Received: 28 December 2016 | Revised: 9 May 2017 | Accepted: 5 July 2017
DOI: 10.1002/ajmg.b.32574

RESEARCH ARTICLE

WILEY American Journal of Medical Genetics Part B Neuropsychiatric Genetics

Exome sequencing reveals *NAA15* and *PUF60* as candidate genes associated with intellectual disability

Jin J. Zhao[1] | Jonatan Halvardson[1] | Cecilia S. Zander[1] | Ammar Zaghlool[1] | Patrik Georgii-Hemming[1,2] | Else Månsson[3] | Göran Brandberg[4] | Helena E. Sävmarker[5] | Carina Frykholm[1] | Ekaterina Kuchinskaya[6] | Ann-Charlotte Thuresson[1] | Lars Feuk[1]

[1] Department of Immunology, Genetics and Pathology, Science for Life Laboratory Uppsala, Uppsala University, Uppsala, Sweden

[2] Department of Molecular Medicine and Surgery, Karolinska Institute, Karolinska University Hospital Solna, Stockholm, Sweden

[3] Department of Pediatrics, Örebro University Hospital, Örebro, Sweden

[4] Pediatric Clinic, Falun, Sweden

[5] Department of Pediatrics, Gävle Hospital, Gävle, Sweden

[6] Department of Clinical Genetics, and Department of Clinical Medicine, Linköping University, Linköping, Sweden

Correspondence
Dr. Lars Feuk, Department of Immunology, Genetics and Pathology, Science for Life Laboratory Uppsala, Uppsala University, Box 815, BMC B11:4, Uppsala University, SE-751 08 Uppsala, Sweden.
Email: lars.feuk@igp.uu.se

Funding information
Sävstaholm Society; Regionala forskningsrådet; European Research Council; Medicinska Forskningsrådet

Intellectual Disability (ID) is a clinically heterogeneous condition that affects 2–3% of population worldwide. In recent years, exome sequencing has been a successful strategy for studies of genetic causes of ID, providing a growing list of both candidate and validated ID genes. In this study, exome sequencing was performed on 28 ID patients in 27 patient-parent trios with the aim to identify de novo variants (DNVs) in known and novel ID associated genes. We report the identification of 25 DNVs out of which five were classified as pathogenic or likely pathogenic. Among these, a two base pair deletion was identified in the PUF60 gene, which is one of three genes in the critical region of the 8q24.3 microdeletion syndrome (Verheij syndrome). Our result adds to the growing evidence that PUF60 is responsible for the majority of the symptoms reported for carriers of a microdeletion across this region. We also report variants in several genes previously not associated with ID, including a de novo missense variant in NAA15. We highlight NAA15 as a novel candidate ID gene based on the vital role of NAA15 in the generation and differentiation of neurons in neonatal brain, the fact that the gene is highly intolerant to loss of function and coding variation, and previously reported DNVs in neurodevelopmental disorders.

KEYWORDS
exome sequencing, intellectual disability, *NAA15*, *PUF60*

1 | INTRODUCTION

Intellectual disability (ID) is an early, before adulthood, onset condition characterized by significantly limited ability in learning, reasoning and communicating (American Association on Intellectual and Developmental Disabilities, 2017; World Health Organization, Reginal office for Europe, 2010). ID is defined by an IQ score lower than 70, and has an estimated worldwide prevalence of 2–3% (Ropers, 2010). While the etiologies of ID are variable and include environmental factors such as infections and injuries, it is well established that genetics play a major role, especially in severe cases of ID (Vissers, Gilissen, & Veltman, 2016).

Genetic causes of ID range from chromosome abnormalities and copy number variations (CNVs) to single nucleotide variants (SNVs),

Ann-Charlotte Thuresson and Lars Feuk contributed equally to this work.

including inherited recessive variants and dominant de novo variants. Early genetic diagnosis of ID was limited to large structural abnormalities and targeted tests of known ID genes, which explain 3–6% of cases (Knight et al., 1999). The introduction of chromosomal microarray analysis (CMA) enables genome-wide investigation with increased resolution, which results in clinical diagnosis of additional 15–20% of ID patients (de Vries et al., 2005; Miller et al., 2010). The implementation of whole exome sequencing (WES) and whole genome sequencing (WGS) has led to a rapid increase in the discovery of genes with variants associated with ID. WES in trio families has been an especially successful strategy to identify genetic causes of unexplained ID cases, yielding clinically significant findings in 20–30% of patients previously screened by CMA (de Ligt et al., 2012; Miller et al., 2010; Tammimies et al., 2015). Despite these advances, nearly half of the patients with sporadic ID subjected to genetic testing do not receive a molecular diagnosis. Novel candidate ID genes can be identified by observing variations in the same gene in multiple unrelated ID patients with phenotypic similarity, which underlines the importance of detailed documentation of patient phenotype and the importance of sharing data between clinical laboratories and research initiatives.

In this study, WES was performed in 27 patient-parent trios, where the proband had an ID diagnosis. We report the identification of 25 DNVs, with five variants classified as pathogenic or likely pathogenic according to the American College of Medical Genetics and Genomics (ACMG) guidelines (Richards et al., 2015). Of the DNVs we identify, we specifically highlight variants in *PUF60* and *NAA15* as variants providing novel insight into the etiology of ID.

2 | MATERIALS AND METHODS

2.1 | Study design and patients

The participating patients and parents were recruited between 2012 and 2015 in collaboration with the Genetic Diagnostics Unit at Uppsala University Hospital. Ethical approval for WES was received from the Uppsala Ethical Review Board and informed consent was obtained from the parents of all patients. The selection criteria for patients included ID/DD and dysmorphology/congenital malformations while parents had to be healthy with no family history of neurodevelopmental disorders. All patients had previously been screened with CMA (250 K Nsp Array, GenomeWide SNP Array V.6.0, or CytoScan HD [Affymetrix, Santa Clara, CA]) and no pathogenic CNVs had been detected. Genomic DNA was extracted from peripheral blood leucocytes according to standard procedures.

