Manuscript Title: Newly Hemizygous Mutations In *L1CAM* In Two Unrelated Probands With Childhood Onset Psychosis

Running Title: Damaging *L1CAM* variants in Childhood Psychosis

Mitra S. Sato[1], Marinos Kyriakopoulos[2], Anthony James[3], Susanne Marwedel[4], Clare Borsay[5], Armandina Almanza Gutierrez[1], Alexandra I. Blakemore[6,7] and Anna C. Need[1,8]

[1]Faculty of Medicine, Department of Brain Sciences, Imperial College London, Du Cane Road, London, W12 0NN, UK.
[2]National and Specialist Acorn Lodge Inpatient Children's Unit, South London and Maudsley NHS Foundation Trust, Bethlem Royal Hospital, Monks Orchard Road, Beckenham, Kent, BR3 3BX, UK
[3]Highfield Adolescent Inpatient Unit, Warneford Hospital, Warneford Lane, Headington, Oxford, OX3 7JX, UK
[4]Chorley and South Ribble CAMHS, Shawbrook House, Balcarres Road, Leyland, PR25 3ED, UK
[5]Oxford Health NHS Foundation Trust, Central Oxon CAMHS, Raglan House, Cowley, Oxford, OX4 3LX, UK
[6]Department of Life Sciences, Brunel University London, Uxbridge, UB8 3PH, UK
[7]Faculty of Medicine, Department of Metabolism, Digestion and Reproduction, Imperial College London, Du Cane Road, London, W12 0NN, UK
[8]William Harvey Research Institute, Queen Mary University of London, London, EC1M 6BQ, UK

Corresponding Author:
Anna C. Need, PhD
William Harvey Research Institute, Queen Mary University of London, London, EC1M 6BQ, UK
Email address: a.need@qmul.ac.uk

Funding
This work was supported by the Marie Curie Integration Grant (grant number CIG 631787).

Abstract:

Objective: To identify genes underlying childhood onset psychosis. Methods: Patients with onset of psychosis at age 13 or younger were identified from clinics across England and they and their parents were exome sequenced and analysed for likely highly-penetrant genetic contributors. Results: We report two male childhood onset psychosis patients of different ancestries carrying hemizygous very rare damaging missense variants (p.Arg846His and p.Pro145Ser) in the *L1CAM* gene. *L1CAM* is an X-linked Mendelian disease gene in which both missense and loss of function variants are associated with syndromic forms of intellectual disability and developmental disorder. Conclusions: This is the first report *L1CAM* variants contributing to a psychiatric disorder. The family history and presence of other significant rare genetic variants in the patients suggest there may be genetic interactions affecting the presentation.

Keywords: exome, COS, childhood-onset, schizophrenia

Introduction:

Schizophrenia is a heterogeneous neurodevelopmental disorder affecting ~1% of the world's population. Childhood-onset schizophrenia (COS) is a very rare and severe sub-type of the disorder with an age of onset of 13 or under. In the UK, childhood onset psychosis (COP), including COS and other childhood-onset schizophrenia spectrum disorders, has an incidence of 1 in 500,000[1].

Previous work on COS has suggested it has a greater Mendelian genetic contribution than adult-onset schizophrenia (AOS), including a higher rate of pathogenic copy number variants[2] (CNVs; 11.9% in COS, vs. 2.5% in AOS), and a higher rate of *de novo* mutations in sporadic cases[3] (1.93×10^{-8} per bp in COS vs. 1.61×10^{-8} in AOS). This is in keeping with other complex genetic disorders in which stratification by an earlier age of onset has revealed a more genetic form of the disease [4–6]. A number of COS cases with damaging genetic variants in Mendelian disease genes have been reported, including probands with variants in *MECP2*[7], *TRRAP*[8] and *UPF3B*[9], as well as CNVs at known schizophrenia-associated regions 15q13.3[10] and 16p13.11[11] . More recently, several unrelated COS cases have been reported to carry *de novo* damaging missense variants in *ATP1A3*, severe pathogenic mutations in which cause alternating hemiplegia of childhood[12,13] .

Here we describe 2 COP patients with damaging variants in another Mendelian disease gene, *L1CAM*. *L1CAM* encodes the L1 adhesion molecule protein[14] expressed on neuronal surfaces and is essential for neuronal migration and organisation, axonal growth and formation of synapses [15–17]. *L1CAM* is intolerant of both loss of function (pLI 1)[18] and missense variation (RVIS[19] bottom 2.5%, constraint score 2.84[18]). Severe pathogenic mutations in *L1CAM* cause L1 syndrome which is a Mendelian disease associated with agenesis of the corpus callosum,

intellectual disability, adducted thumbs, spasticity and hydrocephalus[20–23]. The severity of L1 syndrome is associated with the type of mutation in *L1CAM*, whereby loss of function variants are associated with more severe phenotypes and missense variants are associated with milder phenotypes[24]. From a cohort of 29 cases with COP, we have identified two unrelated probands, of different ethnicities, with rare damaging newly hemizygous missense variants in *L1CAM.* This may represent phenotypic expansion for *L1CAM* and a new genetic association with COP.

Methods:

The probands were recruited through participating NHS trusts in England as part of the Genetics of Childhood-Onset-Psychosis study. Informed consent was obtained from both probands (if aged 16 or over) and their parents or available first-degree relatives. If the proband was <16 years in age, parents provided informed consent on behalf of a child proband and assent was obtained from the probands. Individuals considered by their psychiatrist to have clear non-genetic causes of illness (e.g. history of brain injury, substance abuse, very severe trauma) were excluded from the study. Saliva samples were collected for genetic analyses.

Saliva DNA was extracted according to the GeneFix Saliva isolation protocol[25]. Exome sequencing was performed at the Institute of Genomic Medicine (IGM), Columbia University according to standard protocols, as described here (https://www.nature.com/articles/gim2014191#materials-and-methods). Briefly, raw sequencing reads were processed using FastQC v0.11.2 and aligned to the Human Reference Genome Build 37 (Hg19), duplicated reads were removed using Picard tools v1.85 and calibrated using GATK v3.3-0[26,27].

Variants were put through a list of quality control metrics and all had to meet the following GATK hard-filtering thresholds[27]; QD≥2, MQ≥40, GQ≥20 and QUAL≥30.

