Truncating Variants in *NAA15* Are Associated with Variable Levels of Intellectual Disability, Autism Spectrum Disorder, and Congenital Anomalies

Hanyin Cheng,[1,49] Avinash V. Dharmadhikari,[1,49] Sylvia Varland,[2,45] Ning Ma,[3,4,5] Deepti Domingo,[6] Robert Kleyner,[7] Alan F. Rope,[8] Margaret Yoon,[7] Asbjørg Stray-Pedersen,[9,10] Jennifer E. Posey,[9] Sarah R. Crews,[11] Mohammad K. Eldomery,[9] Zeynep Coban Akdemir,[9] Andrea M. Lewis,[9,12] Vernon R. Sutton,[9] Jill A. Rosenfeld,[9] Erin Conboy,[13] Katherine Agre,[13] Fan Xia,[1,9] Magdalena Walkiewicz,[1,9,14] Mauro Longoni,[15,16] Frances A. High,[15,17,18] Marjon A. van Slegtenhorst,[19] Grazia M.S. Mancini,[19] Candice R. Finnila,[20] Arie van Haeringen,[21] Nicolette den Hollander,[21] Claudia Ruivenkamp,[21] Sakkubai Naidu,[22] Sonal Mahida,[22] Elizabeth E. Palmer,[23,24]

(Author list continued on next page)

N-alpha-acetylation is a common co-translational protein modification that is essential for normal cell function in humans. We previously identified the genetic basis of an X-linked infantile lethal Mendelian disorder involving a c.109T>C (p.Ser37Pro) missense variant in *NAA10*, which encodes the catalytic subunit of the N-terminal acetyltransferase A (NatA) complex. The auxiliary subunit of the NatA complex, NAA15, is the dimeric binding partner for NAA10. Through a genotype-first approach with whole-exome or genome sequencing (WES/WGS) and targeted sequencing analysis, we identified and phenotypically characterized 38 individuals from 33 unrelated families with 25 different *de novo* or inherited, dominantly acting likely gene disrupting (LGD) variants in *NAA15*. Clinical features of affected individuals with LGD variants in *NAA15* include variable levels of intellectual disability, delayed speech and motor milestones, and autism spectrum disorder. Additionally, mild craniofacial dysmorphology, congenital cardiac anomalies, and seizures are present in some subjects. RNA analysis in cell lines from two individuals showed degradation of the transcripts with LGD variants, probably as a result of nonsense-mediated decay. Functional assays in yeast confirmed a deleterious effect for two of the LGD variants in *NAA15*. Further supporting a mechanism of haploinsufficiency, individuals with copy-number variant (CNV) deletions involving *NAA15* and surrounding genes can present with mild intellectual disability, mild dysmorphic features, motor delays, and decreased growth. We propose that defects in NatA-mediated N-terminal acetylation (NTA) lead to variable levels of neurodevelopmental disorders in humans, supporting the importance of the NatA complex in normal human development.

Advances in sequencing technologies such as whole-exome or genome sequencing (WES/WGS) have led to disease-gene association discoveries, functional annotation of the human genome, and improved diagnostic rates in individuals with suspected genetic disorders refractory to conventional diagnostic testing. An estimated diagnostic rate that often exceeds 25% can be achieved when WES/WGS is applied to otherwise undiagnosed complex cases.[1–5] *NAA15* (N-alpha-acetyltransferase 15, MIM: 608000) was previously characterized as one of fifty-two risk genes for neurodevelopmental disorders by targeted sequencing of a large autism spectrum and intellectual disability (ASID) cohort.[6] In another study of *de novo* changes in severe congenital heart disease (CHD), likely gene disrupting (LGD) variants in *NAA15* were identified in two affected individuals in a cohort of 362 severe CHD cases; one of these individuals was known to have additional neurodevelopmental defects.[7] In an effort to further characterize the clinical and molecular spectrum associated with genetic defects in *NAA15*, we ascertained, from 33 unrelated families, 38 individuals with truncating, presumably LGD (nonsense, frameshifting and splice) variants in *NAA15* via a collaborative world-wide effort among multiple institutions. As a result of comprehensive clinical evaluation and molecular analyses in all individuals, we propose that deleterious variants in *NAA15* are

[1]Baylor Genetics, Houston, TX, 77021, USA; [2]Department of Biomedicine, University of Bergen, N-5020 Bergen, Norway; [3]Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA 94305, USA; [4]Division of Cardiology, Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA; [5]Institute for Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA; [6]School of Biological Sciences, Faculty of Genes and Evolution, the University of Adelaide, Adelaide, SA 5000, Australia; [7]Stanley Institute for Cognitive Genomics, 1Bungtown Road, Cold Spring Harbor Laboratory, NY 11724, USA; [8]Department of Medical Genetics, Kaiser Permanente Northwest, Portland, OR 97227, USA; [9]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA; [10]Norwegian National Unit for Newborn Screening, Division of Pediatric and Adolescent Medicine, Oslo University Hospital, N-0424 Oslo, and Institute of Clinical Medicine, University of Oslo, N-0318 Oslo, Norway; [11]Department of Pharmacology, Creighton University Medical School, Omaha, NE, 68178, USA; [12]Department of Pediatrics, Texas Children's Hospital and Baylor College of Medicine, Houston, TX 77030, USA; [13]Department of Clinical Genomics, Mayo Clinic, MN 55905, USA; [14]The National Institute of Allergy and Infectious Disease, The National Institutes of Health, Bethesda, MD 20892, USA; [15]Pediatric Surgical Research Laboratories, Massachusetts General Hospital, Boston, MA 02114, USA; [16]Department of Surgery, Harvard Medical School,

(Affiliations continued on next page)

Lucinda Murray,[23] Derek Lim,[25] Parul Jayakar,[26] Michael J. Parker,[27] Stefania Giusto,[28] Emanuela Stracuzzi,[28] Corrado Romano,[28] Jennifer S. Beighley,[29] Raphael A. Bernier,[29] Sébastien Küry,[30] Mathilde Nizon,[30] Mark A. Corbett,[31] Marie Shaw,[31] Alison Gardner,[31] Christopher Barnett,[32] Ruth Armstrong,[33] Karin S. Kassahn,[34,35] Anke Van Dijck,[36] Geert Vandeweyer,[36] Tjitske Kleefstra,[37] Jolanda Schieving,[37] Marjolijn J. Jongmans,[37] Bert B.A. de Vries,[37] Rolph Pfundt,[37] Bronwyn Kerr,[38,39] Samantha K. Rojas,[40] Kym M. Boycott,[40] Richard Person,[41] Rebecca Willaert,[41] Evan E. Eichler,[42,43] R. Frank Kooy,[36] Yaping Yang,[1,9] Joseph C. Wu,[3,4,5] James R. Lupski,[9,44] Thomas Arnesen,[2,45,46] Gregory M. Cooper,[20] Wendy K. Chung,[47] Jozef Gecz,[6,31,48] Holly A.F. Stessman,[11,50] Linyan Meng,[1,9,50,*] and Gholson J. Lyon[7,50,*]

associated with variable levels of intellectual disability, developmental delay, autism spectrum disorder, dysmorphic features, and congenital cardiac anomalies.

This study was performed in accordance with protocols approved by the institutional review boards of the participating institutions (see Supplemental Data). Three affected individuals were recruited from the UK Deciphering Developmental Disorders (DDD) project (families 20, 27, and 30). Written informed consent was obtained from all study participants. The key clinical features of our cohorts are summarized in Table 1. Detailed clinical summaries for each subject are provided in the Supplemental Data. The use of GeneMatcher, a web-based tool for connecting researchers with an interest in the same gene,[8] facilitated contact between international collaborators.

All subjects have variable degrees of neurodevelopmental disabilities, including impaired motor abilities (HP: 0001270), intellectual disability (HP: 0001249), impaired verbal abilities (HP: 0000750), and autism spectrum disorder (HP: 0000729) (Table 1, Table S1, and Table S2). Many subjects have impaired motor function, including fine-motor difficulties (n = 5, or 12%), mild ataxia (n = 1), abnormality of movement (n = 1), motor delay (n = 22, or 60%), and hypotonia (n = 5, or 14%). Various levels of intellectual disability are reported in almost all study subjects with available data; such disability includes mild, moderate, or severe intellectual disability and learning difficulties with or without behavioral issues (Table 1, Table S1, and Table S2). The majority of affected individuals have verbal issues, including complete absence of speech, delayed language development, the need for sign language, or other speech difficulties. Most subjects also present with autism spectrum disorder (ASD) and/or other behavioral abnormalities. Individual 11 was noted to have marked hypersomnolence in early years, in apparent similarity to what was recently reported in a girl with a missense variant in *NAA10*.[9] Minor facial dysmorphology was reported in some individuals (Table 1), but there were no consistent features noted nor a recognizable pattern of facial dysmorphology (Figure 1 and Figure S1). The birth weight and length were low in some individuals; the most notable feature was a birth weight $\leq$ 1st percentile in 7 out of 25 (28%) individuals with available information (Table S3). Some of the individuals remain small throughout life, whereas others are of normal stature and a few are above average height (Table S3).

Almost all individuals have normal or uncharacterized cardiac function (Table 1), with four exceptions. Individual 2 (Figure 1) has atrial ectopic (multifocal) tachycardia (HP: 0011701), treated with verapamil, and hypertension (HP: 0000822). Individual 3 had a ventricular septal defect (VSD), repaired surgically during infancy. Individual 17

Boston, MA 02114, USA; [17]Department of Pediatrics, Massachusetts General Hospital, Boston, MA 02114, USA; [18]Department of Surgery, Boston Children's Hospital, Boston, MA 02115, USA; [19]Department of Clinical Genetics, Erasmus University Medical Center, 3015 CN Rotterdam, The Netherlands; [20]HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806, USA; [21]Department of Clinical Genetics, Leiden University Medical Center, Leiden, 2333, The Netherlands; [22]Kennedy Krieger Institute, 801 North Broadway Baltimore, MD 21205, USA; [23]Genetics of Learning Disability Service, Hunter Genetics, Waratah, NSW 2298, Australia; [24]School of Women's and Children's Health, University of New South Wales, Sydney, NSW 2031, Australia; [25]West Midlands Regional Genetics Service, Birmingham Women's and Children's NHS Foundation Trust, Mindelsohn Way, Birmingham B15 2TG, UK; [26]Division of Genetics and Metabolism, Nicklaus Children's Hospital, Miami, FL 33155, USA; [27]Sheffield Clinical Genetics Service, Sheffield Children's Hospital, Western Bank, Sheffield S10 2TH, UK; [28]Oasi Research Institute - Istituto di Ricovero e Cura a Carattere Scientifico, Troina 94018, Italy; [29]Department of Psychiatry, University of Washington, Seattle WA, 98195, USA; [30]Department of Medical Genetics, Centre Hospitalier Universitaire, Nantes 44093, France; [31]Adelaide Medical School and Robinson Research Institute, the University of Adelaide, Adelaide, SA 5000, Australia; [32]Paediatric and Reproductive Genetics, South Australian Clinical Genetics Service, SA Pathology (at Women's and Children's Hospital), Adelaide, SA 5006, Australia; [33]East Anglian Medical Genetics Service, Clinical Genetics, Addenbrooke's Treatment Centre, Addenbrooke's Hospital, Cambridge CB2 0QQ, UK; [34]Department of Genetics and Molecular Pathology, SA Pathology, Women's and Children's Hospital, North Adelaide, SA 5006, Australia; [35]School of Biological Sciences, University of Adelaide, Adelaide, SA 5000, Australia; [36]Department of Medical Genetics, University of Antwerp, Antwerp 2000, Belgium; [37]Department of Human Genetics, Radboud University Medical Center, Nijmegen 6500HB, The Netherlands; [38]Manchester Centre for Genomic Medicine, St Mary's Hospital, Manchester University NHS Foundation Trust, Manchester Academic Health Sciences Centre, Manchester M13 9PL, UK; [39]Division of Evolution and Genomic Sciences School of Biological Sciences, University of Manchester, Manchester M13 9PL, UK; [40]Children's Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, ON K1H 8L1, Canada; [41]GeneDx, 207 Perry Parkway, Gaithersburg, MD 20877, USA; [42]Department of Genome Sciences, University of Washington School of Medicine, Seattle, WA 98195, USA; [43]Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA; [44]Human Genome Sequencing Center of Baylor College of Medicine, Houston, TX 77030, USA; [45]Department of Surgery, Haukeland University Hospital, N-5021 Bergen, Norway; [46]Department of Molecular Biology, University of Bergen, N-5020 Bergen, Norway; [47]Departments of Pediatrics and Medicine, Columbia University Medical Center, New York, NY 10032, USA; [48]Healthy Mothers, Babies and Children, South Australian Health and Medical Research Institute, Adelaide, SA 5000, Australia
[49]These authors contributed equally
[50]These authors contributed equally
*Correspondence: lmeng@bcm.edu (L.M.), gholsonjlyon@gmail.com (G.J.L.)
https://doi.org/10.1016/j.ajhg.2018.03.004.

Table 1. Summary of Phenotypes

Phenotype	Number of individuals with phenotype	Number of individuals with relevant data	Percentage
Brain Structure and Function			
Intellectual disability (ID)[a]	23	23	100
ASD, ADHD, or behavioral issues	30	33	91
Abnormal brain MRI	2	11	18
Speech delay	32	33	97
Seizures	6	26	23
Motor Impairments			
Motor delay and related abnormalities	31	32	97
Muscle tone issues	7	18	39
Feeding difficulties	8	14	57
Cardiovascular			
Congenital cardiac defects	4	19	21
Major vessel anomalies	2	19	11
Arrhythmias	1	19	5
Hypertension	1	19	5
Other			
Mild dysmorphism	18	28	64
Skeletal or connective-tissue defects	8	20	40

[a]In individuals > 5 years of age, when IQ testing or other cognitive testing was performed.

has a Marfanoid habitus, with an aortic root at the upper limit of normal. By far the most severely affected, individual 19 has heterotaxy syndrome associated with a complex cardiac diagnosis of dextrocardia involving left superior and inferior venae cavae, total anomalous pulmonary venous return to the innominate vein, tricuspid atresia, hypoplastic right ventricle, double-outlet right ventricle, and transposed great arteries with severe pulmonary stenosis. The variant in this individual (c.1009_1012delGAAA) was previously reported in a cohort of 1,213 subjects with CHD and an increased prevalence of extracardiac congenital anomalies (CAs) and risk of neurodevelopmental disabilities (NDDs).[7,10] Another LGD variant, c.2282C>A (p.Ser761*) in *NAA15*, was first reported in an individual with pulmonary stenosis, single left coronary artery, and tetralogy of Fallot (in the context of no reported neurodevelopmental disability), although we have been unable to obtain additional information on this individual.[7] A more recent analysis of this now expanded cohort of 2,871 CHD probands, including 2,645 parent-offspring trios, did not find any new variants in *NAA15*.[11] Given the low prevalence of CHD in our own cohort of 38 individuals, one caveat is that the expression of severe congenital heart disease could be due to variation at a second locus, a noncoding mutation outside of the exome, or some other additional variation undetected thus far.

A total of 25 presumably LGD variants contained in 12 of the 20 exons and two intron-exon boundaries of *NAA15* were identified from 33 unrelated families (Figure 2, Table S1, Table S2); these included nonsense variants (n = 13), canonical splice-site variants (n = 2), and frameshift variants (n = 10). The inheritance pattern of the variants was determined to be *de novo* for most subjects (22 families) through testing of parental samples. Familial inheritance was observed in three families (families 10, 22, and 28), and the corresponding *NAA15* LGD variant segregated with the neurocognitive phenotypes, including in one mildly affected parent in each family and in affected siblings in families 10 and 28. For Family 10, the read count data did not demonstrate any somatic mosaicism in the blood sample from the mother. Among the 25 variants identified, there were three recurrent variants, including c.228_232delCTTGA (p.Asp76Glufs*20) (families 3 and 4, *de novo*), c.239_240delAT (p.His80Argfs*17) (families 6-11B, *de novo* in families 6 and 7, familial in family 10, and unknown inheritance in the rest of the families), and c.1009_1012delGAAA (p.Glu337Argfs*5) (families 19 and 20, *de novo*). We examined genomic context around the three recurrent loci to look for micro-homology that might increase the propensity for recurrent mutations and found that the most recurrent mutation, c.239_240delAT, occurs in the middle of one of 20 reported mutation hotspots, CATGT.[12] In addition, this recurrent variant, and another one, c.228_232delCTTGA, are close to each other in exon 3 in an area that is computationally predicted[13] to form a quasipalindromic structure (lying distal to an even larger quasipalindromic structure), and the third recurrent mutation c.1009_1012delGAAA in exon 9 lies just distal to a quasipalindromic structure (Figure S2).[14]

