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

Cinthia Aguilera[1], Elisabeth Gabau[2], Ariadna Ramirez-Mallafré[2], Carme Brun-Gasca [2,3], Jana Dominguez-Carral[2], Veronica Delgadillo[2], Steve Laurie[4], Sophia Derdak[4], Natàlia Padilla[5], Xavier de la Cruz[5,6], Núria Capdevila[2], Nino Spataro[1], Neus Baena[1] , Miriam Guitart[1]*and Anna Ruiz[1]*

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

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

[3]Department of Clinical Psychology and Health Psychology, Universitat Autònoma de Barcelona, 08193 Bellatera, Barcelona,Spain.

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

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

[6]Institució 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

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 14 patients with clinical features of AS and no molecular diagnosis. As a result, we identified 10 *de novo* and 1 X-linked pathogenic/likely pathogenic variants in 10 neurodevelopmental genes (*SYNGAP1, VAMP2, TBL1XR1, ASXL3, SATB2, SMARCE1, SPTAN1, 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.

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 13 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 78,5% 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

14 patients (7 girls and 7 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 1 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 9 of 14 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.

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
				Severe developmental delay	Speech impairment	Ataxia or unsteady gait	Apparent happy demeanor/ Stereotypies	Microcephaly	Seizures	Abnormal EEG	
1	M	1	14	+	+ (5-10 words)	+	+/+	-	+	+	Sleep disorder, hypotonia
2	F	8	19	+	+ (less than 5 words)	-	+/+	-	+	+	Sleep disorder, feeding problems, kyphoscoliosis
3	F	1	12	+	+ (Absent speech)	+	-/+	-	+	+	Hypotonia
4	F	11 months	9	+	+ (Absent speech)	+	+/+	+ (Relative)	-	+	Hypotonia, feeding problems
5	F	19 months	20	+	+ (5-10 words)	+	+/-	-	-	+	-
6	F	4	18	+	+ (Absent speech)	-	+/-	+ (Relative)	+	NA	Scoliosis
7	M	5	15	+	+ (less than 5 words)	+	+/+	-	+	+	Feeding problems, wide mouth, hypotonia
8	F	1	14	+	+ (More than 20 words)	+	+/-	+	-	-	Hypotonia
9	M	8	38	+	+ (Absent speech)	+	+/+	+ (Relative)	+	+	Hypotonia, feeding problems (oesophageal reflux), sleep disorder
10	M	1	7	+	+ (less than 5 words)	-	+/-	+ (Relative)	-	-	Wide spaced teeth, brachycephaly
11	F	2	14	+	+ (less than 5 words)	-	+/+	+ (Relative)	+	NA	Feeding problems (dysphagia)
12	M	3	24	+	+ (5-10 words)	-	+/+	+	+	NA	Strabismus, sleep disorder, kyphoscoliosis
13	M	3	9	+	+ (Absent speech)	+	+/+	+ (Relative)	+	-	Sleep disorder, hypotonia
14	M	2	13	+	+ (Absent speech)	+	+/+	-	+	+	Sleep disorder, wide-spaced teeth

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

Whole-exome sequencing and variant interpretation

Trio WES of 13 patients and their parents was performed using the SureSelect Human All Exon V5+UTR kit (Agilent technologies). In patient 4, 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 (S1 Table). 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 PrimeScript^{TM}RT 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) (S1 Fig).

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, 10 *de novo* (*SYNGAP1, VAMP2, TBL1XR1, ASXL3, SATB2, SMARCE1, SPTAN1, KCNQ3* and *SLC6A1)* and 1 X-linked (*LAS1L*) protein altering variants were confirmed in 11 patients, leading to a diagnostic yield of 78,5%. These variants were located in 10 different genes previously reported to be associated with neurodevelopmental disorders [5,10–17].

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* and *KCNQ3*) and chromatin remodeling or transcription regulation (*TBL1XR1, SATB2, SMARCE1, ASXL3* and *LAS1L*) as has been described before in neurodevelopmental diseases [18,19].

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 the patients. In short, cerebellar atrophy in *SPTAN1* [10], hypoplasia of the corpus callosum, hypoplasia of the 5th finger nail, hypertrichosis, sparse scalp hair and aggressive behavior in *SMARCE1* [20], truncal obesity and short stature in *LAS1L* [13], myoclonic atonic seizures in *SLC6A1* [12], aggressive behavior in *SYNGAP1* [14], dysmorphic features and dental anomalies in *ASXL*3 [11] and aggressive behavior and dental anomalies in *SATB2* [16] (Fig 1).