The analysis focuses only on protein-changing (functional) variants. This list included truncating variants (stop gain/loss, start loss/stop lost or frameshift), missense variants and essential splice-site variants (+/- 2bp from exon/intron boundary). All functional variants were required to pass the following additional filters; (1) read depth≥10 in all three

members of the family (2) present in <2 individuals in the in-house controls (n=4,023) in heterozygous/homozygous form for a heterozygous/homozygous variant, respectively (3) RVIS<50 (4) minor allele frequency $<3.09 \times 10^{-3}$ in gnomAD v2.1.1 for a homozygous variant (which is equivalent of observing a maximum of 2 carriers in homozygous/hemizygous form) and maximum allele count of 2 for a heterozygous variant (5) CADD>15 (6) variant not flagged within the Washington Exome Variant Server (EVS) as a SVM or INDEL putative artefact (http://evs.gs.washington.edu/EVS/HelpDescriptions.jsp?tab=tabs-1#FilterStatus) and (8) pass a visual inspection using Integrative Genomics Viewer (IGV)[28].

We shortlisted variants fitting 5 different genetic models: a) known (likely) pathogenic variant analysis based on the ClinVar database b) *de novo* variants absent in available IGM control population (n=4,023) and gnomAD (n=141,456), c) recessive homozygous genotypes with a MAF $<3.09 \times 10^{-3}$ in gnomAD, d) hemizygous variants inherited from a heterozygous mother with MAF $<3.09 \times 10^{-3}$ and e) compound heterozygous variants where neither contributing variant is homozygous in >2 population controls in gnomAD. The genotypes meeting all of these criteria were referred to as "qualifying genotypes" with the genes harbouring the qualifying genotypes referred to as "qualifying genes".

For validation of the variants identified through exome sequencing, DNA from both probands and their first degree relatives was amplified by polymerase chain reaction (PCR) on the G-STORM GS4 thermal cycler system. The primers were designed using the Primer3Plus program[29] and optimal annealing temperature was determined by performing gradient PCR. Raw Sanger sequencing traces (ab1.files) were processed using the CLC Genomics Workbench (https://www.qiagenbioinformatics.com/products/clc-genomics-workbench/

). Briefly, both forward and reverse reads were trimmed using the generic parameters and mapped to the region of interest in the *L1CAM* gene. The reads were visually assessed for confirmation of presence/absence of variant of interest, as shown in Figures 1B and 2B.

Pedigrees for both families were drawn using a customised version of the Kinship-2[30] package in R.

Results:

Clinical presentation and family history

Proband 1

Proband 1 is a 13-year-old boy of Ghanaian ancestry with a diagnosis of COP from the age of 9. He was born with no complications at full term via normal vaginal delivery to non-consanguineous parents. He fed well and slept normally in his early months and met his early developmental milestones, sitting and standing at 6 months and walking unassisted by 9 months. Parents recall that the proband's first words used meaningfully were delayed, which led to difficulties in articulating himself early on. There were no concerns regarding hyperactivity, inattention or impulsivity until the age of 9, when the parents report that the proband became hyperactive and started displaying unsafe behaviours such as engaging in physical aggression. At the same time he presented with thought disorder, frequent hallucinations, and emotional and behavioural disturbances. He was initially reviewed by paediatrics and known organic causes of his presentation were ruled out. Investigations included extensive blood investigations, EEG and brain MRI. He then had his first admission to a mental health inpatient children's unit which lasted for 7 months. His physical examination at the time of hospitalisation revealed no abnormalities with his cardiovascular, abdominal and respiratory system. His neurological examination was also normal. His ECG showed right ventricular dominant forces by voltage criteria but was otherwise normal. Following this episode, he was diagnosed with COP on a background of some early language delay and received aripiprazole gradually increased to 3mg od to which he responded well. A higher dose of aripiprazole was associated with increased agitation and aggression. His cognitive assessment with the Wechsler Abbreviated Scale of Intelligence - Second Edition

(WASI-II) at age 9 years 6 months revealed an IQ score of 72, which was classed as borderline intellectual functioning. At that time, he was not able to engage in assessment with the Wechsler Intelligence Scale for Children – Fourth Edition (WISC IV). He subsequently underwent assessment for autism including the Autism Diagnostic Observation Schedule (ADOS) and was diagnosed with autism spectrum disorder (ASD).

The proband was re-hospitalised after 9 months due to deterioration associated with motor tics which were continuously present but fluctuated in severity throughout his admission and eventually received the diagnosis of chronic motor tic disorder. Extensive blood investigations were repeated and did not reveal a cause for his psychosis. His clinical presentation seemed to be treatment resistant and this second hospitalisation lasted for 22 months. Antipsychotics tried included an increased dose of aripiprazole, lurasidone, and olanzapine. A trial of a small dose of sertraline for anxiety did not seem to be helpful, was associated with increased excitability, and was quickly discontinued. Episodes of elevated mood, possibly having a seasonal exacerbation in the autumn, were observed and the diagnosis of childhood-onset schizoaffective disorder was confirmed. Sodium valproate was introduced and gradually titrated, and metformin was initiated for weight management. Assessment with the WISC IV at 10 years 7 months revealed a full-scale IQ (FSIQ) score of 51, and with WISC V at 11 years 7 months an FSIQ score of 52, corresponding to moderate intellectual disability. The proband was discharged on olanzapine 12.5mg daily, sodium valproate 400mg bd, metformin 250mg bd, and movicol for constipation.

There is no positive family history for neuropsychiatric illness (Figure 1A). The proband lives with his mother, father, an older brother and an older sister all of whom are reportedly well. His eldest sister attends school in Ghana. The proband is a second-generation immigrant

from Ghana and lives in a family with reportedly significant financial strain. The proband has a male first cousin (Figure 1A) on the maternal side who is reported to have had unspecified speech problems since birth. There are no other reports of neuro-psychiatric/developmental illness in the extended family.

Proband 2

Proband 2 is a 14-year-old boy of white British ancestry with a diagnosis of COP from the age of 10. He was born with no complications to non-consanguineous parents, two weeks before a planned Caesarean section date. His birthweight was 3700g. He fed well, slept normally and met all his early developmental milestones, although his speech development was delayed. As a toddler, he was keen on having routine and order. He struggled with social interactions, had significant sensory issues and behavioural rigidity. He reacted with behavioural outbursts and self-directed aggression if his rigidity and routines were challenged. From the age of 2 or 3 he had an imaginary friend he spoke to and was reportedly seen muttering and talking to himself throughout nursery years.

At age 10, he presented with significant auditory hallucinations, at times associated with headaches, as well as violent outbursts, excessive daytime sleepiness with poor night time sleep, night terrors and joint pains, which required him to use a wheel chair for longer distances. He also had a history of nose bleeds. Proband 2 had investigations for narcolepsy, epilepsy and organic causes for psychosis, including an MRI of the brain, lumbar puncture, sleep studies and video EEG, as well a number of metabolic investigations, which were are all reported as normal. Neurological examination was also normal. Following the conclusion of the investigations, Proband 2 received a diagnosis of non-organic psychosis and was started on aripiprazole. In view of his daytime sleepiness, he was seen at a Chronic Fatigue Syndrome (CFS) clinic. Whilst acknowledging that he has symptoms of fatigue, he was not

given a diagnosis of CFS. Proband 2 was given a diagnosis of Autism Spectrum Disorder (ASD) at the age of 11y10m. Aged 12 and 13, the proband attempted suicide in response to auditory command hallucinations.

Proband 2 is maintained on aripiprazole 5mg with residual, but manageable psychotic symptoms and no significant side effects. After a period of home education as a result of his fatigue, he was gradually and successfully re-integrated into a mainstream school.

There is a positive and complex family history of neuropsychiatric illness (Figure 2A). The proband's mother has a diagnosis of obsessive-compulsive disorder (OCD), social anxiety and has been on sertraline since the age of 21 (currently 33). She suffers from cluster headaches and also has a history of an acoustic neuroma. The proband's father is reported to have schizophrenia with a strong family history of personality disorder and alcohol abuse but was not available to take part in the study.

On the maternal side, the proband has one full sister, one older half-brother and one younger half-sister. His half-siblings have different fathers. The proband's full-sister was born with congenital hand malformation and dilated kidneys, but she is reportedly otherwise well and sociable with no cognitive problems. The proband's half-brother has a diagnosis of high functioning ASD and Tourette's syndrome. This sibling's father comes from a family with a history of bipolar disorder. The proband's half-sister was born with congenital hypothyroidism and cerebral atrophy. She is otherwise well. On the maternal side, there are four half-uncles, two from the maternal grandfather and two from the maternal grandmother (Figure 2A, FII-4 and FII-5). Both biological sons of the maternal grandfather are affected, FII-4 with early-onset depression and anxiety currently on sertraline and FII-5 with epilepsy (since age 12, currently in his 20's). The other two uncles (FII-2 and FII-3) are reportedly well. The proband's maternal grandmother (FI-4) has fibromyalgia and diabetes the maternal

grandfather (FI-3) was dismissed from his employment in the armed forces due to personality disorder.

Genetic analyses

We performed exome sequencing and analysis of a cohort of 29 families with COP (Methods). Table 2 summarises the number of qualifying variants per analysis type at each stage. Among 11 and 8 qualifying variants for probands 1 and 2, respectively, we noted that each has a damaging newly hemizygous variant in *L1CAM*. This was the only gene that had a qualifying variant in more than one family.

p.Arg846His in proband 1

The variant in proband 1 is a missense NM_001278116.1:c.2537G>A (p.Arg846His), falling on the Fibronectin type III-3 (Fn3) domain of the L1 protein. This variant has an allele frequency of 0.0000327 in gnomAD, carried in heterozygous form by six females and in hemizygous form in one Latino male. None of these carriers belong to the gnomAD psychiatric cohorts. The variant is predicted to be tolerated by SIFT[31], probably damaging by polyphen2[32] and has a CADD[33] score of 20.5, suggesting that is within the top 1% of deleterious variants. This variant has been reported once in ClinVar[34] as a variant of uncertain significance in an individual with a history of neurodevelopmental disorder. This variant is absent in SCHEMA[35].

p.Pro145Ser in proband 2

The variant in proband 2 is a missense NM_001278116.1:c.433C>T (p.Pro145Ser), falling on the Ig-like C2-type 2 (Ig2) domain of the L1 protein. This variant has an allele frequency of 0.00001471 in gnomAD, carried in heterozygous form by two females and in hemizygous

form in one African male. None of these carriers belong to the gnomAD psychiatric cohorts. The variant is predicted to be deleterious by SIFT, possibly damaging by polyphen2 and has a CADD score of 23, suggesting that it is within the top 1% of deleterious variants. This variant has never been reported in ClinVar. This variant has been observed in 3/194,644 controls (2 of African ancestry and 1 of European ancestry) and 0/48,496 cases in SCHEMA.

Observing two such rare damaging variants in L1CAM by chance is unlikely

In order to assess how likely it was to see two such rare, damaging variants *L1CAM* in our cohort of 29 probands, we looked at how frequent similarly damaging variants are in gnomAD, after applying the same filters we used to identify our reported variants. There are a maximum of 1,104 male individuals out of 76,702 total males (1.44%) with a *L1CAM* missense variant in gnomAD (sum of the allele counts across all 398 missense variants) assuming that these variants are each present in a different male, since there is no way to know if any male carries more than one. After removing variants seen in more than 2 males (Methods) the number is reduced to 217 (180 variants) (0.29%). Filtering for damaging missense variants indicated by polyphen probably/possibly damaging and CADD>15, reduces the count to 76 males (n=65 variants) (0.09%). This gives a population frequency for similar variants of 0.09% (76/76,702 total number of males in gnomAD), vs. 6.9% (2/29) in our cohort.

Sanger sequencing confirmation

Sanger sequencing confirmed the presence of the variants in the probands and their mothers (010132 and 011312) and their absence in the first-degree relatives (010133, 011314), as shown in Figures 1B and 2B. Extended family members were not available for analysis.

Discussion:

This paper describes the first report of two newly hemizygous rare damaging missense variants in two unrelated patients with Childhood Onset Psychosis (COP) in *L1CAM*. *L1CAM* is located in the Xq28 region and encodes the 220kDA neural cell adhesion module L1, a transmembrane glycoprotein.