Data from the Exome Aggregation Consortium (ExAC) study of 60,706 control individuals show that *NAA15* is likely intolerant to LGD variants (pLI = 1.00),[15] that the residual variation intolerance score (RVIS) = −0.89 (among the 10.2% most LGD intolerant of human genes), and that LoF-FDR[ExAC] = 0.000224349.[16] Excluding small cohorts (<100 probands, Table S4), we in total identified fourteen *de novo* variants in *NAA15* from six independent rare disease cohorts with a total sample size of ~36,731. Ten out of 14 cases are reported in detail here; the remaining four lack sufficient phenotype information. Our aggregate frequency of *de novo* LGD variants in affected individuals (~4.0 per 10,000) is significantly higher than the background rates estimated by Samocha et al.[17] for LGD

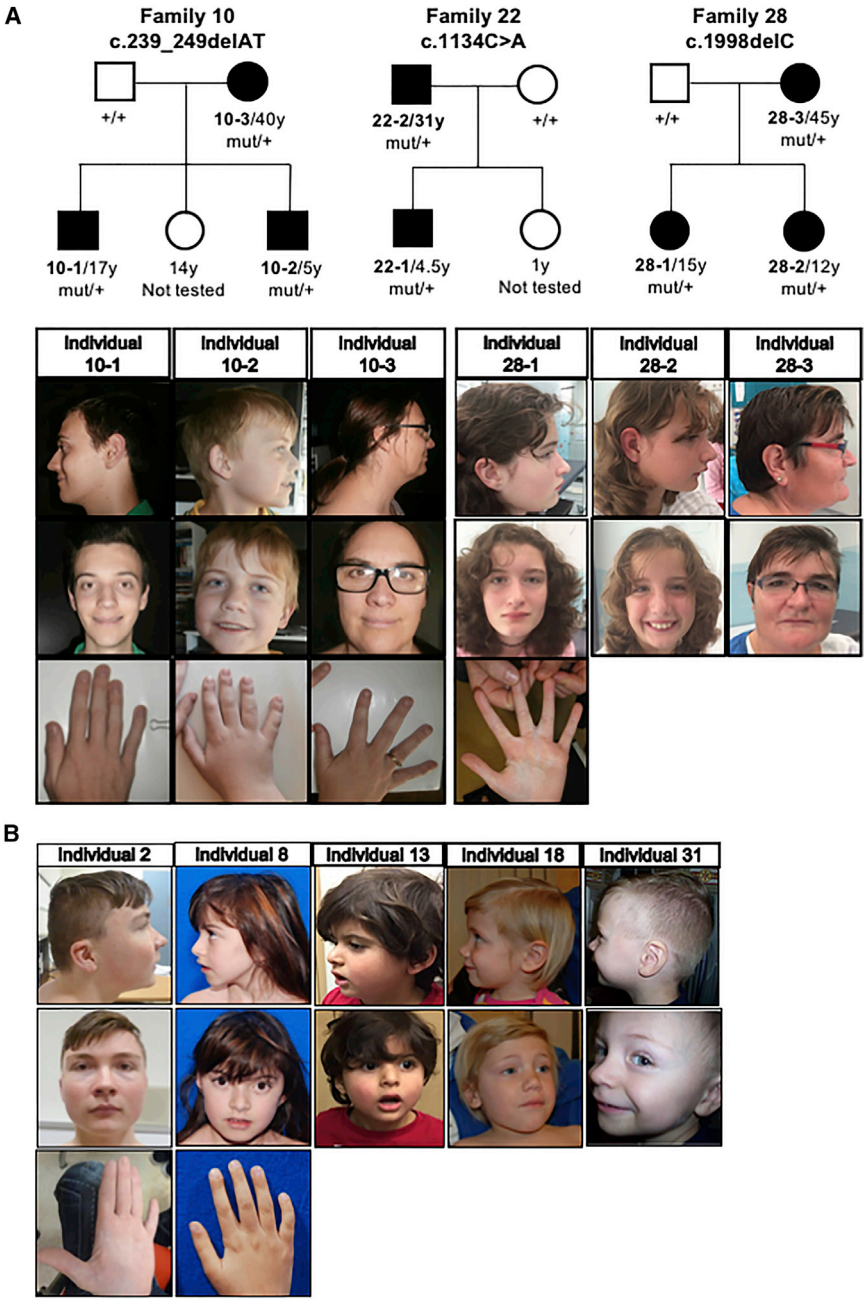


Figure 1. Pedigrees, Mild Facial Dysmorphology, and Hands of Individuals with Familial or *de novo NAA15* LGD Variants

(A) Pedigrees are shown for the three families with inherited variants. **Family 10, Individual 10-1:** at age 17 years and 6 months, with prominent eyebrows, broad nose, and prominent chin. Hand appears normal. **Individual 10-2:** at 6 years and 6 months, with very well-developed philtral pillars. Hand appears normal. **Individual 10-3:** mother, with long mentum of the chin and relatively thick alae nasi. Hand appears normal. **Family 28, Individual 28-1:** at age 15 years, partial syndactyly in one hand, but otherwise not with particularly notable dysmorphology. **Individual 28-2:** sister, at age 12 years, who was not noted to have any obvious dysmorphology. **Individual 28-3:** Mother at age 45 years, with broad nose but otherwise not with notable dysmorphology.

(B) Minor facial dysmorphology was noted in some probands, but there were no reliably consistent features shared among them. **Individual 2:** at 17 years old, noted to have brachycephaly, appearance of ocular hypertelorism with short palpebral fissures, prominent nose tip with a longer columella of the nose, trapezoidal philtrum, and micrognathia without retrognathia. Also noted are small low-set, posteriorly rotated ears, with thickened and overfolded helix; hypoplastic distal phalanges on digits 2, 3, and 4; 5th finger with brachyclinodactyly; and persistence of fetal finger pads on the 3rd and 4th digit. **Individual 8:** at the age 8 years 9 months, noted to have thin philtrum, bulbous nasal tip, and 5th finger with brachyclinodactyly. **Individual 13:** at 4 years old, no facial dysmorphism noted. **Individual 18:** at 4 years and 3 months, with bulbous nose tip, thick alae nasi and anteverted nares, prominent cupid's bow and philtrum, long mentum of the chin, and simple ears. **Individual 31**: with epicanthus inversus, smooth philtrum, thin vermilion border of the upper lip, and sparse lateral eyebrows.

mutations (expected ~0.038 per 10,000; $p < 2.2 \times 10^{-16}$). If we further restrict the analysis to the three largest cohorts, each of which included more than 5,000 probands, the observed enrichment remains highly significant (nine *de novo* LGD variants among 33,831 total probands; $p = 2.48 \times 10^{-14}$). We acknowledge that there are limitations to comparing results from ExAC to a clinically ascertained cohort, particularly when one undertakes a genotype-first approach by actively searching for singleton cases with variants in *NAA15* by using different sequencing platforms and coverage levels.[18] However, the average coverage for *NAA15* in ExAC and gnomAD databases is 47× versus approximately 20× coverage levels provided by clinically offered exome tests, suggesting that the increased number of LGD variants in the current study is not due to higher exon coverage levels in clinical sequencing. Only six LGD variants in *NAA15* are reported in ExAC (Table S5), and 11 *NAA15* LGD variants are reported in the Genome Aggregation Database (gnomAD) (Table S6). Two of the variants that are recurrent and *de novo* in our research cohort (c.239_240delAT [p.His 80Argfs*17] and c.228_232delCTTGA [p.Asp76Glufs*20]) are present one time each in ExAC (and also duplicated in gnomAD, given that gnomAD includes many variants from ExAC). It should be noted that phenotypic information as well as the variant inheritance are not available on these individuals in ExAC or gnomAD. Given that the three parents in the inherited families (families 10, 22,

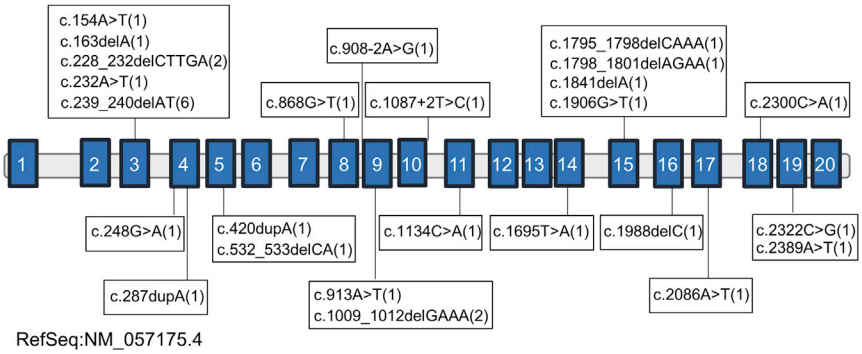


Figure 2. Exonic Localization of *NAA15* LGD Variants Identified in Subjects in This Study

Schematic representation of the genomic structure of human *NAA15*. Solid blue rectangles indicate exons, and the horizontal bars represent introns. *NAA15* variants with their relative positions in the gene are shown, and the number of affected individuals with the specific variants is shown in parentheses.

and 28) were only mildly affected, it is possible that such individuals could be found in cohorts such as ExAC or gnomAD. A recent study showed that ~2.8% of the ExAC population is associated with possible disease-associated genotypes,[19] and it is well-known that genetic background can influence the expressivity of any given variant.

We sought to confirm whether any of the LGD variants might trigger nonsense-mediated decay (NMD) of the respective mutant RNA. For this, we made use of two research-subject-derived cell lines, including one lymphoblastoid cell line (LCL) from individual 10-1 (c.239_240delAT) (Figures 3A–3C and Supplemental Methods) and one induced pluripotent stem cell (iPS) line from individual 19 (c.1009_1012delGAAA) (Figures 3D–3F). Quantitative RT-PCR with primers 3′ to the mutation demonstrated approximately 50% decreased total RNA in one cell passage from the LCLs from individual 10-1 in comparison to three control LCLs (Figure 3C, left panel), whereas the same assay (Figure 3F, right panel) and an additional Taqman assay (Figure S3) showed more variability in total RNA isolated from three different passages of the iPS line from individual 19 than from one control iPS line and a control human embryonic stem cell (hESC) line. Nonetheless, RT-PCR with primers spanning the mutation sites, followed by Sanger sequencing, did demonstrate substantially reduced mutant transcript in the LCL from individual 10-1 (Figure 3B) and almost complete absence of the mutant transcript in three different passages of the iPS line from individual 19 (a representative result from passage 16 is shown in Figure 3E). This reduction most likely occurs because the variant transcript is targeted for degradation via the nonsense-mediated decay (NMD) pathway.[20]

We further explored the functional effects for two of the other LGDs in a yeast assay in which the human NatA complex can functionally replace yeast NatA, as shown by complementation of growth phenotypes[21,22] and partial rescue of the NatA-specific Nt-acetylome.[23] Mutant *NAA15* (p.Thr55Hisfs*2 [c.163delA] from family 2 and p.Lys305* [c.913A>T] from family 18) failed to rescue the temperature-sensitive growth phenotype of yNatAΔ (Figure 4A and Tables S7, S8, and S9), suggesting that the two variants lead to reduced or abolished NatA activity, at least as assessed in this heterologous system. We further verified human NatA expression in the yNatA deletion strain by immunoblotting (Figure 4B) against the HA epitope that was incorporated N-terminal to NAA15. In the context of overexpression from a plasmid, we detected both full-length HA-NAA15 and HA-NAA15 p.Lys305*, but not HA-NAA15 p.Thr55Hisfs*2, suggesting that the mRNA for HA-NAA15 p.Thr55Hisfs*2 is most likely undergoing complete NMD and/or that this truncated mini-protein is unstable, whereas truncated mini-protein HA-NAA15 p.Lys305* can be expressed in this system but nonetheless does not provide functional rescue.

Distributed throughout the entire gene of *NAA15*, the 25 LGD variants we reported here are predicted to undergo NMD, leading to degradation of the mutant mRNA and thus loss of the aberrant protein product. Expression analysis from research-subject-derived lymphoblast cells or IPSCs confirmed under-representation of the mutant transcript in cDNA. In addition, the functional deficiency of human mutated *NAA15* was further supported by the growth rescue experiment in the yeast NatA-deficient strain, in which mutant human *NAA15* failed to restore the growth-deficiency phenotype. In light of these results, we propose haploinsufficiency of *NAA15* as the most likely mechanism for this newly recognized disease, although we readily acknowledge that some of the LGDs might not trigger complete NMD or might do so differentially in different tissues, leaving open the possibility for expression of a truncated NAA15 protein, which could possibly act via a dominant-negative or gain-of-function mechanism in some individuals. *De novo* missense variants (c.1014G>T [p.Lys338Asn] and c.841G>C [p.Glu281Gln]) have been previously reported in two individuals with autism and intellectual disability, respectively;[24,25] however, the deleterious effect of these missense variants has not been established and so will also require further functional studies, segregation in families, and/or proof of recurrence in multiple affected individuals. Further supporting our postulated mechanism of haploinsufficiency, when we searched the DECIPHER database[26] and our clinical cohorts for individuals with small microdeletions involving *NAA15*, the smallest deletion we could find is in a 31-year-old man carrying a *de novo* 2.73 Mb deletion, including *NAA15* and 17 other predicted genes. This man was noted as having mild intellectual disability, mild

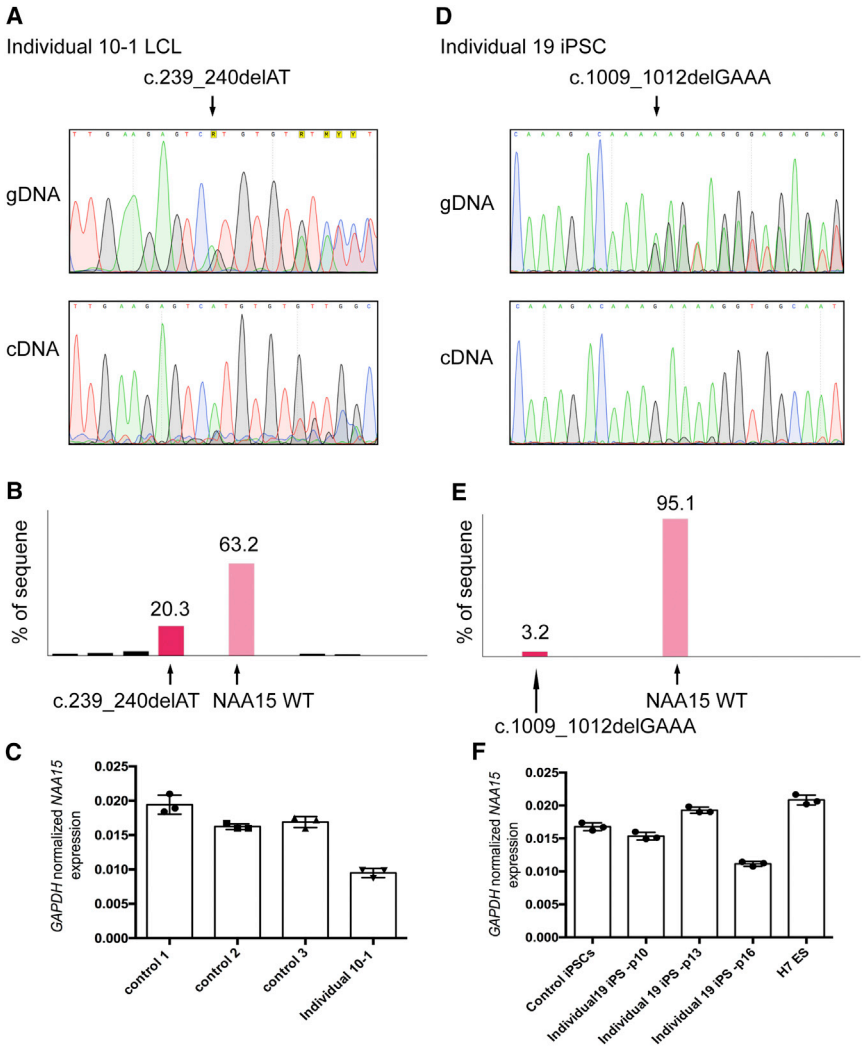


Figure 3. Expression Analysis of *NAA15* in Research-Subject-Derived Cell Lines

(A and D) Sanger sequencing of genomic DNA (top panel) and reverse-transcribed cDNA (bottom panel) isolated from a lymphoblastoid cell line (LCL) of individual 10-1 (c.239_240delAT) (A) and an induced pluripotent stem cell (iPS) line (passage 16) of individual 19 (c.1009_1012delGAAA) (D). (B and E) Quantification of different cDNA species from cDNA Sanger sequencing showing the relative ratio of WT *NAA15* versus c.239_240delAT (LCL line) (B) and (c.1009_1012delGAAA) (passage 16 iPS cell line) (E).

(C and F) *NAA15* mRNA expression level analyzed by qPCR in research subject-derived cell lines (at passage numbers p10, p13, and p16), as compared to control cell lines (at passage 16). Error bars are standard deviation (SD), and the assay was performed three times per sample.

dysmorphic features, motor delays in childhood, a low birth weight (−2SD), and adult height, weight, and head circumference all at the 10th centile (Figure S4). He has poor vision as a result of cortical visual impairment (CVI), which was not reported (but also not formally screened for) in any of the above reported individuals but which was found in some of the individuals with *NAA10* mutations.[27] It is also possible that his CVI could be due to some other missing gene in the CNV interval. There are currently 18 large heterozygous CNV deletions, including *NAA15* in the DECIPHER database;[26] these deletions range in size from 3.27 Mb to 24.30 Mb, and many are noted to be associated with global developmental delay or intellectual disability, supporting the case for haploinsufficiency of at least some of the genes in these CNV intervals. One individual with a *de novo* 5.2 Mb deletion died from a sudden cardiac event at the age of 35 (see Supplemental Case Reports).

Human *NAA15* encodes an 866 amino acid (~105 kDa) protein, NAA15, containing tetratricopeptide repeat domains and a putative bipartite nuclear localization signal.[28] Many studies have shown that NAA15 acts as the auxiliary subunit binding with the catalytic subunit NAA10 and localizes it to the ribosome, where this complex (named the NatA complex) serves as an N-terminal acetyltransferase (NAT).[29] This complex is evolutionarily conserved from yeast to vertebrates,[23] and the X-ray crystal structure of the 100 kDa holo-NatA complex from *Schizosaccharomyces pombe* shows that the NatA-NAA15 auxiliary subunit contains 13 tetratricopeptide motifs and adopts a ring-like topology that wraps around the NatA-NAA10 subunit, an interaction that alters the NAA10 active site for substrate-specific acetylation.[30] Mutation or loss of the NatA subunits in yeast (*Saccharomyces cerevisiae*) or human HeLa cells results in inhibited cell growth, cell apoptosis, and failure to enter the G_0 phase in the cell cycle.[31,32] *Nat1* (ortholog of *NAA15*) knock-down flies have impaired locomotor activity and early adult lethality.[6] *NAA10* and *NAA15* are both highly expressed in regions of cell division and migration during brain development and are downregulated as neurons differentiate in early postnatal development.[33,34] *NAA15* has been shown to be expressed at low levels in most adult tissues (e.g., nervous system, heart, and reproductive system) (see GTEx Portal). However, RNA-seq data from human brain tissue suggests that upregulation of *NAA15* occurs in utero at eight weeks after conception and is developmentally downregulated thereafter, the highest expression being in the occipital neocortex and anterior cingulate (medial prefrontal) cortex (Figures S5A and S5B), supporting a role for *NAA15* in development of the nervous system. Similarly, in mice, upregulated expression of *NAA15* has been shown in regions of neuronal

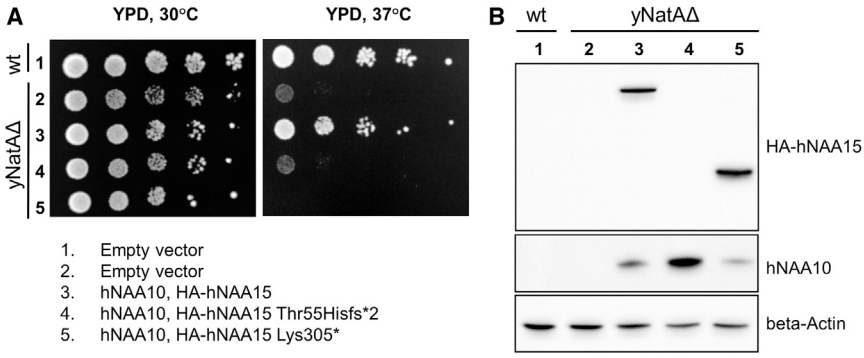


Figure 4. Truncation Mutations of Human NAA15 Impair NatA Function and Yeast Viability

(A) Serial dilution spot assay depicting the sensitivity of human *NAA15* Thr55Hisfs*2 and Lys305* mutants to increased temperature in a *ynaa10Δ*, *ynaa15Δ* double-deletion background (*yNatAΔ*).

(B) Confirmation of human NatA expression by immunoblot analysis with anti-hNAA10 and anti-HA (for HA-hNAA15 detection) along with anti-beta Actin as a loading control.

migration, and proliferation in the neonatal mouse brain has been shown along with reduced expression as neurons differentiate during early postnatal development.[33,34]

Genetic defects in *NAA10*, which is X-linked and encodes another member of the NatA complex, are associated with Ogden syndrome (MIM: 300855), Lenz microphthalmia (MIM: 309800), and intellectual disability (with variable cardiac involvement).[27,35–39] In the case of Ogden syndrome, a total of eight boys from two families had a distinct combination of dysmorphology, hypotonia, global developmental delays, cardiac anomalies, cardiac arrhythmias and cardiomegaly, and the identical missense mutation segregated in multiple affected individuals in two unrelated families.[40] Different variants in *NAA10* have been reported, sometimes with only a mild intellectual-disability phenotype in heterozygous females, but also sometimes with hydrocephaly, supernumerary vertebrae, congenital heart defects, and arrhythmias, which are always more severe in the males.[9,27,35–37] Although developmental delay and/or intellectual disability might be the only presenting feature, the additional cardiac, growth, dysmorphic features and other findings vary in type and severity. For the one family in which affected members had Lenz microphthalmia syndrome and a splice-site variant in *NAA10,* and in which proband-derived fibroblasts lacked expression of full-length NAA10 and displayed a cell-proliferation defect,[41] it is not known why this family alone has such a dramatic ocular phenotype, although it is worth noting that 9/13 (69%) of the female subjects reported with missense variants in *NAA10* had some milder form of eye anomalies, including astigmatism, hyperopia and/or myopia.[27] Most studies have reported that missense mutations in *NAA10* decrease the enzymatic function of NAA10 and/or decrease its binding to NAA15.[21,22,27,35,39,40]

In total, the presentations involving *NAA10* and *NAA15* appear to have phenotypic overlap but variability, and as such should be referred to more broadly as "*NAA10*-related syndrome" and "*NAA15*-related syndrome." The extensive phenotypic variability is most likely related to genetic background differences and also to the spatial and temporal tissue-specific acetylation of a few N-terminal acetylation substrates by the NatA complex, although there are also suggested N-terminal acetylation (NTA)-independent functions for *NAA10*.[38,42] In the past few years, the first instance of NTA with relevance to cardiac function was reported and involved NTA of the cardiac voltage-gated sodium channel, Nav1.5, in tissues from individuals with end-stage heart failure.[43,44] Indeed, protein quality control is of major relevance in heart failure.[45] Also, a 2015 study linked NTA and N-end-rule degradation to blood pressure regulation[46,47] and demonstrated that N-terminal mutants of Rgs2, a key G-protein regulator, are differentially processed by NATs and the two branches of the N-end-rule pathway, leading to an imbalance in the signaling governing blood pressure. In regard to more common diseases and basic biology, there is emerging evidence that NTA of proteins are overexpressed or otherwise dysregulated in a variety of cancers, including lung, prostate, and liver cancers.[48–54] NTA has been linked to neurodegenerative diseases such as Parkinson, Alzheimer, and Huntington disease, and NatA/NAA10 has been shown to contribute to the regulation of amyloid β-protein generation, to modulate the stabilization of Sup35 amyloid formation, and to prevent aggregation of Htt,[55–60] supporting the importance of NTA in the progression of these diseases. Current findings link NTA to degradation of some proteins via Ac/N-degron-mediated recruitment of specific ubiquitin ligases.[47,61–64] NTA might also influence protein-complex formation, as exemplified by the NEDD8 ligation enzymes,[65] along with prion formation.[60] Also, protein-specific targeting to membranes of the nucleus,[66] Golgi[67,68] and lysosomes[69] was shown to require NTA, but a general role in targeting is not supported.[39,40,70]

In conclusion, we propose that disruption of NatA complex functionality can cause developmental disorders with variable expressivity. Future identification of additional affected individuals and studies in model organisms will be required if we are to continue to refine the clinical phenotype and determine the underlying mechanism whereby reduced expression or perturbed function of *NAA15* results in these phenotypes.

Supplemental Data

Supplemental Data include Supplemental Case reports, Figures S1–S4, a Supplemental Note, Supplemental Materials and Methods, Figures S1–S5, Tables S1–S9, Supplemental References,

and Supplemental Acknowledgments and can be found with this article online at https://doi.org/10.1016/j.ajhg.2018.03.004.