2.2 | Exome sequencing and analysis

Exome was enriched using SureSelect v2−5 (Agilent, Santa Clara, CA) and sequenced on SOLiD, Illumina, or IonProton platforms to at least 30x coverage. All reads were mapped to human reference genome Hg19. SOLiD reads were mapped using Bioscope (Life Technologies) and Life Scope (Life Technologies). Illumina reads were mapped using BWA (Li & Durbin, 2009), and IonProton reads were mapped using the Torrent suit software (Life Technologies). All programs used for mapping were run using default settings. Variants were called from SOLiD and Illumina reads using Genome Analysis Toolkit (GATK) and the standard GATK workflow (Broad Institute). Variants were called from IonProton reads using the Torrent suit software (Life Technologies). SNVs were filtered against our in-house database containing previously identified variants as well as dbSNP V.42 (non-flagged). Identified DNVs and inherited variants were validated by Sanger sequencing using standard protocols. Validated variants were interpreted according to the ACMG guidelines. The number of DNVs identified in patient cases and controls in a selected set of previous exome sequencing studies were counted.

3 | RESULTS

3.1 | Exome sequencing

Exome sequencing was performed in a total of 27 trio families (28 patients in 27 families) with ID/DD patients. Families were enrolled in the study if results from CMA were negative. A total of 25 heterozygous DNVs were identified within the protein-coding region of 16 patients, with detection of de novo events ranging from 0 to 4 sites per patient. All identified variants were validated by Sanger sequencing.

Out of the 25 validated DNVs we determined five variants in six patients in the genes *SMARCA4*, *FGFR2*, *SETD5*, *ASXL3* (two brothers) and *PUF60* to be pathogenic or likely pathogenic (Table 1), following the standards for interpretation of sequencing variants recommended by the ACMG (Richards et al., 2015). This results in a diagnostic yield of 21% (Table 1). We also identified a de novo heterozygous nonsynonymous variant of uncertain significance (VUS) in the gene *SLC13A5*. Although *SLC13A5* is a known ID gene, the inheritance pattern in previously reported patients was recessive, and the main symptom of the patients was early infantile epileptic encephalopathy, which is not present in our patient. Several identified DNVs in other genes represent interesting candidates, however, without previous reports and well-established functional studies, these variants were also classified as VUS (Supplemental Table S1).

3.2 | CADD prediction and previously reported data

In order to predict the deleteriousness of the identified DNVs the combined annotation-dependent depletion (CADD) score was calculated for all variants (Supplemental Table S1). CADD scores are a relative measurement of pathogenicity of genetic variants, with higher CADD scores indicating higher pathogenicity (Kircher et al., 2014). It is reported that pathogenic variants in the OMIM database are significantly enriched for CADD scores over 20 (Shyr et al., 2014). All variants classified as pathogenic or likely pathogenic in the

TABLE 1 A list of variants classified as pathogenic in this study, OMIM IDs for the associated disorder, phenotype of the patients and CADD score for each variant

Family	Gene (transcript)	Variant (cDNA level)	Variant (protein level)	Type	Associated disorder (OMIM designation and number)	Phenotype	CADD
Fam1	SMARCA4 (NM_001128849)	c.4543G>A	p.Glu1515Lys	MS	Coffin-Siris syndrome 4 MIM: 614609	Mild ID, hypotonia, delayed gross motor development, delayed speech and language development (at 4 years speech is fluent), prominent forehead, thin hair and eyebrows, hypertelorism, epicanthal folds, broad nasal bridge and tip, thin upper lip, slightly everted lower lip, high palate, small teeth, prognatism bilateral clinodactyli of fifth finger, hypoplasia of finger and toe nails, height -1SD	26.4
Fam2	FGFR2 (NM_000141.4)	c.1977G>T	p.Lys659Asn	MS	Apert syndrome MIM: 101200 Craniofacial-skeletal-dermatologic dysplasia MIM: 101600 Jackson-Weiss syndrome MIM: 123150 Pfeiffer syndrome MIM: 101600 Scaphocephaly, maxillary retrusion, and mental retardation MIM: 609579	Delayed psychomotor development, delayed speech and language development, ventriculomegaly, macrocephaly (OFC +2SD), prominent forehead and metopic ridge, epicanthus inversus, hypertelorism, proptosis, ptosis, strabismus, prognathism, recurrent otitis media	26.7
Fam3	SETD5 (NM_001080517.1)	c.2302C>T	p.Arg768*	NS	Mental retardation, autosomal dominant 23, MIM: 615761	Delayed psychomotor development, speech delay, square face, hypertelorism, low set protruding ears, large earlobe, preauricular ear pit (left), thick ala nasi, anteverted nares, long philtrum, macrostomia, thin lips, widely spaced teeth, widely spaced nipples, mitral insufficiency, neonatal pulmonary adaptation disturbance, thrombocytopenia	37
Fam4	PUF60 (NM_078480.1)	c.698_699del	p.Val233Alafs*8	FS	8q23.3 deletion syndrome/Verheij syndrome MIM: 615583	ID, speech delay, pre- and postnatal growth retardation (height -3.5SD), flat occiput, short neck, square face, full cheeks, bitemporal narrowing, low set eyebrows, hypotelorism, strabismus, hyperopia, low set ears with a thick helix, broad nasal tip, long philtrum, high palate, thin upper lip, bilateral bridged palmar crease, wide and short thorax, short and broad fingers and toes, tetralogy of fallot, broncial obstruction, recurrent epistaxis, mild hearing deficit, sleeping problems, hypermetropi	33
Fam5	ASXL3 (NM_030632.2)	c.4143dupC	p.Leu1395Profs*5	FS	Bainbridge-Ropers syndrome MIM: 615485	Two brothers, with variable expression, the older more severly affected than the younger. Both have ID (severe/moderate), autism, feeding difficulties in infancy, hypotonia, hypermobility, prominent arched eye-brows, slight hypertelorism, down-slanting eyes with a marked nasal bridged, high palate. In the elder brother complete absence of speech, a facial phenotype including a long face with temporal narrowing, deep set eyes, long eye lashes, reduced mimics, drooling and sweating from hands and feet was also noted.	35

FS, frameshift; ID, intellectual disability; MS, missense; NS, nonsense; SD, standard deviation.