The *L1CAM* family (*L1, CHL1, NrCAM* and neurofascin) are highly expressed in the brain and they play an important role in human disease[36]. Mutations in *L1CAM* are a widely reported cause of X-linked intellectual disability through a neurological phenotype termed CRASH (OMIM #303350) or L1 syndrome, which encompasses partial agenesis of corpus callosum, hydrocephalus, MASA syndrome, spasticity and adducted thumbs[20]. In most families with L1 syndrome, affected males tend to die soon after birth, however phenotypic variability has been observed both within and between families[37,38]. Gecz and colleagues report one missense variant (p.Asp202Asn) segregating in two unrelated families with mild-moderate intellectual disability without obvious L1 syndrome features[39] suggesting *L1CAM* mutations are subject to variable penetrance and expressivity. To date, at least 390 disease-causing *L1CAM* mutations have been identified[34]. Mutations in the intracellular domains are associated with severe phenotypes, with missense mutations in the extracellular domain leading to a milder phenotype and those in the cytoplasmic domain having the least severe effects[40]. Both missense variants described in this report localise to the extracellular domain of the L1 protein and, as such, if pathogenic, would be expected to be associated with milder forms of L1 syndrome. The majority of known pathogenic variants in this gene are loss of functions however there is evidence for missense pathogenicity in this gene. Out of the 122 missense variants reported in *L1CAM* in ClinVar, 64 (52.45%) have been assigned as (likely) pathogenic, 47 as uncertain significance and the remainder as (likely) benign. Indeed, there

seems to be no clustering of pathogenic missense variants to any specific gene region or exons. In addition, other genes in the *L1CAM* family, namely *NrCAM* and *CHL1* have previously been linked to autism[41] and schizophrenia[42], respectively.

Proband 1 has no positive family history of neuropsychiatric illness. Both his mother who is a heterozygous carrier of the mutation and his brother and sisters (for whom we have no genetic data) are unaffected. The mother's father and brothers are reportedly unaffected. One of the unaffected brothers has a son with unspecified speech problems (Figure 1A). This suggests either variable penetrance or that the variant is *de novo* in the mother. It is interesting to note that the locus of this variant (X-153131169-C-T) is multiallelic. There is a C/A variant (as opposed to C/T, which is our variant of interest) that results in a leucine (p.Arg846Leu) rather than a histidine (p.Arg84His) amino acid change. The C/A variant leading to a change of arginine to leucine is very rare with an allele frequency of 0.0001072, present in five European males in hemizygous form and carried by 22 European females in heterozygous form.

Proband 2 has a positive history of neuropsychiatric illness on both the maternal and paternal side. His mother is heterozygous for the variant and has a diagnosis of OCD, social anxiety, cluster headaches with a history of an acoustic neuroma. Since the mother is a heterozygous carrier, we tested to see if the proband's half-brother with high functioning Asperger's and Tourette syndrome also carries the variant. We found that he does not, however his biological father (FII-6) is reported to have personality disorder and therefore the half-brother (FIII-1) could have a different genetic cause of illness to proband 2. Proband 2's maternal grandfather (FI-3) may be carrying the L1CAM variant and manifesting as psychiatric phenotype as he was reportedly discharged from the armed services due to having personality disorder. His

two sons, the maternal half uncles are both affected with neuropsychiatric illness (Figure 2A). FII-4 suffers from early-onset depression and anxiety and is on medication for sertraline and FII-5 has suffered from epilepsy since age 12 (currently in his 20's). We do not have DNA available for any of the affected maternal male relatives, however phenotypically we can see a pattern of neuropsychiatric illness in the male relatives of the proband's mother and the phenotypic segregation appears to resemble an X-linked recessive disease model.

Genetic background can modulate the severity of the cognitive and neurodevelopmental phenotypes[43–45]. Both probands carry a number of rare and private damaging variants in genes previously implicated in neurodevelopment. It is likely that the psychosis observed is a result of the *L1CAM* variants(s) occurring on a background of other rare damaging variants leading to a general neuronal vulnerability in the probands. Table 3 summarises the list of all other qualifying variants the probands carry. Proband 1 carries 10 other rare damaging variants, four inherited from his father and six from his mother, both reportedly unaffected. Three of the genes with a qualifying variant in proband 1, namely *L1CAM*, *SCN1A* and *SPTAN1* fall on the L1-Ankyrins interaction pathway. The L1 protein binds to Ankyrins (an actin/spectrin adapter protein) through a conserved motif in the cytoplasmic domain of the L1 protein[46,47]. Studies in mice indicate that disruption of this pathway can lead to a decrease in synapse formation at GABAergic inhibitory interneurons in the developing mouse cortex[48]. However, since the conserved motif is situated in the cytoplasmic domain (residues 1142-1257) and the *L1CAM* variant seen in proband 1 is on residue 846 (extracellular domain), it is unlikely that this interaction is being directly affected. Interestingly, both *SPTAN1* and *SCN1A* are neurodevelopmental genes[49] and disruptions in these genes have been reported in childhood-onset schizophrenia[50], epilepsy[51,52], intellectual disability[53,54], autism spectrum disorders[55,56] and developmental delay[57].

In addition to the *L1CAM* variant, proband 2 has inherited four other rare damaging variants from his mother, in genes previously associated with neurodevelopmental disorders[58,59]. He carries a damaging missense variant in *TCF20* which is either *de novo* or inherited from his affected father. *TCF20* is a neurodevelopmental gene[49] and disruptions in this gene have been reported in a range of neurodevelopmental disabilities including intellectual disability[60–62] and autism[61].

In this paper we have outlined the first report of rare newly hemizygous damaging missense variants in *L1CAM* in two patients with COP. There have been no previous reports of this gene being associated with psychosis. The evidence that *L1CAM* is important in childhood-onset psychosis is strong but remains inconclusive until additional cases are identified.

Figures and Tables

Phenotypic features	Proband 1	Proband 2
Neurodevelopmental features		
Age of onset of psychosis	9	7
Age at diagnosis of COP	9	10
Cognitive functioning	Borderline IQ (WASI-II FSIQ 72)	Average IQ (WASI-II FSIQ 105)
Autism	yes	yes
Epilepsy	no	no
Other neurodevelopmental disorder	no	no
speech delay	yes	yes
L1 syndrome features:		
Spasticity	no	no
Adducted thumbs	no	no
Agenesis of corpus callosum	no	no
Hydrocephalus	no	no
Other medical conditions	Tics	Fatigue, Joint pains, Headaches, Transient memory loss, Poor concentration, Nosebleeds, Sensitivity to noise

Table 1. Summary of key clinical features
Clinical information was obtained from clinical records and interview notes with the parents.

