) of the 15q11.2-q13 region is carried out when there is a strong clinical suspicion of Angelman or Prader Willi syndrome. MS-MLPA analysis simultaneously assesses the methylation status and genomic dosage changes at the 15q11.2-q13 region and can confirm the diagnosis and identify the presence of a causative deletion in 70% of cases. In the case of a normal methylation pattern and if the clinical suspicion remains high, sequence analysis of the *UBE3A* gene is performed, as 10% of Angelman patients harbor *UBE3A* mutations.

In addition, we propose here that the analysis of the copy number variations within the *UBE3A* gene should be taken into account, once the MS-MLPA test and the sequencing of *UBE3A* have been carried out with a negative result. Although rare, intragenic deletions may account for a small proportion of AS patients without a genetic diagnosis and if present, they have a recurrence risk of up to 50%, depending on the carrier status of the

mother. In the two patients reported here the mutation has appeared de novo but in the other intragenic deletions previously reported they were transmitted by healthy mothers to their affected children (Table 1). Moreover, MLPA analysis of the *UBE3A* gene can be used to confirm and precisely establish the size of an intragenic deletion detected by MS-MLPA. The SALSA MLPA kit P336-A2 *UBE3A* contains probes for all coding exons of the *UBE3A* gene (exons 1–10, NM_130838.1).

To date, few reports have evaluated the contribution of exonic deletions to the spectrum of *UBE3A* mutations. Here, we report two patients with AS who present two novel intragenic deletions within the *UBE3A* gene that together with those reported in the literature show that *UBE3A* intragenic deletions may represent a small fraction of AS patients. Testing for *UBE3A* intragenic exonic deletions/duplications should be included in those AS patients with a normal methylation pattern and no mutations in the *UBE3A* gene.

Abbreviations

AS: Angelman syndrome; E6-AP: E6- Associated Protein; HECT: Homologous to E6-AP COOH terminus; MLPA: Multiplex Ligation-Dependent Probe Amplification; MS-MLPA: Methylation Specific Multiplex Ligation-Dependent Probe Amplification

Acknowledgements

We are grateful to the patients and their families for their collaboration.

Funding

The financial support for carrying out this work was received from Fundació Parc Taulí- Institut d'Investigació i Innovació Parc Taulí I3PT (CIR2015/040), Asociación Española de Síndrome de Angelman and Instituto de Salud Carlos III (PI16/01411).

Availability of data and materials

Additional data that support the findings of this study are available from the corresponding author.

Authors' contributions

Design of the work: MG, CA, AR, NB, MV. Data collection: CA, MV, NB, EG, CF, NC, SC, AS, MM, DR. Data analysis and interpretation: CA, MV, NB, AR, MG. Manuscript preparation: CA, AR. Manuscript revising, all authors. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Informed consent for this investigation was obtained from all participating patients and parents, and the project was carried out in agreement with the Ethical Committee of the Corporació Sanitària Parc Taulí.

Consent for publication

Consent to publish from the parents to report patient data has been obtained.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Genetics Laboratory, UDIAT-Centre Diagnòstic, Parc Taulí Hospital Universitari, Institut d'Investigació i Innovació Parc Taulí I3PT, Universitat Autònoma de Barcelona, Parc del Taulí 1, 08208 Barcelona, Sabadell, Spain. ²Paediatric Unit, Parc Taulí Hospital Universitari, Institut d'Investigació i Innovació Parc Taulí I3PT, Universitat Autònoma de Barcelona, Sabadell, Spain. ³Laboratory for Medical Genetics, Mother and Child Health Care Institute of Serbia "Dr Vukan Cupic", Belgrade, Serbia. ⁴Department of Metabolism and Clinical Genetics, Mother and Child Health Care Institute of Serbia "Dr Vukan Cupic", Belgrade, Serbia. ⁵School of Medicine, University of Belgrade, Belgrade, Serbia.