Acknowledgments

T.A. thanks Nina Glomnes for technical support. We appreciate the study participants and their families, without whom this work would not have been possible. Megan Cho at GeneDx facilitated the collaborations with the clinicians who submitted samples to GeneDx. The Baylor-Hopkins Center for Mendelian Genomics (BHCMG) has been funded by the US National Human Genome Research Institute (NHGRI)/National Heart Lung and Blood Institute (NHLBI) grant number UM1HG006542. The DDD Study presents independent research commissioned by the Health Innovation Challenge Fund (grant HICF-1009-003), a parallel funding partnership between the Wellcome Trust and the Department of Health, and the Wellcome Trust Sanger Institute (grant WT098051). H.A.F.S. is supported by funds from the LB692 Nebraska Tobacco Settlement Biomedical Research Development Program. J.E.P. was supported by NHGRI grant K08 HG008986. This research was further supported in part by the following to E.E.E.: the Simons Foundation Autism Research Initiative (SFARI 303241 and 337701) and National Institutes of Health (R01MH101221). E.E.E. is an investigator of the Howard Hughes Medical Institute. The collection of the proband from HudsonAlpha was supported by Clinical Sequencing Evidence-Generating Research (CSER) study grant number (NIH/NHGRI 4UM1HG007301-04). S.V. and T.A. were funded by the Research Council of Norway (249843), the Norwegian Cancer Society (PR-2009-0222), and the Norwegian Health Authorities of Western Norway (912176). G.J.L. is supported by funds from the Stanley Institute for Cognitive Genomics at Cold Spring Harbor Laboratory. Additional acknowledgments can be found in the Supplemental Data.

Declaration of Interests

G.J.L serves on advisory boards for GenePeeks and Seven Bridges Genomics. The Department of Molecular and Human Genetics at BCM derives revenue from molecular testing offered at Baylor Genetics Laboratories. J.R.L has stock ownership in 23 and Me, is a paid consultant for Regeneron Pharmaceuticals, has stock options in Lasergen, and is a co-inventor on multiple United States and European patents related to molecular diagnostics for inherited neuropathies, eye diseases, and bacterial genomic fingerprinting. E.E.E. is on the scientific advisory board of DNAnexus. W.K.C. is on the scientific advisory board of the Regeneron Genetics Center. Richard Person and Rebecca Willaert are employees of GeneDx, a wholly owned subsidiary of OPKO Health.

Received: November 13, 2017
Accepted: February 27, 2018
Published: April 12, 2018

Web Resources

BrainSpan: Atlas of the Developing Brain, http://www.brainspan.org/ (accessed 09/22/17)
ExAC, http://exac.broadinstitute.org/
gnomAD, http://gnomad.broadinstitute.org/
GTEx, https://www.gtexportal.org/ (accessed 10/26/17)
OMIM, http://www.omim.org

References

1. Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N. Engl. J. Med. *369*, 1502–1511.
2. Yang, Y., Muzny, D.M., Xia, F., Niu, Z., Person, R., Ding, Y., Ward, P., Braxton, A., Wang, M., Buhay, C., et al. (2014). Molecular findings among patients referred for clinical whole-exome sequencing. JAMA *312*, 1870–1879.
3. Farwell, K.D., Shahmirzadi, L., El-Khechen, D., Powis, Z., Chao, E.C., Tippin Davis, B., Baxter, R.M., Zeng, W., Mroske, C., Parra, M.C., et al. (2015). Enhanced utility of family-centered diagnostic exome sequencing with inheritance model-based analysis: results from 500 unselected families with undiagnosed genetic conditions. Genet. Med. *17*, 578–586.
4. Lee, H., Deignan, J.L., Dorrani, N., Strom, S.P., Kantarci, S., Quintero-Rivera, F., Das, K., Toy, T., Harry, B., Yourshaw, M., et al. (2014). Clinical exome sequencing for genetic identification of rare Mendelian disorders. JAMA *312*, 1880–1887.
5. Retterer, K., Juusola, J., Cho, M.T., Vitazka, P., Millan, F., Gibellini, F., Vertino-Bell, A., Smaoui, N., Neidich, J., Monaghan, K.G., et al. (2016). Clinical application of whole-exome sequencing across clinical indications. Genet. Med. *18*, 696–704.
6. Stessman, H.A., Xiong, B., Coe, B.P., Wang, T., Hoekzema, K., Fenckova, M., Kvarnung, M., Gerdts, J., Trinh, S., Cosemans, N., et al. (2017). Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. Nat. Genet. *49*, 515–526.
7. Zaidi, S., Choi, M., Wakimoto, H., Ma, L., Jiang, J., Overton, J.D., Romano-Adesman, A., Bjornson, R.D., Breitbart, R.E., Brown, K.K., et al. (2013). De novo mutations in histone-modifying genes in congenital heart disease. Nature *498*, 220–223.
8. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum. Mutat. *36*, 928–930.
9. Sidhu, M., Brady, L., Tarnopolsky, M., and Ronen, G.M. (2017). Clinical manifestations associated with the N-terminal-acetyltransferase NAA10 gene mutation in a girl: Ogden Syndrome. Pediatr. Neurol. *76*, 82–85.
10. Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J., et al. (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. Science *350*, 1262–1266.
11. Jin, S.C., Homsy, J., Zaidi, S., Lu, Q., Morton, S., DePalma, S.R., Zeng, X., Qi, H., Chang, W., Sierant, M.C., et al. (2017). Contribution of rare inherited and de novo variants in 2,871 congenital heart disease probands. Nat. Genet. *49*, 1593–1601.
12. Růžička, M., Kulhánek, P., Radová, L., Čechová, A., Špačková, N., Fajkusová, L., and Réblová, K. (2017). DNA mutation motifs in the genes associated with inherited diseases. PLoS ONE *12*, e0182377.
13. Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. *31*, 3406–3415.
14. Ripley, L.S. (1982). Model for the participation of quasi-palindromic DNA sequences in frameshift mutation. Proc. Natl. Acad. Sci. USA *79*, 4128–4132.

15. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.
16. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. PLoS Genet. *9*, e1003709.
17. Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnström, K., Mallick, S., Kirby, A., et al. (2014). A framework for the interpretation of de novo mutation in human disease. Nat. Genet. *46*, 944–950.
18. Barrett, J.C., Buxbaum, J., Cutler, D., Daly, M., Devlin, B., Gratten, J., Hurles, M.E., Kosmicki, J.A., Lander, E.S., MacArthur, D.G., et al. (2017). New mutations, old statistical challenges. bioRxiv. https://doi.org/10.1101/115964.
19. Tarailo-Graovac, M., Zhu, J.Y.A., Matthews, A., van Karnebeek, C.D.M., and Wasserman, W.W. (2017). Assessment of the ExAC data set for the presence of individuals with pathogenic genotypes implicated in severe Mendelian pediatric disorders. Genet. Med. *19*, 1300–1308.
20. Lykke-Andersen, S., and Jensen, T.H. (2015). Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nat. Rev. Mol. Cell Biol. *16*, 665–677.
21. Van Damme, P., Støve, S.I., Glomnes, N., Gevaert, K., and Arnesen, T. (2014). A Saccharomyces cerevisiae model reveals in vivo functional impairment of the Ogden syndrome N-terminal acetyltransferase NAA10 Ser37Pro mutant. Mol. Cell. Proteomics *13*, 2031–2041.
22. Dörfel, M.J., Fang, H., Crain, J., Klingener, M., Weiser, J., and Lyon, G.J. (2017). Proteomic and genomic characterization of a yeast model for Ogden syndrome. Yeast *34*, 19–37.
23. Arnesen, T., Van Damme, P., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., Varhaug, J.E., Vandekerckhove, J., Lillehaug, J.R., Sherman, F., and Gevaert, K. (2009). Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. Proc. Natl. Acad. Sci. USA *106*, 8157–8162.
24. De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al.; DDD Study; Homozygosity Mapping Collaborative for Autism; and UK10K Consortium (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature *515*, 209–215.
25. Zhao, J.J., Halvardson, J., Zander, C.S., Zaghlool, A., Georgii-Hemming, P., Mansson, E., Brandberg, G., Savmarker, H.E., Frykholm, C., Kuchinskaya, E., et al. (2018). Exome sequencing reveals NAA15 and PUF60 as candidate genes associated with intellectual disability. Am. J. Med. Genet. B. Neuropsychiatr. Genet. *177*, 10–20.
26. Bragin, E., Chatzimichali, E.A., Wright, C.F., Hurles, M.E., Firth, H.V., Bevan, A.P., and Swaminathan, G.J. (2014). DECIPHER: database for the interpretation of phenotype-linked plausibly pathogenic sequence and copy-number variation. Nucleic Acids Res. *42*, D993–D1000.
27. Saunier, C., Støve, S.I., Popp, B., Gérard, B., Blenski, M., Ahmew, N., de Bie, C., Goldenberg, P., Isidor, B., Keren, B., et al. (2016). Expanding the phenotype associated with NAA10-related N-terminal acetylation deficiency. Hum. Mutat. *37*, 755–764.
28. Fluge, Ø., Bruland, O., Akslen, L.A., Varhaug, J.E., and Lillehaug, J.R. (2002). NATH, a novel gene overexpressed in papillary thyroid carcinomas. Oncogene *21*, 5056–5068.
29. Arnesen, T., Anderson, D., Baldersheim, C., Lanotte, M., Varhaug, J.E., and Lillehaug, J.R. (2005). Identification and characterization of the human ARD1-NATH protein acetyltransferase complex. Biochem. J. *386*, 433–443.
30. Liszczak, G., Goldberg, J.M., Foyn, H., Petersson, E.J., Arnesen, T., and Marmorstein, R. (2013). Molecular basis for N-terminal acetylation by the heterodimeric NatA complex. Nat. Struct. Mol. Biol. *20*, 1098–1105.
31. Mullen, J.R., Kayne, P.S., Moerschell, R.P., Tsunasawa, S., Gribskov, M., Colavito-Shepanski, M., Grunstein, M., Sherman, F., and Sternglanz, R. (1989). Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. EMBO J. *8*, 2067–2075.
32. Arnesen, T., Gromyko, D., Pendino, F., Ryningen, A., Varhaug, J.E., and Lillehaug, J.R. (2006). Induction of apoptosis in human cells by RNAi-mediated knockdown of hARD1 and NATH, components of the protein N-alpha-acetyltransferase complex. Oncogene *25*, 4350–4360.
33. Sugiura, N., Adams, S.M., and Corriveau, R.A. (2003). An evolutionarily conserved N-terminal acetyltransferase complex associated with neuronal development. J. Biol. Chem. *278*, 40113–40120.
34. Sugiura, N., Patel, R.G., and Corriveau, R.A. (2001). N-methyl-D-aspartate receptors regulate a group of transiently expressed genes in the developing brain. J. Biol. Chem. *276*, 14257–14263.
35. Popp, B., Støve, S.I., Endele, S., Myklebust, L.M., Hoyer, J., Sticht, H., Azzarello-Burri, S., Rauch, A., Arnesen, T., and Reis, A. (2015). De novo missense mutations in the NAA10 gene cause severe non-syndromic developmental delay in males and females. Eur. J. Hum. Genet. *23*, 602–609.
36. Casey, J.P., Støve, S.I., McGorrian, C., Galvin, J., Blenski, M., Dunne, A., Ennis, S., Brett, F., King, M.D., Arnesen, T., and Lynch, S.A. (2015). NAA10 mutation causing a novel intellectual disability syndrome with Long QT due to N-terminal acetyltransferase impairment. Sci. Rep. *5*, 16022.
37. Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., Di Donato, N., et al. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet *380*, 1674–1682.
38. Dörfel, M.J., and Lyon, G.J. (2015). The biological functions of Naa10 - From amino-terminal acetylation to human disease. Gene *567*, 103–131.
39. Myklebust, L.M., Van Damme, P., Støve, S.I., Dörfel, M.J., Abboud, A., Kalvik, T.V., Grauffel, C., Jonckheere, V., Wu, Y., Swensen, J., et al. (2015). Biochemical and cellular analysis of Ogden syndrome reveals downstream Nt-acetylation defects. Hum. Mol. Genet. *24*, 1956–1976.
40. Rope, A.F., Wang, K., Evjenth, R., Xing, J., Johnston, J.J., Swensen, J.J., Johnson, W.E., Moore, B., Huff, C.D., Bird, L.M., et al. (2011). Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency. Am. J. Hum. Genet. *89*, 28–43.
41. Esmailpour, T., Riazifar, H., Liu, L., Donkervoort, S., Huang, V.H., Madaan, S., Shoucri, B.M., Busch, A., Wu, J., Towbin, A., et al. (2014). A splice donor mutation in NAA10 results in the dysregulation of the retinoic acid signalling pathway and causes Lenz microphthalmia syndrome. J. Med. Genet. *51*, 185–196.

42. Lee, C.C., Peng, S.H., Shen, L., Lee, C.F., Du, T.H., Kang, M.L., Xu, G.L., Upadhyay, A.K., Cheng, X., Yan, Y.T., et al. (2017). The role of N-alpha-acetyltransferase 10 protein in DNA methylation and genomic imprinting. Mol. Cell *68*, 89–103.e7.
43. Beltran-Alvarez, P., Tarradas, A., Chiva, C., Pérez-Serra, A., Batlle, M., Pérez-Villa, F., Schulte, U., Sabidó, E., Brugada, R., and Pagans, S. (2014). Identification of N-terminal protein acetylation and arginine methylation of the voltage-gated sodium channel in end-stage heart failure human heart. J. Mol. Cell. Cardiol. *76*, 126–129.
44. Marionneau, C., and Abriel, H. (2015). Regulation of the cardiac Na+ channel NaV1.5 by post-translational modifications. J. Mol. Cell. Cardiol. *82*, 36–47.
45. Ferreira, J.C., Boer, B.N., Grinberg, M., Brum, P.C., and Mochly-Rosen, D. (2012). Protein quality control disruption by PKCβII in heart failure; rescue by the selective PKCβII inhibitor, βIIV5-3. PLoS ONE *7*, e33175.
46. Aksnes, H., Drazic, A., and Arnesen, T. (2015). (Hyper)tension release by N-terminal acetylation. Trends Biochem. Sci. *40*, 422–424.
47. Park, S.E., Kim, J.M., Seok, O.H., Cho, H., Wadas, B., Kim, S.Y., Varshavsky, A., and Hwang, C.S. (2015). Control of mammalian G protein signaling by N-terminal acetylation and the N-end rule pathway. Science *347*, 1249–1252.
48. Min, L., Xu, H., Wang, J., Qu, L., Jiang, B., Zeng, Y., Meng, L., Jin, H., and Shou, C. (2013). N-α-acetyltransferase 10 protein is a negative regulator of 28S proteasome through interaction with PA28β. FEBS Lett. *587*, 1630–1637.
49. Shim, J.H., Chung, Y.H., Kim, J.A., Lee, D., Kim, K.M., Lim, Y.S., Lee, H.C., Lee, Y.S., Yu, E., and Lee, Y.J. (2012). Clinical implications of arrest-defective protein 1 expression in hepatocellular carcinoma: A novel predictor of microvascular invasion. Dig. Dis. *30*, 603–608.
50. Wang, Z., Wang, Z., Guo, J., Li, Y., Bavarva, J.H., Qian, C., Brahimi-Horn, M.C., Tan, D., and Liu, W. (2012). Inactivation of androgen-induced regulator ARD1 inhibits androgen receptor acetylation and prostate tumorigenesis. Proc. Natl. Acad. Sci. USA *109*, 3053–3058.
51. Wang, Z.H., Gong, J.L., Yu, M., Yang, H., Lai, J.H., Ma, M.X., Wu, H., Li, L., and Tan, D.Y. (2011). Up-regulation of human arrest-defective 1 protein is correlated with metastatic phenotype and poor prognosis in breast cancer. Asian Pac. J. Cancer Prev. *12*, 1973–1977.
52. Hua, K.T., Tan, C.T., Johansson, G., Lee, J.M., Yang, P.W., Lu, H.Y., Chen, C.K., Su, J.L., Chen, P.B., Wu, Y.L., et al. (2011). N-α-acetyltransferase 10 protein suppresses cancer cell metastasis by binding PIX proteins and inhibiting Cdc42/Rac1 activity. Cancer Cell *19*, 218–231.
53. Lee, C.-F., Ou, D.S.C., Lee, S.-B., Chang, L.-H., Lin, R.-K., Li, Y.-S., Upadhyay, A.K., Cheng, X., Wang, Y.-C., Hsu, H.-S., et al. (2010). hNaa10p contributes to tumorigenesis by facilitating DNMT1-mediated tumor suppressor gene silencing. J. Clin. Invest. *120*, 2920–2930.
54. Kalvik, T.V., and Arnesen, T. (2013). Protein N-terminal acetyltransferases in cancer. Oncogene *32*, 269–276.
55. Asaumi, M., Iijima, K., Sumioka, A., Iijima-Ando, K., Kirino, Y., Nakaya, T., and Suzuki, T. (2005). Interaction of N-terminal acetyltransferase with the cytoplasmic domain of beta-amyloid precursor protein and its effect on A beta secretion. J. Biochem. *137*, 147–155.
56. Pezza, J.A., Villali, J., Sindi, S.S., and Serio, T.R. (2014). Amyloid-associated activity contributes to the severity and toxicity of a prion phenotype. Nat. Commun. *5*, 4384.
57. Raychaudhuri, S., Sinha, M., Mukhopadhyay, D., and Bhattacharyya, N.P. (2008). HYPK, a Huntingtin interacting protein, reduces aggregates and apoptosis induced by N-terminal Huntingtin with 40 glutamines in Neuro2a cells and exhibits chaperone-like activity. Hum. Mol. Genet. *17*, 240–255.
58. Arnesen, T., Starheim, K.K., Van Damme, P., Evjenth, R., Dinh, H., Betts, M.J., Ryningen, A., Vandekerckhove, J., Gevaert, K., and Anderson, D. (2010). The chaperone-like protein HYPK acts together with NatA in cotranslational N-terminal acetylation and prevention of Huntingtin aggregation. Mol. Cell. Biol. *30*, 1898–1909.
59. Holmes, W.M., Klaips, C.L., and Serio, T.R. (2014). Defining the limits: Protein aggregation and toxicity in vivo. Crit. Rev. Biochem. Mol. Biol. *49*, 294–303.
60. Holmes, W.M., Mannakee, B.K., Gutenkunst, R.N., and Serio, T.R. (2014). Loss of amino-terminal acetylation suppresses a prion phenotype by modulating global protein folding. Nat. Commun. *5*, 4383.
61. Hwang, C.-S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. Science *327*, 973–977.
62. Hwang, C.S., Shemorry, A., Auerbach, D., and Varshavsky, A. (2010). The N-end rule pathway is mediated by a complex of the RING-type Ubr1 and HECT-type Ufd4 ubiquitin ligases. Nat. Cell Biol. *12*, 1177–1185.
63. Shemorry, A., Hwang, C.-S., and Varshavsky, A. (2013). Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. Mol. Cell *50*, 540–551.
64. Kim, H.K., Kim, R.R., Oh, J.H., Cho, H., Varshavsky, A., and Hwang, C.S. (2014). The N-terminal methionine of cellular proteins as a degradation signal. Cell *156*, 158–169.
65. Scott, D.C., Monda, J.K., Bennett, E.J., Harper, J.W., and Schulman, B.A. (2011). N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. Science *334*, 674–678.
66. Murthi, A., and Hopper, A.K. (2005). Genome-wide screen for inner nuclear membrane protein targeting in Saccharomyces cerevisiae: roles for N-acetylation and an integral membrane protein. Genetics *170*, 1553–1560.
67. Behnia, R., Panic, B., Whyte, J.R.C., and Munro, S. (2004). Targeting of the Arf-like GTPase Arl3p to the Golgi requires N-terminal acetylation and the membrane protein Sys1p. Nat. Cell Biol. *6*, 405–413.
68. Setty, S.R.G., Strochlic, T.I., Tong, A.H.Y., Boone, C., and Burd, C.G. (2004). Golgi targeting of ARF-like GTPase Arl3p requires its Nalpha-acetylation and the integral membrane protein Sys1p. Nat. Cell Biol. *6*, 414–419.
69. Hofmann, I., and Munro, S. (2006). An N-terminally acetylated Arf-like GTPase is localised to lysosomes and affects their motility. J. Cell Sci. *119*, 1494–1503.
70. Aksnes, H., Osberg, C., and Arnesen, T. (2013). N-terminal acetylation by NatC is not a general determinant for substrate subcellular localization in Saccharomyces cerevisiae. PLoS ONE *8*, e61012.