TABLE 2 Results from previous studies showing the number of DNVs identified in cases and controls in previous exome sequencing studies, as well as the pLI and Z scores from ExAC for all genes with nonsense or nonsynonymous DNVs identified in this study

Gene	Cases	Controls	pLI	*Z* score
SMARCA4	2	0	1	8.36
FGFR2	0	1	1	2.74
SETD5	6	1	1	−0.04
PUF60	2	1	0.85	4.51
ASXL3	2	0	1	−0.94
NAA15	3	0	1	3.12
UTP14C	2	0	0	−0.88
TAOK1	1	0	1	5.17
PARP1	0	0	0.01	0.57
GJB4	0	0	0	−0.59
KLF16	0	0	0.47	3.17

present study had scores >20. Nonsense and loss of function (LoF) variants showed the highest scores (>30), including variants located in *SETD5*, *PUF60*, and *UTP14C*. We also note that nonsynonymous variants in the genes *NAA15*, *KLF16*, *TAOK1*, *IRF2BPL*, and *PARP1* all had CADD scores >20. Several of the nonsynonymous DNVs classified as VUS received a score <10, as did all synonymous DNVs identified.

To find further evidence for pathogenicity of the DNVs that were not scored as pathogenic, we combined data from 13 previous studies in ID and other neurodevelopmental disorders with overlapping genetic causes, including epilepsy and autism spectrum disorder. The combined data includes a total of 5 338 patient trios and 2 181 control trios. Of the 11 genes with nonsense, frameshift, or nonsynonymous DNVs identified in this study, seven genes had DNVs reported in patients in these 13 previous exome studies (Table 2). There were between 0 and 6 variants reported in cases for the pathogenic variants identified in this study. The most striking result is that three DNVs have previously been reported in patients with neurodevelopmental disorders in the gene *NAA15*, while no DNVs have been reported in controls. The DNV identified in *NAA15* in our study also had a high CADD score of 24.2. We followed up by searching the DECIPHER database of the gene *NAA15*, and note that another nonsynonymous DNV in a patient with global developmental delay is reported there (Firth et al., 2009). In addition, a study has now been published that specifically targeted candidate genes in ID and autism, identifying another 12 likely pathogenic variants in *NAA15* (Stessman et al., 2017). An overview of *NAA15* with the DNVs identified in patients is shown in Figure 1.

To further evaluate the genes where DNVs had been identified in patients also in previous studies, we used the Exome Aggregation Consortium (ExAC) browser (Lek et al., 2016). ExAC reports measures of constraint for all genes and we focused on two relevant metrics reported in ExAC. The measure of Probability of intolerance to loss of function (pLI) score is based on the difference in observed and expected loss of function variants in the gene, with a score >0.9 indicating extremely LoF intolerant genes where heterozygous LoF is not tolerated. The missense z score represents the deviation of observed counts from the expected number of missense variants in a gene, where a positive value means less variants than expected and a negative score means more variants than expected. The pLI and z scores are reported in Table 2. Four of the five pathogenic variants are located in genes that receive a pLI score of 1 (extremely intolerant to heterozygous LoF), while the *PUF60* gene receives a score of 0.85. Of the genes containing VUS, we note that *NAA15* and *TAOK1* both receive a pLI of 1 as well. The z score correlates less well with pathogenicity in our limited sample with scores ranging from negative (*ASXL3*, z = −0.94) to very high (*SMARCA4*, z = 8.36) for the pathogenic variants. Both *NAA15* and *TAOK1* have positive z scores (>3) indicating that they contain less coding variation than expected.

4 | DISCUSSION

This WES study identified clinically significant DNV in 6 out of 28 patients with ID, yielding an 21% diagnostic rate, a relatively low yield compared to previous WES studies in trio families with ID patients, which typically ranged from 20–30% (de Ligt et al., 2012; Miller et al., 2010; Tammimies et al., 2015). Two potential factors could have affected the diagnostic rate. Firstly, the 28 patients were chosen from a larger ID patient cohort without epilepsy. WES results of patients with both ID and epilepsy in the full cohort were previously reported by us with a diagnostic rate of 28% (Halvardson et al., 2016). The ID patients with epilepsy had generally more severe

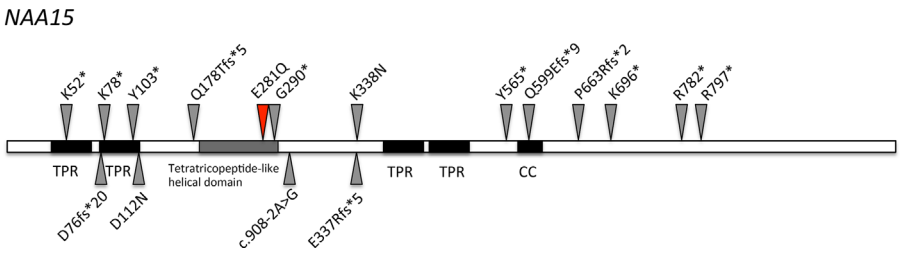

FIGURE 1 Schematic model of NAA15, showing previously reported DNVs in patient cases. The red triangle marks the DNV identified in this study. CC, coiled coil domain; TPR, tetratricopeptide repeat

phenotypes compared to the ID only patients. This is in line with earlier observations that clinical yield from genetic diagnostic screening is higher in patients with more severe phenotype (de Ligt et al., 2012; Gilissen et al., 2014; Rauch et al., 2012; Vissers et al., 2010). Secondly, WES of the patients presented here has been performed sequentially over a long period of time, with different sequencing techniques, sequencing depth, and enrichment kits. Potential causal DNVs could therefore have been missed, especially in the earlier WES of patient trios.

Of the pathogenic variants identified in this study, the de novo two base pair deletion identified in *PUF60* provides further evidence that the heterozygous loss-of-function of this gene causes ID. *PUF60* is one of three genes in the critical region of the 8q24.3 deletion syndrome (MIM#615583), also known as Verheij Syndrome. Patients with deletions in this region present with ID, craniofacial dysmorphology, vertebra anomalies, and short stature. Functional studies of the three genes have proved *SCRIB* and *PUF60* as the main contributors for the phenotype associated with the 8q24.3 deletion syndrome. *PUF60* was first identified as part of a novel complex PUF, which enhanced the efficiency of splicing of several introns (Liu et al., 2000). Depletion of *PUF60* in human cells indicates that *PUF60* is important for neuronal splicing (Hastings, Allemand, Duelli, Myers, & Krainer, 2007). Knockout of the *PUF60* ortholog in *C. elegans* is embryonic lethal, while studies in *Drosophila* implicates a critical role for the *PUF60* ortholog in alternative splicing affecting developmental regulation (MacMorris, Brocker, & Blumenthal, 2003; Park, Parisky, Celotto, Reenan, & Graveley, 2004). To date, there have been 16 reported cases with SNVs in *PUF60* (Figure 2), of which the most common phenotypes include Intellectual disability/developmental delay, short stature, cardiac defects and a recognizable facial phenotype consisting of square face, full cheeks, prominent forehead, low set eyebrows, wide nasal bridge, broad nasal tip, long philtrum, and a thin upper lip (Table 3) (Dauber et al., 2013; Deciphering Developmental Disorders, 2017; El Chehadeh et al., 2016; Santos-Simarro et al., 2017). Moreover, brain and skeletal anomalies also seems to be common among reported cases. These are symptoms that fit well with the phenotype of the patient reported here, a girl with ID, short stature (−3.5 SD), tetrology of fallot, scoliosis, hemivertibrae, and facial features that fits well with the cases described by Dauber and El Chehadeh (Table 3). However, speech delay seems to be unique for our patient. Thus, our data further support the role for *PUF60* as the major contributor to the 8q24.3 deletion phenotype (Dauber et al., 2013).

Of the DNVs we identified that were not scored as pathogenic, we searched for evidence of pathogenicity using results from previous exome studies and by using different available tools to score the variants and genes. We find *NAA15*, encoding the N(alpha)-acetyltransferase 15, NatA auxiliary subunit, to be the most interesting new candidate ID gene. The variant identified in *NAA15* has a high CADD score (24.2), and the gene is under very strong selective constraint based on results in the ExAC browser. Another score used to measure constraint is the Residual Variation Intolerance Score (RVIS) (Petrovski, Wang, Heinzen, Allen, & Goldstein, 2013), and based on its RVIS score *NAA15* is ranked among the top 10% of genes most intolerant to variation.

De novo variants (DNVs) in *NAA15* have been previously reported in three studies of neurodevelopmental disorders. Two DNVs (one nonsense and one nonsynonymous) were previously reported in a study of autism trios (De Rubeis et al., 2014). Unfortunately no additional information on phenotype is available for these patients. One patient was reported in the first publication from the Deciphering Developmental Disorders (DDD) project

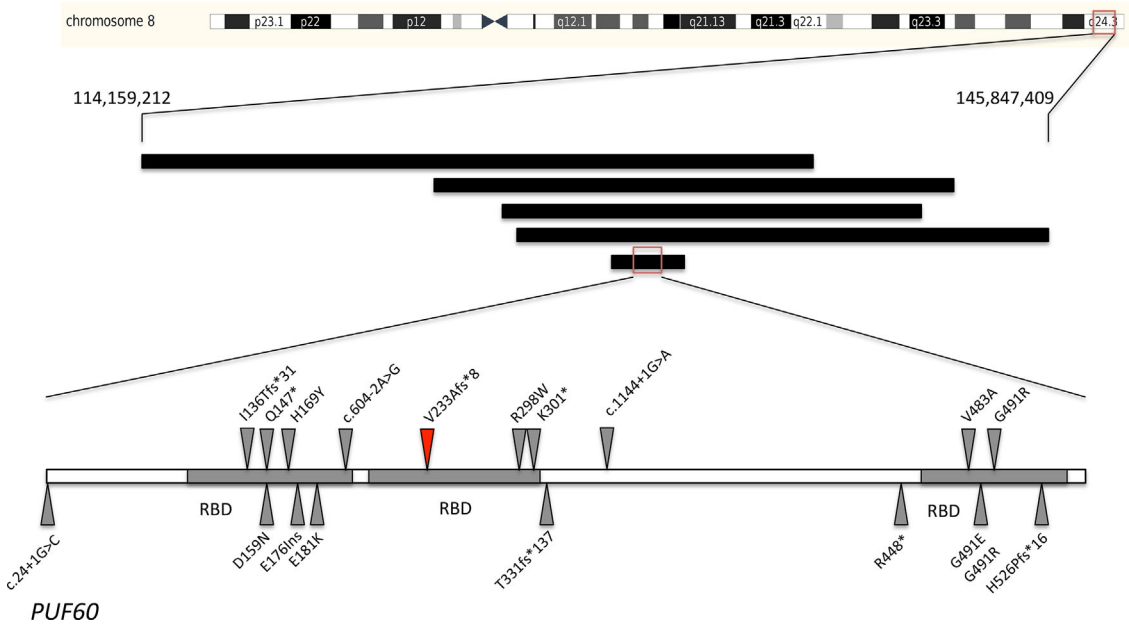

FIGURE 2 Schematic model of chromosome 8 and PUF60, showing five reported deletions and previously reported DNVs. The red triangle marks the 2 bp stop-gain-causing deletion site at amino acid 233 identified in this study. RBD, RNA binding domain