Received: 28 February 2017 Accepted: 14 November 2017

Published online: 21 November 2017

References

- Williams CA, Zori RT, Hendrickson J, Stalker H, Marum T, Whidden E, Driscoll DJ. Angelman Syndrome. *Curr Probl Pediatr.* 1995;25(7):216–31.
- Williams CA, Angelman H, Clayton-Smith J, Driscoll DJ, Hendrickson JE, Knoll JHM, Magenis RE, Schinzel A, Wagstaff J, Midden EM, et al. Angelman Syndrome. Consensus for diagnostic criteria. *Am J Med Genet.* 1995;56(2):237–8.
- Williams CA, Beaudet AL, Clayton-Smith J, Knoll JH, Kyllerman M, Laan LA, Magenis RE, Moncla A, Schinzel AA, Summers JA, et al. Angelman syndrome 2005: updated consensus for diagnostic criteria. *Am J Med Genet A.* 2006; 140A(5):413–8.
- Buiting K. Prader-Willi syndrome and Angelman syndrome. *Am J Med Genet C: Semin Med Genet.* 2010;154C(3):365–76.
- Ramsden SC, Clayton-Smith J, Birch R, Buiting K. Practice guidelines for the molecular analysis of Prader-Willi and Angelman syndromes. *BMC Medical Genetics.* 2010;11(1)
- Beleza-Meireles A, Cerqueira R, Sousa SB, Palmeiro A, Ramos L. Novel deletion encompassing exons 5–12 of the UBE3A gene in a girl with Angelman syndrome. *European Journal of Medical Genetics.* 2011;54(3):348–50.
- Boyes L, Wallace AJ, Krajewska-Walasek M, Chrzanowska KH, Clayton-Smith J, Ramsden S. Detection of a deletion of exons 8–16 of the UBE3A gene in familial Angelman syndrome using a semi-quantitative dosage PCR based assay. *European Journal of Medical Genetics.* 2006;49(6):472–80.
- Calli F, Ragalmuto A, Chiavetta V, Calabrese G, Fichera M, Vinci M, Ruggeri G, Schinocca P, Sturnio M, Romano S, et al. Novel deletion of the E3A ubiquitin protein ligase gene detected by multiplex ligation-dependent probe amplification in a patient with Angelman syndrome. *Experimental and Mol Med.* 2010;42(12):842.
- Piard J, Depienne C, Keren B, Fédirkó E, Trouillard O, Charles P, Heron D. Intragenic deletion of UBE3A gene in 2 sisters with Angelman syndrome detected by MLPA. *Am J Med Genet A.* 2011;155(12):3170–3.
- Bürger J, Horn D, Tönnies H, Neitzel H, Reis A. Familial interstitial 570 kbp deletion of the UBE3A gene region causing Angelman syndrome but not Prader-Willi syndrome. *Am J Med Genet.* 2002;111(3):233–7.
- Dagli A, Buiting K, Williams CA. Molecular and clinical aspects of Angelman syndrome. *Molecular Syndromology.* 2011;
- Ramamoorthy S, Nawaz Z. E6-associated protein (E6-AP) is a dual function coactivator of steroid hormone receptors. *Nucl Recept Signal.* 2008;6:e006.
- Bird L. Angelman syndrome: review of clinical and molecular aspects. *Appl Clin Genet.* 2014;7:93–104.
- Sadikovic B, Fernandes P, Zhang VW, Ward PA, Miloslavskaya I, Rhead W, Rosenbaum R, Gin R, Roa B, Fang P. Mutation update for UBE3A variants in Angelman syndrome. *Hum Mutat.* 2014;35(12):1407–17.
- Tzagkaraki E, Sofocleous C, Helen F-K, Dinopoulos A, Goulielmos G, Mavrou A, Sofia K-T, Kanavakis E. Screening of UBE3A gene in patients referred for Angelman syndrome. *Eur J Paediatr Neurol.* 2013;17(4):366–73.
- Williams CA, Driscoll DJ, Dagli AI. Clinical and genetic aspects of Angelman syndrome. *Genetics in Medicine.* 2010;12(7):385–95.
- Gentile JK, Tan W-H, Horowitz LT, Bacino CA, Skinner SA, Barbieri-Welge R, Bauer-Carlin A, Beaudet AL, Bichell TJ, Lee H-S, et al. A neurodevelopmental survey of Angelman syndrome with genotype-phenotype correlations. *Journal of Developmental & Behavioral Pediatrics.* 2010;31(7):592–601.
- Lossie A, Whitney M, Ammidon D, Dong H, Chen P. Distinct phenotypes distinguish the molecular classes of AS. *J Med Genet.* 2001;38B34–45.
- Moncla A, Malzac P, Voelckel M-A, Auquier P, Girardot L, Mattei M-G, Philip N, Mattei J-F, Lalonde M, Livet M-O. Phenotype-genotype correlation in 20 deletion and 20 non-deletion Angelman syndrome patients. *Eur J Hum Genet.* 1999;7:131–9.
- Varela MC, Kok F, Otto PA, Koiffmann CP. Phenotypic variability in Angelman syndrome: comparison among different deletion classes and between deletion and UPD subjects. *Eur J Hum Genet.* 2004;12(12):987–92.
- Sahoo T, Bacino CA, German JR, Shaw CA, Bird LM, Kimonis V, Anselm I, Waisbren S, Beaudet AL, Peters SU. Identification of novel deletions of 15q11q13 in Angelman syndrome by array-CGH: molecular characterization and genotype-phenotype correlations. *Eur J Hum Genet.* 2007;15(9):943–9.
- Mertz LGB, Thaulov P, Trillingsgaard A, Christensen R, Vogel I, Hertz JM, Østergaard JR. Neurodevelopmental outcome in Angelman syndrome: genotype-phenotype correlations. *Res Dev Disabil.* 2014;35(7):1742–7.
- Thibert RL, Larson AM, Hsieh DT, Raby AR, Thiele EA. Neurologic manifestations of Angelman syndrome. *Pediatr Neurol.* 2013;48(4):271–9.
- LaSalle JM, Reiter LT, Chamberlain SJ. Epigenetic regulation of UBE3A and roles in human neurodevelopmental disorders. *Epigenomics.* 2015;7(7): 1213–28.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit