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

Patient	Gene	NM number	Nucleotide change	Amino acid change	Variant Type	Pattern of inheritance	Varsome Classification	Described before	Protein function
1	*VAMP2*	NM_014232.2	c.128_130delTGG	p.Val43del	In-frame	*De novo*	Pathogenic	Yes Salprieto et al., 2019	*VAMP2* is a member of the SNARE family of proteins, which are involved in membrane fusion of synaptic vesicles.
2	*SYNGAP1*	NM_006772.2	c.1861C>T	p.Arg621*	Nonsense	*De novo*	Pathogenic	No	*SYNGAP1* is a RAS-GTPase-activating protein with a critical role in synaptic development, structure, function and plasticity.
3	*TBL1XR1*	NM_024665.5	c.1000T>C	p.Cys334Arg	Missense	*De novo*	Likely pathogenic	No	*TBL1XR1* is part of the repressive NCoR/SMRT complex acting as a transcriptional regulator.
4	*TBL1XR1*	NM_024665.5	c.1043A>G	p.His348Arg	Missense	*De novo*	Likely pathogenic	No	
5	*SATB2*	NM_001172509.1	c.1826delA	p.Asp609Alafs*15	Frameshift	*De novo*	Pathogenic	No	*SATB2* participates in chromatin remodeling and transcription regulation.
6	*KCNQ3*	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	*KCNQ3* is a voltage-gated potassium channel subunits that underlay the neuronal M-Current.
7	*SMARCE1*	NG_032163.1 (NM_003079.4)	c.237+1G>T	p.Ala53_Lys79del	Splice site	*De novo*	Likely pathogenic	No	*SMARCE1* is part of the SWI/SNF chromatin remodeling complex involved in transcriptional activation.
8	*SPTAN1*	NM_001130438.2	c.6592_6597dupCTGCAG	p.Leu2198_Gln2199dup	In-frame	*De novo*	Likely pathogenic	No	*SPTAN1* is an α-II spectrin involved in stabilization and activation of membrane channels, transporters and receptors.
9	*ASXL3*	NM_030632.2	c.3106C>T	p.Arg1036*	Nonsense	*De novo*	Pathogenic	Yes Kuechler A et al., 2016	*ASXL3* plays a role in the regulation of gene transcription and histone deubiquitination.
10	*LAS1L*	NM_031206.4	c.1237G>A	p.Gly413Arg	Missense	X-linked	Likely pathogenic	No	*LAS1L* is involved in the 60S ribosomal subunit synthesis, maturation of 28S rRNA and regulation of transcription.
14	*SLC6A1*	NM_003042.3	c.889G>A	p.Gly297Arg	Missense	*De novo*	Pathogenic	Yes Carvill GL et al., 2015	*SLC6A1* gene encodes for the GAT-1 GABA transporter.

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.

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* [14]; similarly the patient harboring a pathogenic variant in *SATB2* did not show sialorrhea and feeding difficulties [16]. Finally, the ataxia of gait, stereotypies and hypotonia were not observed for the patient with a pathogenic variant in *KCNQ3* [17].

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 [21], supporting a loss of function mechanism of disease for the *HSF2* gene (S2 Fig).

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*.

183 However, a significant enrichment in GO terms related to the nervous system
184 development (biological process), protein N-terminus binding (molecular function) and
185 nuclear lumen (cellular component) were observed. Overall, these results suggest that
186 the genes found in our cohort, the *UBE3A* gene and the known AS-like genes may not
187 interact directly but carry out biological functions involved in synapsis and brain
188 development, leading to overlapping phenotypes (S1 Fig).
189

190 # DISCUSSION

191 We identified causal variants in 11 out of 14 patients with an AS-like phenotype.
192 The global yield diagnostic of WES in this is study is 78,5%, which is higher to what has
193 been reported in the literature for other neurodevelopmental disorders (24-68%) [22].
194 The results of WES led to the identification of 10 new genes to cause an AS-like
195 phenotype (*SYNGAP1, VAMP2, TBL1XR1, ASXL3, SATB2, SMARCE1, SPTAN1, KCNQ3,*
196 *SLC6A1* and *LAS1L*), all of them previously associated with other neurodevelopmental
197 disorders. In addition, we propose *HSF2* (Heat Shock Factor) as a new candidate gene
198 for the AS-like phenotype. Although *HSF2* has not been previously associated with any
199 human disease, the gene is highly expressed in the brain and highly intolerant to loss of
200 function variation (pLI 0.92). *HSF2* knockout mice show defects in spermatogenesis and
201 in the development of the central nervous system [23,24]. The identification of
202 additional patients with loss of function variants in *HSF2* and functional studies in neural
203 cells will contribute to elucidate the role of *HSF2* in the AS-like phenotype.
204 *De novo* variants have been described to account for approximately half of the
205 genetic architecture of severe developmental disorders [25]. In our cohort, 10 of the 11

206 pathogenic and likely pathogenic variants were *de novo*, accounting for 90% of diagnosis
207 and highlighting the power of using trio-WES for the molecular diagnosis of severe
208 developmental disorders. Only in one case, the X-linked variant in *LAS1L* was inherited
209 from the mother, who was a healthy mosaic carrier (20%).

210 All patients had received an initial diagnosis of AS, supported by the presence of
211 consistent and frequent clinical features. In the majority of our patients (11/14) the
212 initial diagnosis was done during infancy or early childhood (before five years old). At
213 the time of initial diagnosis, all of them presented severe global DD and speech
214 impairment in addition to ataxia of gait or the characteristic happy disposition. The
215 review of patient's clinical reports upon WES results showed the presence of additional
216 clinical features generally not described for AS, but that were then associated with the
217 new identified genes. Pathogenic/likely pathogenic variants in *SMARCE1, SATB2,*
218 *SYNGAP1, SLC6A1, ASXL3*, *SPTAN1* and *LAS1L* genes are associated with
219 neurodevelopment disorders that overlap with AS and with some differential features
220 that were present in our patients (Fig 1). On the other hand, *VAMP2, KCNQ3* and
221 *TBL1XR1* genes are associated with ID, autism spectrum disorder and epilepsy [5,15,17],
222 features that are shared by AS-like patients.

223 Lack of molecular diagnosis in 10-15% of clinically diagnosed AS patients has
224 been used to define the AS-like group. Our results indicate that the 78,5% of AS-like
225 patients are carriers of pathogenic variants in genes involved in neurodevelopmental
226 disorders whose features overlap with AS, showing the wide genetic heterogeneity in
227 AS-like (Fig 1). AS-like new genes do not interact directly with *UBE3A* gene product but
228 are involved in synapsis and nervous system development. Except for the *SYNGAP1*
229 gene, none of the genes identified here have been previously described in the

differential diagnosis of AS [26]. 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.

Acknowledgements

We thank the patients and families for their participation in this study.

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: 413–418. doi:10.1002/ajmg.a

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;27: 1326–1340. doi:10.1038/s41431-019-0435-0

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: 975–992. doi:10.1002/ajmg.a.36416

4. Luk H-M. Angelman-Like Syndrome: A Genetic Approach to Diagnosis with Illustrative Cases. Case Rep Genet. 2016;2016: 1–6. doi:10.1155/2016/9790169

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;104: 721–730. doi:10.1016/j.ajhg.2019.02.016

6. Bell L, Wittkowski A, Hare DJ. Movement Disorders and Syndromic Autism: A Systematic Review. J Autism Dev Disord. 2019;49: 54–67. doi:10.1007/s10803-018-3658-y

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. doi:10.1002/0471250953.bi1110s43

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: 405–423. doi:10.1038/gim.2015.30

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–D368. doi:10.1093/nar/gkw937

10. 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;140: 2322–2336. doi:10.1093/brain/awx195

11. 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: 537–543. doi:10.1136/jmedgenet-2016-104360

12. 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. doi:10.1111/epi.13986

13. 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: 133–148. doi:10.1038/mp.2014.193

14. 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;53: 511–522. doi:10.1136/jmedgenet-2015-103451

15. 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: 927–935. doi:10.1093/hmg/ddv625

16. 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: 925–935. doi:10.1002/ajmg.a.38630

17. 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. doi:10.1002/ana.25522

18. Hormozdiari F, Penn O, Borenstein E, Eichler EE. The discovery of integrated gene networks for autism and related disorders. Genome Res. 2015;25: 142–154. doi:10.1101/gr.178855.114

19. 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;98: 149–164. doi:10.1016/j.ajhg.2015.11.024

20. 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. doi:10.1093/bioinformatics/bty897

21. Lykke-Andersen S, Jensen TH. Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat Rev Mol Cell Biol. 2015;16: 665–677. doi:10.1038/nrm4063

22. 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. doi:10.1038/s41525-018-0053-8

23. 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. doi:10.1002/gene.10200

24. 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;21: 2591–2601. doi:10.1093/emboj/21.11.2591

25. 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;542: 433–438. doi:10.1038/nature21062

26. 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: 2231–2237. doi:10.1002/ajmg.a.37189

Supporting information

S1 Fig. 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.

S2 Fig. 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 12 and a control sample shows a reduction in the percentage of the allele with the variant in the cDNA compared to DNA.