	Proband 1	Proband 2
Total Qualifying Variants	**11**	**8**
Present in ClinVar	0	0
*de novo**	0	3
Newly homozygous	0	0
Newly hemizygous	1	1
Compound Heterozygous	0	0
Inherited rare damaging	10	4

Table 2. Variant counts
The total qualifying variants as described in section 4.3 were 11 in proband 1 and 8 in proband 2.
* indicates that the *de novo* variant is either *de novo* or inherited from a father with schizophrenia (but we cannot confirm as proband's father was not available to take part into the study).

ID	Variant (hg19)	Gene	HGVS change	gnomAD frequency (v2.1.1)	Polyphen (Humvar)	CADD score	RVIS (%)	Type of Qualifying variant	Pat / mat
1	2-96780549-C-G	*ADRA2B*	p.Trp447Ser	Absent	benign (0.11)	27.1	36	inherited	mat
1	1-201020185-C-T	*CACNA1S*	p.Gly1347Glu	Absent	probably (1)	26.3	NA	inherited	pat
1	18-56964086-C-T	*CPLX4*	p.Met109Ile	1 AFR female, 1 Other* male (Non-Neuro) 1 AFR male (Psych)	probably (0.954)	24.8	46	inherited	mat
1	10-103530238-G-A	*FGF8*	p.Arg195Trp	2 AFR females and 1 NFE female (Non-Neuro)	probably (1)	32	31	inherited	mat
1	8-30550490-G-C	*GSR*	p.Ser293Cys	2 AFR (male and female) Non-Neuro and 1 FIN male (Psych)	possibly (0.638)	26.5	29	inherited	mat
1	6-160501274-G-A	*IGF2R*	p.Asp1934Asn	1 SA female Non-Neuro and 2 AFR females (Psych)	possibly (0.5)	24.5	2	inherited	pat
1	X-153131169-C-T	*L1CAM*	p.Arg846His	1 Hemizygous Latino male (Non-Neuro)	possibly (0.881)	32	5	newly hemizygous	mat
1	19-47542737-G-A	*NPAS1*	p.Ala293Thr	1 AFR male (Psych)	possibly (0.546)	25.2	NA	inherited	pat
1	2-166900202-C-G	*SCN1A*	p.Asp674His	Absent	probably (0.942)	25.9	4	inherited	mat
1	9-131370209-G-A	*SPTAN1*	p.Gly1409Arg	1 Latino, 1 NFE male (Non-Neuro)	probably (0.986)	26.1	0.3	inherited	pat
1	10-114710703-A-G	*TCF7L2*	p.Glu63Gly	Absent	probably (0.995)	23.6	38	inherited	mat
2	17-36666568-G-T	*ARHGAP23*	p.Arg1279Leu	Absent	possibly (0.814)	22.8	NA	*de novo***	-

2	11-10800235-G-C	*CTR9*	p.Arg1035Ser	Absent	benign (0.227)	19.24	9	inherited	mat
2	6-170592661-A-C	*DLL1*	p.Val569Gly	1 NFE female (Non-Neuro)	benign (0.183)	16.57	44	*de novo***	-
2	8-132997193-T-A	*EFR3A*	p.Asn585Lys	2 NFE males (Non-Neuro)	probably (0.999)	23.7	33	inherited	mat
2	X-153136602-G-A	*L1CAM*	p.Pro145Ser	1 Hemizygous AFR male (Non-Neuro)	probably (1)	23	3	newly hemizygous	mat
2	11-33891153-G-C	*LMO2*	c.-221C>G	1 Latino female (Non-Neuro)	NA	15.42	37	inherited	mat
2	22-42606541-C-A	*TCF20*	p.Ala1591Ser	Absent	benign (0.116)	21.4	0.8	*de novo***	-
2	16-74990420-C-A	*WDR59*	p.Val65Leu	Absent	benign (0.015)	28.8	16	inherited	mat

Table 3. List of all qualifying variants identified in the two probands.

*In gnomAD v2.1.1. the population 'Other' indicates samples did not fit into the six principle components of ancestry (European, African, South Asian, East Asian, Latino, Ashkenazi Jewish) computed for all samples in the dataset.

** indicates that the *de novo* variant is either *de novo* or inherited from a father with schizophrenia

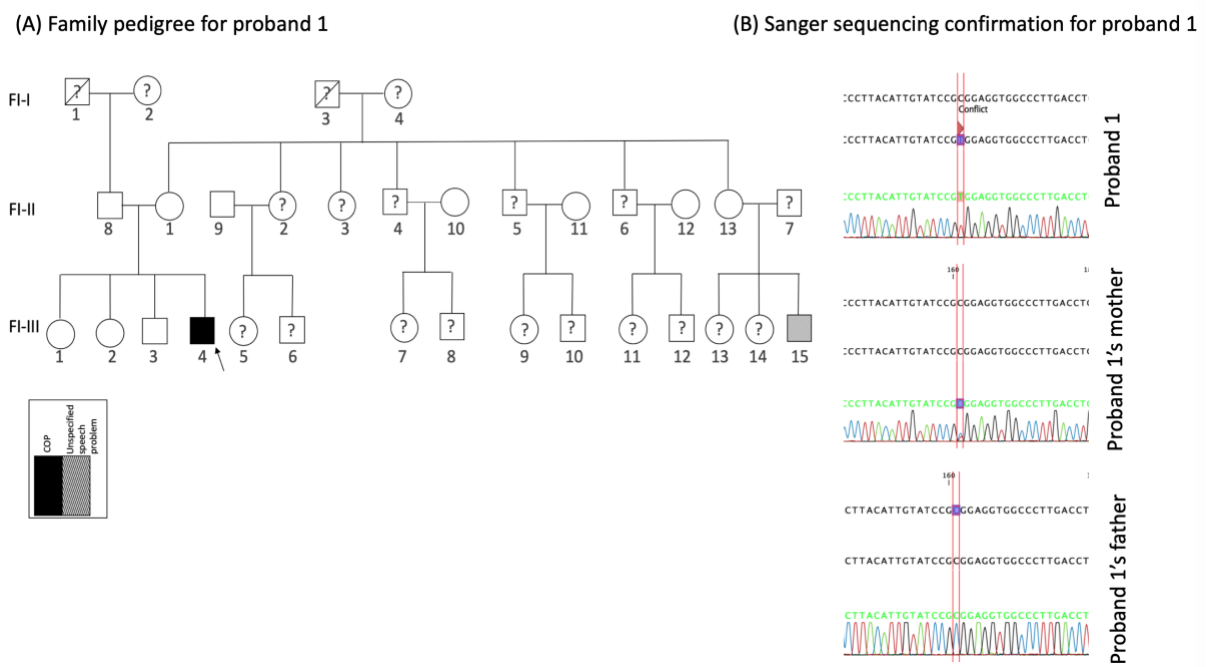

Figure 1. Figure showing the family pedigree and sanger sequencing results for proband 1. Figure 1A depicts the proband marked by the arrow. Figure 1B shows the sanger sequencing traces in proband 1, his mother and his father.