TABLE 3 Phenotype of the patient with PUF60 variant compared with previously reported patients

PUF60	Present study	Dauber et al. (2013)	El Chehadeh et al. (2016) P1	El Chehadeh et al. (2016) P2	El Chehadeh et al. (2016) P3	El Chehadeh et al. (2016) P4	El Chehadeh et al. (2016) P5	Santos-Simarro et al. (2017) P1	Santos-Simarro et al. (2017) P2	Santos-Simarro et al. (2017) P3	Total
Variant detected	c.698_699del, p.Val233Alafs*8	c.505C>T, p.His169Tyr	c.24+1G>C, p.?	c.142C>T, p.Arg448*	c.1448T>C, p.Val483Ala	c.407_410delTCTA, p.Ile136Thrfs*31	c.901A>T, p.Lys301*	c.439C>T, p.Gln147*	c.541G>A, p.Glu181Lys	c.1144+1G>A, p.?	
ID/DD	+	+	+	+	+	+	+	+	+	+	10/10
Speech delay	+	ND	ND	ND	ND	+	ND	ND	+	+	4/4
Brain malformation	Mild hypotrophy of brain	Enlarged 3rd and lateral ventricles, thin corpus callosum	Posterior pituitary ectopia	ND	–	ND	Corpus callosum hypoplasia	White matter subcortical and periventricular abnormalities	ND	ND	5/6
Short stature	+	+	+	–	+	+	+	+	+	+	9/10
Square face	+	+	+	+	–	+	+	ND	ND	ND	6/7
Full cheeks	+	–	+	+	+	+	+	ND	ND	ND	6/7
Bitemporal narrowing	+	ND	+	–	+	+	–	ND	ND	ND	4/6
Prominent forehead	ND	+	+	+	+	+	+	ND	ND	ND	6/6
Low set eyebrows	+	+	+	+	+	+	+	ND	ND	ND	7/7
Wide nasal bridge	–	+	+	+	+	+	+	ND	ND	ND	6/7
Broad nasal tip	+	+	+	+	+	+	+	ND	ND	ND	7/7
Anteverted nares	–	+	+	+	+	+	–	ND	ND	ND	5/7
Long philtrum	+	+	+	+	+	+	+	+	ND	ND	8/8
Thin upper lip	+	+	+	–	+	+	+	+	ND	ND	7/8
Micro- or retrognathism	–	+	+	–	–	+	–	ND	ND	ND	3/7
Short neck	+	+	+	–	–	+	+	ND	ND	ND	5/7
Branchial arch	–	–	Branchial cleft cysts	–	–	Branchial cleft cysts, uvula dimple	–				2/7

(Continues)

TABLE 3 (Continued)

PUF60	Present study	Dauber et al. (2013)	El Chehadeh et al. (2016) P1	El Chehadeh et al. (2016) P2	El Chehadeh et al. (2016) P3	El Chehadeh et al. (2016) P4	El Chehadeh et al. (2016) P5	Santos-Simarro et al. (2017) P1	Santos-Simarro et al. (2017) P2	Santos-Simarro et al. (2017) P3	Total
anomalies											
Cardiac defects	Tetrology of Fallot	VSD, Bicuspid aortic valve	VSD	–	VSD, Truncus arteriosus	Aortic insufficiency	AVSD	–	VSD	TOF	8/9
Abnormality of the eyes	Strabismus, hypertelorism, hyperopia	–	Coloboma, Bilateral microphtalmia and optic nerve hypoplasia	Strabismus	–	ND	Bilateral optic nerve hypoplasia	Bilateral iris coloboma and unilateral left chorioretinal coloboma, strabismus, deep-set eyes with epicanthic folds	ND	Left iridoretinal coloboma	6/9
Abnormality of the hands and feet	Single transverse palmar crease, short and broad fingers and toes	Absent right thumb, hypoplastic metacarpal, scaphoid bones	Clinodactyly dig V, valgus feet deformity	Clinodactyly dig V	–	Valgus feet deformity, brachymesophalangy dig V, overlying proximal thumb	Brachydactyly, pes planus, short hallux, sandal gap	Bilateral talipes	–	–	7/10
Skeletal anomalies	Hemivertebrae, scoliosis	Hip dislocation	Hip dislocation,	ND	Fusion of left facet joints C5-C6	Unfused arch C7, scoliosis	ND	Delayed bone maturation		ND	5/5
Other	Flat occiput, high palate, low set ears with a thick helix, wide and short thorax, broncial obstruction, recurrent epistaxis, mild hearing deficit, sleeping problems, hypermetropi	Bronchopulmonary dysplasia, cleft palate, microcephaly	Nevus flammeus on the forehead and eyelids, pectus excavatum, micromelia		Microcephaly, micromelia	Pectus excavatum, hypoplastic kidneys	Trilobar left lung, Pelvic left kidney with unilateral VUR	Mild facial hypotonia, limited ability to chew and eat solids, broad forehead, short nose with a flat nasal root, simple detached ears. Small mouth with dental crowding	Hypoplasia of depressor anguli oris muscle		

+, presence of phenotype; −, absence of phenotype; AVSD, atrioventricular septal defect; DD, developmental delay; ID, intellectual disability; ND, not determined; VSD, ventricular septal defect; VUR, vesicoureteral reflux.