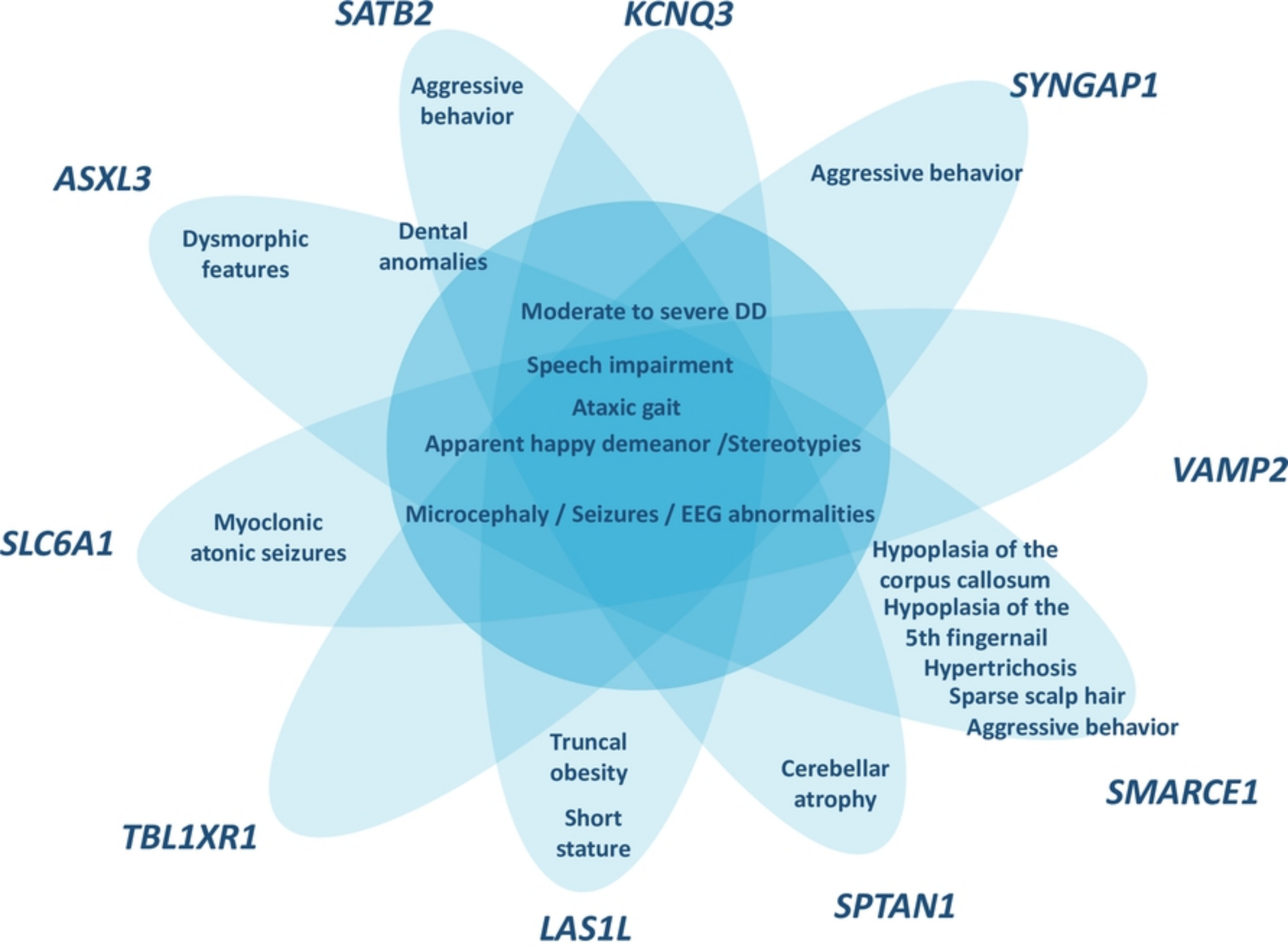
SATB2
KCNQ3
SYNGAP1
Aggressive behavior
Aggressive behavior
ASXL3
Dysmorphic features
Dental anomalies
Moderate to severe DD
Speech impairment
Ataxic gait
Apparent happy demeanor /Stereotypies
Microcephaly / Seizures / EEG abnormalities
VAMP2
SLC6A1
Myoclonic atonic seizures
Hypoplasia of the corpus callosum
Hypoplasia of the 5th fingernail
Hypertrichosis
Sparse scalp hair
Aggressive behavior
SMARCE1
Truncal obesity
Cerebellar atrophy
Short stature
TBL1XR1
SPTAN1
LAS1L

Fig 1