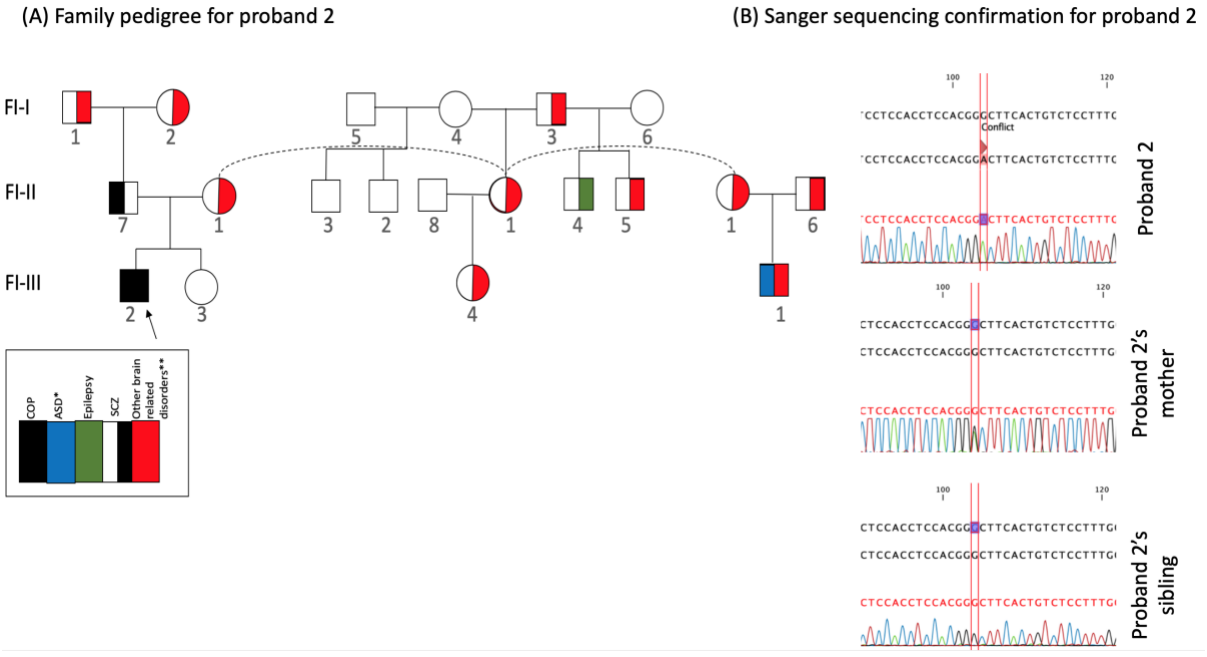

Figure 2. Figure showing the family pedigree and sanger sequencing results for proband 2. Figure 2A depicts proband marked by the arrow. Figure 2B shows the sanger sequencing traces in proband 2, his mother and half-brother.

Conflict of Interest statement
The authors have declared that there are no conflicts of interest in relation to the subject of this study.

Acknowledgements
We would like to thank the patients and their families for participating in the study.
We would like to acknowledge the following individuals or groups for the contributions of control samples: D. Valle; J. Hoover-Fong; N. Sobriera; R. Buckley; R. Ottman; V. Shashi; Epilepsy Genetics Initiative, A Signature Program of CURE; S. Berkovic; I. Scheffer; B. Grinton; C. Depondt; S. Sisodiya; G. Cavalleri; N. Delanty; the Epi4K Consortium and Epilepsy Phenome/Genome Project; S. Hirose; the ALS Sequencing Consortium; the Washington University Neuromuscular Genetics Project; G. Nestadt; J. Samuels; Y. Wang; E. Pras; D. Lancet; M. Hauser; S. Delaney; Undiagnosed Diseases Network; Utah Foundation of Biomedical Research. The collection of control samples and data was funded in part by: NINDS Award (RC2NS070344); NIH (U54 NS078059); NIH (P01 HD080642); The J. Willard and Alice S. Marriott Foundation; The Muscular Dystrophy Association; The Nicholas Nunno Foundation; The JDM Fund for Mitochondrial Research; The Arturo Estopinan TK2 Research Fund; New York-Presbyterian Hospital; Columbia University College Physicians and Surgeons; Columbia University Medical Center; UCB; Epi4K Gene Discovery in Epilepsy study (NINDS U01-NS077303) and the Epilepsy Genome/Phenome Project (EPGP - NINDS U01-NS053998); Ellison Medical Foundation New Scholar award (AG-NS-0441-08 to O.C.); Biogen; OCD Rare (1R01MH097971-01A1); an American Academy of Child and Adolescent Psychiatry (AACAP) Pilot Research Award; NIH (U01HG007672); the Stanley Institute for Cognitive Genomics at Cold Spring Harbor Laboratory.

References

1. Tiffin PA, Kitchen CEW. Incidence and 12-month outcome of childhood non-affective psychoses: British national surveillance study. *Br J Psychiatry*. 2015;206(6):517-518. doi:10.1192/bjp.bp.114.158493
2. Ahn K, Gotay N, Andersen TMTM, et al. High rate of disease-related copy number variations in childhood onset schizophrenia. *Mol Psychiatry*. 2013;19(5):568-572. doi:10.1038/mp.2013.59
3. Ambalavanan A, Girard SL, Ahn K, et al. De novo variants in sporadic cases of childhood onset schizophrenia. *Eur J Hum Genet*. 2015;24(August):1-5. doi:10.1038/ejhg.2015.218
4. Bateman RJ, Aisen PS, De Strooper B, et al. Autosomal-dominant Alzheimer's disease: a review and proposal for the prevention of Alzheimer's disease. *Alzheimers Res Ther*. 2011;3(1):1. doi:10.1186/alzrt59
5. Vaxillaire M, Froguel P. Genetic basis of maturity-onset diabetes of the young. *Endocrinol Metab Clin North Am*. 2006;35(2):371-384, x. doi:10.1016/j.ecl.2006.02.009
6. Wang L, Fan C, Topol SE, Topol EJ, Wang Q. Mutation of MEF2A in an Inherited Disorder with Features of Coronary Artery Disease. *Science (80-)*. 2003;302(5650):1578-1581. doi:10.1126/science.1088477
7. Curie A, Lesca G, Bussy G, et al. Asperger syndrome and early-onset schizophrenia associated with a novel MECP2 deleterious missense variant. *Psychiatr Genet*. 2017. doi:10.1097/YPG.0000000000000165
8. Mavros CF, Brownstein CA, Thyagrajan R, et al. De novo variant of TRRAP in a patient with very early onset psychosis in the context of non-verbal learning disability

and obsessive-compulsive disorder: a case report. *BMC Med Genet*. 2018. doi:10.1186/s12881-018-0711-9