TABLE 4 Phenotypes of the patient with NAA15 variant compared with previously reported patients

NAA15	Present study	Stessman et al. (2017) P1	Stessman et al. (2017) P2 (rare)	Stessman et al. (2017) P3	Stessman et al. (2017) P4	Stessman et al. (2017) P5 (rare)	Stessman et al. (2017) P6 (rare)	Stessman et al. (2017) P7	Stessman et al. (2017) P8	Stessman et al. (2017) P9	Stessman et al. (2017) P10	Stessman et al. (2017) P11	Stessman et al. (2017) P12	Stessman et al. (2017) P13
Variation detected	c.841G>C, p. Glu281Gln	c.154A>T, p.Lys52*	c.239_240del, p. His80Argfs*17	c.532_533del, p. Gln178Thrfs*5	c.868G>T, p. Gly290*	c.1348A>G, p. Lys450Glu	c.1424C>T, p. Ala475Val	c.1695T>A, p. Tyr565*	c.1988del, p. Pro663Argfs*2	c.2086A>T, p.Lys696*	c.2344C>T, p. Arg782*	c.225_230delTGACTTinsT, p. Asp76Glufs*20	c.334G>A, p. Asp112Asn	c.2389A>T, p. Arg797*
ID/DD	+	+	+		+	+	+	+	+	+	+	ND	–	+
Short stature	+	+	+	ND	ND	–	–	+	ND	ND	ND	ND	ND	–
Behavior problems	–	ND	ND	ASD	Autistic traits, hair-pulling,	Selective mutism	ASD	Aggressive behavior	ASD, hyperactivity	ASD	ND	ASD	Asperger	NOS Pervasive Developmental Disorder, motor instability, opposition
Motor delay	+	ND	+	ND	–	–	ND	ND	ND	ND	ND	ND	ND	+
Speech delay	+	ND	+	ND	+	–	+	+	ND	ND	ND	ND	ND	+
Other	SGA, thick helix, widely spaced nipples. No facial dysmorphology, slight joint hypermobility, speech dyspraxia. Speech is her most prominent problem.	Cutis marmorata, small head	Neonatal cyanosis, microcephaly, facial asymmetry, hypertelorism, prominent anthelix on the ears, a single café-au-lait spot on the left side of the trunk, joint hyperlaxity, genu recurvatum, right congenital hip dysplasia, hyperopia	ND	ND	Hypertelorism, broad nasal bridge	ND	Unilateral cryptorchidia, ptosis of the left eye, mild hypertelorism, flat philtrum, mild pectus carinatum upper side and excavatum below, low set ears	ND	3 pre-auricular tags on the right side	ND	ND	ND	ND

Rare: Variants in Patient P2, P5, and P6 are not private, but considered ultrarare (allele count ≤3), and are found in more then one family in Stessman et al. (2017).
+, presence of phenotype; −, absence of phenotype; ASD, autism spectrum disorder; DD, developmental delay; ID, intellectual disability; ND, not determined; SGA, small for gestational age.

(Deciphering Developmental Disorders, 2015). This patient has global developmental delay, abnormal facial gestalt and obesity, and carries a 4 bp deletion that results in a frameshift (p.Glu337Argfs*5). In a recent update of the DECIPHER database (Firth et al., 2009), another de novo variant (nonsense, p.Ser767*) was reported in a patient with global developmental delay, stereotypic behavior, and eczema. Recently, *NAA15* was identified as a neurodevelopmental disorder (NDD) risk gene in a study that combined single-molecule molecular inversion probles (smMIPs) with exome sequencing. Variants in *NAA15* were discovered in 13 patients, of which four variants were confirmed to be de novo. The patients shared several key phenotypes including ID (91%), speech delay (83%), Autism Spectrum Disorder (ASD), and diagnosis (63%) (Stessman et al., 2017). These symptoms fit well with the patient reported here who was small for gestational age, with mild ID, motor, and speech delay including speech dyspraxia (Table 4). Most of the reported *NAA15* variants are nonsense variants, and only three *NAA15* missense variants have been reported. However, the phenotypes of the three patients range from Asperger syndrome, DD, to severe ID, and do not seem to differ from the patients with nonsense mutations.

NAA15 is regulated by the N-methyl-D-aspartate class of glutamate receptors (NMDAR), which is responsible for the transmission of signals between neurons (Fluge, Bruland, Akslen, Varhaug, & Lillehaug, 2002; Sugiura, Patel, & Corriveau, 2001). Electrical activity occurs early in developing neurons, which regulates neuronal differentiation and migration, before a full synaptic network is established. *NAA15* was first discovered in NMDAR knockout mice, where it was shown to be regulated by physiological levels of NMDAR function in developing neurons *in vivo*, and was termed mNAT1 (Sugiura et al., 2001). mNAT1 is expressed at high levels in the neonatal brain in regions of neuronal proliferation and migration, and is dramatically down-regulated during early postnatal development (Sugiura et al., 2001). The mouse homolog of *NAA15*, mNAT1, was also shown by in situ hybridization to be highly expressed in areas of cell division and migration and are down-regulated as neurons differentiate, suggesting an important role in the generation and differentiation of neurons (Sugiura, Adams, & Corriveau, 2003). Knockdown of Nat1 (*NAA15* ortholog) in drosophila caused adult early lethality, erect wings, and impaired locomotor activity (Stessman et al., 2017). Another knockdown of Nat1 in drosophila was performed with a presumably weaker RNAi line, which led to normal locomotion and morphology, but impaired response in the light-off jump paradigm, as well as habituation deficits (Stessman et al., 2017). It has been shown that mNAT1 interacts with the protein mARD1 and forms an acetyltransferase complex (Arnesen et al., 2005). This is noteworthy as the human homolog of mARD1 is *NAA10*, which together with *NAA15* forms the major Nα-terminal acetyltransferase complex (NatA) (Sanchez-Puig & Fersht, 2006). The *NAA15* missense variant identified in this study is located in the tetraticopeptide-like helical domain, which is a structural motif that mediates protein-protein interactions and the assembly of multiprotein complexes (Figure 1). This variant could potentially affect the binding between *NAA10* and *NAA15*. *NAA10* is associated with Ogden syndrome, an X-linked neurodevelopment disorder (MIM#300013). Ogden syndrome can be caused by both X-linked dominant inheritance with a milder phenotype, or X-linked recessive inheritance with severe phenotypes, and the syndrome is characterized by postnatal growth failure, delayed psychomotor development, ID, hypotonia, and dysmorphic features (Casey et al., 2015; Popp et al., 2015; Rope et al., 2011). The phenotype thus shows some shared features with the phenotype reported for patients with *NAA15* DNVs.