9. Addington AM, Gauthier J, Piton A, et al. A novel frameshift mutation in UPF3B identified in brothers affected with childhood onset schizophrenia and autism spectrum disorders. *Mol Psychiatry*. 2011. doi:10.1038/mp.2010.59
10. Zhou D, Gochman P, Broadnax DD, Rapoport JL, Ahn K. 15q13.3 duplication in two patients with childhood-onset schizophrenia. *Am J Med Genet Part B Neuropsychiatr Genet*. 2016;171(6):777-783. doi:10.1002/ajmg.b.32439
11. Brownstein CA, Engle EC, Towne MC, et al. Overlapping 16p13.11 deletion and gain of copies variations associated with childhood onset psychosis include genes with mechanistic implications for autism associated pathways: Two case reports. *Am J Med Genet Part A*. 2016;170(5):1165-1173. doi:10.1002/ajmg.a.37595
12. Smedemark-Margulies N, Brownstein CA, Vargas S, et al. A novel de novo mutation in *ATP1A3* and childhood-onset schizophrenia. *Mol Case Stud*. 2016. doi:10.1101/mcs.a001008
13. Chaumette B, Ferrafiat V, Ambalavanan A, et al. Missense variants in ATP1A3 and FXYD gene family are associated with childhood-onset schizophrenia. *Mol Psychiatry*. doi:10.1038/s41380-018-0103-8
14. Rosenthal A, Jouet M, Kenwrick S. Aberrant splicing of neural cell adhesion molecule L1 mRNA in a family with X–linked hydrocephalus. *Nat Genet*. 1992. doi:10.1038/ng1092-107
15. Cohen NR, Taylor JSH, Scott LB, Guillery RW, Soriano P, Furley A. Errors in corticospinal axon guidance in mice lacking the neural cell adhesion molecule L1. *Curr Biol*. 1998. doi:10.1016/S0960-9822(98)70017-X
16. Demyanenko GP, Tsai AY, Maness PF. Abnormalities in neuronal process extension,

hippocampal development, and the ventricular system of L1 knockout mice. *J Neurosci*. 1999. doi:10.1523/JNEUROSCI.19-12-04907.1999

17. Dahme M, Bartsch U, Martini R, Anliker B, Schachner M, Mantei N. Disruption of the mouse L1 gene leads to malformations of the nervous system. *Nat Genet*. 1997. doi:10.1038/ng1197-346

18. Lek M, Karczewski KJ, Minikel E V., et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285-291. doi:10.1038/nature19057

19. Gussow AB, Petrovski S, Wang Q, Allen AS, Goldstein DB. The intolerance to functional genetic variation of protein domains predicts the localization of pathogenic mutations within genes. *Genome Biol*. 2016;17(1). doi:10.1186/s13059-016-0869-4

20. Hamosh A, Scott AF, Amberger J, Valle D, McKusick VA. Online Mendelian Inheritance in Man (OMIM). *Hum Mutat*. 2000;15(1):57-61. doi:10.1002/(SICI)1098-1004(200001)15:1<57::AID-HUMU12>3.0.CO;2-G

21. Fransen E, Van Camp G, Vits L, Willems PJ. L1-associated diseases: Clinical geneticists divide, molecular geneticists unite. *Hum Mol Genet*. 1997. doi:10.1093/hmg/6.10.1625

22. Kanemura Y, Okamoto N, Sakamoto H, Shofuda T, Kamiguchi H, Yamasaki M. Molecular mechanisms and neuroimaging criteria for severe L1 syndrome with X-linked hydrocephalus. *J Neurosurg*. 2006. doi:https://doi.org/10.3171/ped.2006.105.5.403

23. Weller S, Gärtner J. Genetic and clinical aspects of X-linked hydrocephalus (L1 disease): Mutations in the *L1CAM* gene. *Hum Mutat*. 2001;18(1):1-12. doi:10.1002/humu.1144

24. Vos YJ, de Walle HEK, Bos KK, et al. Genotype-phenotype correlations in L1

syndrome: a guide for genetic counselling and mutation analysis. *J Med Genet*. 2010;47(3):169-175. doi:10.1136/jmg.2009.071688

25. Isohelix. *Instructions for Isohelix GeneFix ™ Saliva-Prep DNA Kit : GSP- 48 / GSP-12 / GSP-2*.; 2015.

26. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010. doi:10.1101/gr.107524.110.20

27. Depristo MA, Banks E, Poplin R, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet*. 2011. doi:10.1038/ng.806

28. IGV (Integrative Genomic Viewer). Integrative Genomics Viewer. *Broad Inst*. 2013. doi:10.1038/nbt0111-24

29. Untergasser A, Nijveen H, Rao X, Bisseling T, Geurts R, Leunissen JAM. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Res*. 2007;35(Web Server issue):W71-4. doi:10.1093/nar/gkm306

30. Sinnwell JP, Therneau TM, Schaid DJ. The kinship2 R package for pedigree data. *Hum Hered*. 2014;78(2):91-93. doi:10.1159/000363105

31. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*. 2009;4(7):1073-1082. doi:10.1038/nprot.2009.86

32. Adzhubei IA, Schmidt S, Peshkin L, et al. *A Method and Server for Predicting Damaging Missense Mutations*. Vol 7.; 2010:248-249. doi:10.1038/nmeth0410-248

33. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46(3):310-315. doi:10.1038/ng.2892