In summary, our exome sequencing study revealed a limited number of pathogenic variants, but we highlight two variants that we believe contribute new biologically relevant information. We report a small deletion in *PUF60* that gives rise to a phenotype similar to Verheij Syndrome, providing further evidence that *PUF60* is the major causative gene within the previously reported microdeletion. We also provide multiple lines of evidence suggesting that variants in *NAA15* are associated with an ID syndrome. Additional patients and more detailed patient phenotype comparisons will now be required to further refine the role of *PUF60* and *NAA15* in intellectual disability, and better define the core symptoms associated with causative variants in these genes.

ACKNOWLEDGMENTS

We are very grateful to the participating families for their cooperation. We thank Professor Göran Annerén for initially seeing some of the patients and sharing his knowledge. This study makes use of data generated by the DECIPHER community. A full list of centers which contributed to the generation of the data is available from http://decipher.sanger.ac.uk and via email from decipher@sanger.ac.uk. Funding for the project was provided by the Wellcome Trust. The authors would like to thank the Exome Aggregation Consortium and the groups that provided exome variant data for comparison. A full list of contributing groups can be found at http://exac.broadinstitute.org/about. Computational analyses were performed on resources provided by SNIC through Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX). This work was supported by grants from the Sävstaholm Society and the "Regionala forskningsrådet" (to ACT), and the European Research Council ERC Starting Grant Agreement n. 282330 and the Swedish Medical Research Council (to LF).

CONFLICTS OF INTEREST

There is no conflict of interest to declare.

ORCID

Jin J. Zhao http://orcid.org/0000-0001-8367-8391

REFERENCES

American Association on Intellectual and Developmental Disabilities. (2010). American Association on Intellectual and Developmental Disabilities' definition manual, *Intellectual Disability: Definition, Classification, and Systems of Supports* (Eleventh edition).

Arnesen, T., Anderson, D., Baldersheim, C., Lanotte, M., Varhaug, J. E., & Lillehaug, J. R. (2005). Identification and characterization of the human ARD1-NATH protein acetyltransferase complex. *Biochemical Journal*, *386*, 433–443.

Casey, J. P., Stove, S. I., McGorrian, C., Galvin, J., Blenski, M., Dunne, A., … Lynch, S. A. (2015). NAA10 mutation causing a novel intellectual disability syndrome with Long QT due to N-terminal acetyltransferase impairment. *Scientific Reports*, *5*, 16022.

Dauber, A., Golzio, C., Guenot, C., Jodelka, F. M., Kibaek, M., Kjaergaard, S., … Katsanis, N. (2013). SCRIB and PUF60 are primary drivers of the multisystemic phenotypes of the 8q24.3 copy-number variant. *American Journal of Human Genetics*, *93*(5), 798–811.

de Ligt, J., Willemsen, M. H., van Bon, B. W., Kleefstra, T., Yntema, H. G., Kroes, T., … Vissers, L. E. (2012). Diagnostic exome sequencing in persons with severe intellectual disability. *New England Journal of Medicine*, *367*(20), 1921–1929.

De Rubeis, S., He, X., Goldberg, A. P., Poultney, C. S., Samocha, K., Cicek, A. E., … Buxbaum J. D. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*, *515*(7526), 209–215.

de Vries, B. B., Pfundt, R., Leisink, M., Koolen, D. A., Vissers, L. E., Janssen, I. M., … Veltman, J. A. (2005). Diagnostic genome profiling in mental retardation. *American Journal of Human Genetics*, *77*(4), 606–616.

Deciphering Developmental Disorders S. (2015). Large-scale discovery of novel genetic causes of developmental disorders. *Nature*, *519*(7542), 223–228.

Deciphering Developmental Disorders S. (2017). Prevalence and architecture of de novo mutations in developmental disorders. *Nature*, *542*(7642), 433–438.

El Chehadeh, S., Kerstjens-Frederikse, W. S., Thevenon, J., Kuentz, P., Bruel, A. L., Thauvin-Robinet, C., … Faivre, L. (2016). Dominant variants in the splicing factor PUF60 cause a recognizable syndrome with intellectual disability, heart defects and short stature. *European Journal of Human Genetics*, *25*(1), 43–51.

Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., … Carter, N. P. (2009). DECIPHER: Database of chromosomal imbalance and phenotype in humans using ensembl resources. *American Journal of Human Genetics*, *84*(4), 524–533.

Fluge, O., Bruland, O., Akslen, L. A., Varhaug, J. E., & Lillehaug, J. R. (2002). NATH, a novel gene overexpressed in papillary thyroid carcinomas. *Oncogene*, *21*(33), 5056–5068.

Gilissen, C., Hehir-Kwa, J. Y., Thung, D. T., van de Vorst, M., van Bon, B. W., Willemsen, M. H., … Veltman, J. A. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature*, *511*(7509), 344–347.