34. Landrum MJ, Lee JM, Benson M, et al. ClinVar: Public archive of interpretations of clinically relevant variants. *Nucleic Acids Res*. 2016;44(D1):D862-D868. doi:10.1093/nar/gkv1222

35. SCHEMA browser. http://schema.broadinstitute.org/. Accessed October 16, 2019.

36. Chen L, Zhou S. "CRASH"ing with the worm: Insights into L1CAM functions and mechanisms. *Dev Dyn*. 2010. doi:10.1002/dvdy.22269

37. Otter M, Wevers M, Pisters M, et al. A novel mutation in L1CAM causes a mild form of L1 syndrome: a case report. *Clin case reports*. 2017;5(8):1213-1217. doi:10.1002/ccr3.1038

38. Schrander-Stumpel C, Höweler C, Jones M, et al. Spectrum of X-linked hydrocephalus (HSAS), MASA syndrome, and complicated spastic paraplegia (SPG1): Clinical review with six additional families. *Am J Med Genet*. 1995;57(1):107-116. doi:10.1002/ajmg.1320570122

39. Shaw M, Yap TY, Henden L, et al. Identical by descent L1CAM mutation in two apparently unrelated families with intellectual disability without L1 syndrome. *Eur J Med Genet*. 2015. doi:10.1016/j.ejmg.2015.04.004

40. Bateman A, Jouet M, MacFarlane J, Du JS, Kenwrick S, Chothia C. Outline structure of the human L1 cell adhesion molecule and the sites where mutations cause neurological disorders. *EMBO J*. 1996;15(22):6050-6059. http://www.ncbi.nlm.nih.gov/pubmed/8947027. Accessed September 7, 2019.

41. Marui T, Funatogawa I, Koishi S, et al. Association of the neuronal cell adhesion molecule (NRCAM) gene variants with autism. *Int J Neuropsychopharmacol*. 2009. doi:10.1017/S1461145708009127

42. Sakurai K, Migita O, Toru M, Arinami T. An association between a missense polymorphism in the close homologue of L1 (CHL1, CALL) gene and schizophrenia.

Mol Psychiatry. 2002. doi:10.1038/sj.mp.4000973

43. Todarello G, Feng N, Kolachana BS, et al. Incomplete penetrance of NRXN1 deletions in families with schizophrenia. *Schizophr Res*. 2014;155(1-3):1-7. doi:10.1016/j.schres.2014.02.023

44. Hammer MF, Ishii A, Johnstone L, et al. Rare variants of small effect size in neuronal excitability genes influence clinical outcome in Japanese cases of SCN1A truncation-positive Dravet syndrome. *PLoS One*. 2017. doi:10.1371/journal.pone.0180485

45. Pizzo L, Jensen M, Polyak A, et al. Rare variants in the genetic background modulate the expressivity of neurodevelopmental disorders. 2018. doi:10.1101/257758

46. Davis JQ, Bennett V. Ankyrin binding activity shared by the neurofascin/L1/NrCAM family of nervous system cell adhesion molecules. *J Biol Chem*. 1994. http://www.jbc.org/content/269/44/27163.long.

47. Bennett V, Baines AJ. Spectrin and ankyrin-based pathways: Metazoan inventions for integrating cells into tissues. *Physiol Rev*. 2001. doi:10.1152/physrev.2001.81.3.1353

48. Guan H, Maness PF. Perisomatic GABAergic innervation in prefrontal cortex is regulated by ankyrin interaction with the L1 cell adhesion molecule. *Cereb Cortex*. 2010. doi:10.1093/cercor/bhq016

49. Wright CF, Fitzgerald TW, Jones WD, et al. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data Summary Background. *Lancet*. 2015;385(9975):1305-1314. doi:10.1016/S0140-6736(14)61705-0

50. Papp-Hertelendi R, Tényi T, Hadzsiev K, Hau L, Benyus Z, Csábi G. First report on the association of SCN1A mutation, childhood schizophrenia and autism spectrum disorder without epilepsy. *Psychiatry Res*. 2018. doi:10.1016/j.psychres.2018.07.028

51. Tohyama J, Nakashima M, Nabatame S, et al. SPTAN1 encephalopathy: Distinct phenotypes and genotypes. *J Hum Genet*. 2015. doi:10.1038/jhg.2015.5

52. Escayg A, Goldin AL. Sodium channel SCN1A and epilepsy: Mutations and mechanisms. *Epilepsia*. 2010. doi:10.1111/j.1528-1167.2010.02640.x

53. Rubinstein M, Patowary A, Stanaway IB, et al. Association of rare missense variants in the second intracellular loop of Na v 1.7 sodium channels with familial autism. *Mol Psychiatry*. 2018. doi:10.1038/mp.2016.222

54. Hamdan FF, Saitsu H, Nishiyama K, et al. Identification of a novel in-frame de novo mutation in SPTAN1 in intellectual disability and pontocerebellar atrophy. *Eur J Hum Genet*. 2012. doi:10.1038/ejhg.2011.271

55. O'Roak BJ, Deriziotis P, Lee C, et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet*. 2011;43(6):585-589. doi:10.1038/ng.835

56. Weiss LA, Escayg A, Kearney JA, et al. Sodium channels SCN1A, SCN2A and SCN3A in familial autism. *Mol Psychiatry*. 2003. doi:10.1038/sj.mp.4001241

57. McRae JF, Clayton S, Fitzgerald TW, et al. Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017;542(7642):433-438. doi:10.1038/nature21062

58. Weckhuysen S, Marsan E, Lambrecq V, et al. Involvement of GATOR complex genes in familial focal epilepsies and focal cortical dysplasia. *Epilepsia*. 2016. doi:10.1111/epi.13391

59. Gupta AR, Pirruccello M, Cheng F, et al. Rare deleterious mutations of the gene EFR3A in autism spectrum disorders. *Mol Autism*. 2014. doi:10.1186/2040-2392-5-31

60. Lelieveld SH, Reijnders MRFF, Pfundt R, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci*. 2016;19(9):1194-

1196. doi:10.1038/nn.4352

61. Torti E, Keren B, Palmer EE, et al. Variants in TCF20 in neurodevelopmental disability: description of 27 new patients and review of literature. doi:10.1038/s41436-019

62. Schäfgen J, Cremer K, Becker J, et al. De novo nonsense and frameshift variants of TCF20 in individuals with intellectual disability and postnatal overgrowth. *Eur J Hum Genet*. 2016. doi:10.1038/ejhg.2016.90