Halvardson, J., Zhao, J. J., Zaghlool, A., Wentzel, C., Georgii-Hemming, P., Mansson, E., … Feuk, L. (2016). Mutations in HECW2 are associated with intellectual disability and epilepsy. *Journal of Medical Genetics*, *53*(10), 697–704.

Hastings, M. L., Allemand, E., Duelli, D. M., Myers, M. P., & Krainer, A. R. (2007). Control of pre-mRNA splicing by the general splicing factors PUF60 and U2AF(65). *PLoS ONE*, *2*(6), e538.

Kircher, M., Witten, D. M., Jain, P., O'Roak, B. J., Cooper, G. M., & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics*, *46*(3), 310–315.

Knight, S. J., Regan, R., Nicod, A., Horsley, S. W., Kearney, L., Homfray, T., … Flint, J. (1999). Subtle chromosomal rearrangements in children with unexplained mental retardation. *The Lancet*, *354*(9191), 1676–1681.

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., … Exome Aggregation Consortium. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*(7616), 285–291.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760.

Liu, J. H., He, L. S., Collins, I., Ge, H., Libutti, D., Li, J. F., … Levens, D. (2000). The FBP interacting repressor targets TFIIH to inhibit activated transcription. *Molecular Cell*, *5*(2), 331–341.

MacMorris, M., Brocker, C., & Blumenthal, T. (2003). UAP56 levels affect viability and mRNA export in Caenorhabditis elegans. *RNA*, *9*(7), 847–857.

Miller, D. T., Adam, M. P., Aradhya, S., Biesecker, L. G., Brothman, A. R., Carter, N. P., … Ledbetter, D. H. (2010). Consensus statement: Chromosomal microarray is a first-Tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *American Journal of Human Genetics*, *86*(5), 749–764.

Park, J. W., Parisky, K., Celotto, A. M., Reenan, R. A., & Graveley, B. R. (2004). Identification of alternative splicing regulators by RNA interference in Drosophila. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(45), 15974–15979.

Petrovski, S., Wang, Q., Heinzen, E. L., Allen, A. S., & Goldstein, D. B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genetics*, *9*(8), e1003709.

Popp, B., Stove, S. I., Endele, S., Myklebust, L. M., Hoyer, J., Sticht, H., … Reis, A. (2015). De novo missense mutations in the NAA10 gene cause severe non-syndromic developmental delay in males and females. *European Journal of Human Genetics*, *23*(5), 602–609.

Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., … Strom, T. M. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: An exome sequencing study. *The Lancet*, *380*(9854), 1674–1682.

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., … Acmg Laboratory Quality Assurance Committee. (2015). 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. *Genetics in Medicine*, *17*(5), 405–424.

Rope, A. F., Wang, K., Evjenth, R., Xing, J., Johnston, J. J., Swensen, J. J., … Lyon, G. J. (2011). Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency. *American Journal of Human Genetics*, *89*(1), 28–43.

Ropers, H. H. (2010). Genetics of early onset cognitive impairment. *Annual Review of Genomics and Human Genetics*, *11*, 161–187.

Sanchez-Puig, N., & Fersht, A. R. (2006). Characterization of the native and fibrillar conformation of the human Nalpha-acetyltransferase ARD1. *Protein Science*, *15*(8), 1968–1976.

Santos-Simarro, F., Vallespin, E., Del Pozo, A., Ibanez, K., Silla, J. C., Fernandez, L., … Palomares-Bralo, M. (2017). Eye coloboma and complex cardiac malformations belong to the clinical spectrum of PUF60 variants. *Clinical Genetics*. [Epub ahead of print].

Shyr, C., Tarailo-Graovac, M., Gottlieb, M., Lee, J. J., van Karnebeek, C., & Wasserman, W. W. (2014). FLAGS, frequently mutated genes in public exomes. *BMC Medical Genomics*, *7*, 64.

Stessman, H. A., Xiong, B., Coe, B. P., Wang, T., Hoekzema, K., Fenckova, M., … Eichler, E. E. (2017). Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nature Genetics*, *49*(4), 515–526.

Sugiura, N., Adams, S. M., & Corriveau, R. A. (2003). An evolutionarily conserved N-terminal acetyltransferase complex associated with neuronal development. *Journal of Biological Chemistry*, *278*(41), 40113–40120.

Sugiura, N., Patel, R. G., & Corriveau, R. A. (2001). N-methyl-D-aspartate receptors regulate a group of transiently expressed genes in the developing brain. *Journal of Biological Chemistry*, *276*(17), 14257–14263.

Tammimies, K., Marshall, C. R., Walker, S., Kaur, G., Thiruvahindrapuram, B., Lionel, A. C., … Fernandez, B. A. (2015). Molecular diagnostic yield of chromosomal microarray analysis and whole-Exome sequencing in children with autism spectrum disorder. *Jama-Journal of the American Medical Association*, *314*(9), 895–903.

Vissers, L. E., de Ligt, J., Gilissen, C., Janssen, I., Steehouwer, M., de Vries, P., … Veltman, J. A. (2010). A de novo paradigm for mental retardation. *Nature Genetics*, *42*(12), 1109–1112.

Vissers, L. E., Gilissen, C., & Veltman, J. A. (2016). Genetic studies in intellectual disability and related disorders. *Nature Reviews. Genetics*, *17*(1), 9–18.

World Health Organization, Reginal office for Europe. (2010). Better health, better lives: children and young people with intellectual disability and their families.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Zhao JJ, Halvardson J, Zander CS, et al. Exome sequencing reveals *NAA15* and *PUF60* as candidate genes associated with intellectual disability. *Am J Med Genet Part B*. 2018;177B:10–20. https://doi.org/10.1002/ajmg.b.32574