The American Journal of Human Genetics, Volume *102*

Supplemental Data

Truncating Variants in *NAA15* Are Associated with Variable Levels of Intellectual Disability, Autism Spectrum Disorder, and Congenital Anomalies

Hanyin Cheng, Avinash V. Dharmadhikari, Sylvia Varland, Ning Ma, Deepti Domingo, Robert Kleyner, Alan F. Rope, Margaret Yoon, Asbjørg Stray-Pedersen, Jennifer E. Posey, Sarah R. Crews, Mohammad K. Eldomery, Zeynep Coban Akdemir, Andrea M. Lewis, Vernon R. Sutton, Jill A. Rosenfeld, Erin Conboy, Katherine Agre, Fan Xia, Magdalena Walkiewicz, Mauro Longoni, Frances A. High, Marjon A. van Slegtenhorst, Grazia M.S. Mancini, Candice R. Finnila, Arie van Haeringen, Nicolette den Hollander, Claudia Ruivenkamp, Sakkubai Naidu, Sonal Mahida, Elizabeth E. Palmer, Lucinda Murray, Derek Lim, Parul Jayakar, Michael J. Parker, Stefania Giusto, Emanuela Stracuzzi, Corrado Romano, Jennifer S. Beighley, Raphael A. Bernier, Sébastien Küry, Mathilde Nizon, Mark A. Corbett, Marie Shaw, Alison Gardner, Christopher Barnett, Ruth Armstrong, Karin S. Kassahn, Anke Van Dijck, Geert Vandeweyer, Tjitske Kleefstra, Jolanda Schieving, Marjolijn J. Jongmans, Bert B.A. de Vries, Rolph Pfundt, Bronwyn Kerr, Samantha K. Rojas, Kym M. Boycott, Richard Person, Rebecca Willaert, Evan E. Eichler, R. Frank Kooy, Yaping Yang, Joseph C. Wu, James R. Lupski, Thomas Arnesen, Gregory M. Cooper, Wendy K. Chung, Jozef Gecz, Holly A.F. Stessman, Linyan Meng, and Gholson J. Lyon

Supplementary Information Case Reports

Individual 1 in Family 1: Subject ID: D1.11.09640 (Leiden) – was listed as Subject 1 in Stessman et al. 2017 Nature Genetics paper (Eichler panel)
Event: inheritance unknown *NAA15* nonsense variant c.154A>T (p.Lys52*)

Child (female) was born in 1994 presented with ID, cutis marmorata, and small head circumference with proportionate short stature (both -2 SDs). Positive family history; both parents had lower education.

Individual 2 in Family 2: subject ID 8524555 (Rotterdam)
Event: *de novo NAA15* frameshift variant c.163delA (p.Thr55Hisfs*2)
This 17 years old boy presented with neonatal multifocal atrial ectopic tachycardia, requiring Sotalol treatment in the neonatal period. At the age of 2 years the medication was replaced by labetalol. At the moment, the cardiologist has put him under the regimen of 40mg verapamil t.i.d. From the age of 4 years he developed absence epilepsy treated with valproate. Medication was discontinued at the age of 13 years and he is seizure free. He also shows a global developmental delay and autism. He is overweight and has hypertension. A brain MRI scan at the age of 13 years showed mild global cerebellar atrophy, atrophy of the left hippocampus and thin corpus callosum.

Individual 3 in Family 3:
Event: de novo NAA15 frameshift variant c.228_232delCTTGA (p.Asp76Glufs*20)
Male individual was 8 years old at the time of diagnosis: NAA15; Chr4(GRCh37):g.140258090_140258094del; NM_057175.3:c.228_232del (p.(Asp76fs)) de novo. He is the first child of unrelated parents. He was born after an uneventful pregnancy and delivery at 40 weeks and 5 days. His birth weight was 3510 gram and his APGAR score 9 after 1 minute and 10 after 5 minutes. There were some difficulties maintaining breast feeding due to a weak suck, and he was changed to formula feeds at 2 weeks.

At the age of 4 weeks during regularly control a cardiac murmur was heard and he was sent to the pediatric cardiology award. There were no clinical complaints. Ultrasound of the heart revealed a ventricular spetum defect with hemodynamic consequences. He was operated at the age of 3 months.

There was hypotonia from birth on. His motor milestones were delayed. Walking independently at the age of two years. First words also at the age of two years. At the age of three years he developed a low frequency paroxysmal movement disorder with dystonia in his left leg. This disappeared spontaneously at the age of five years. No problems with seeing, hearing, sleeping, no epilepsy.

At examination a mild synophrys, high forehead, length on -1 SD, weight on -1 SD, head circumference at -1 SD.

He has a mild intellectual disability with delayed development. His total IQ is 65 (WISC-III-NL 2017), with similar verbal and performal IQ. He was diagnosed with autism spectrum disorder He was also noted to have attention issues. The boy is rather shy and easily anxious. His motor development score on the m-ABC-2 in 2017 was p0,1. Reflexes at examination were low.
Family history negative for developmental delay and cardiac diseases.
Examinations:

- Cerebral MRI showed no abnormalities.
- Array showed a normal male profile
- metabolic screening of blood and urine was normal
- Sanger sequencing of ATP1A3 and GOSR2 were normal.

Internal problems:

- profound perspiration from birth
- constipation
- easily tired
- eczema

Individual 4 in Family 4 : AU025403 (AGRE) also 03C14733
Event: d*e novo NAA15* frameshift variant c.228_232delCTTGA (p.Asp76Glufs*20)
Subject is male with a diagnosis of autism spectrum disorder (ASD). He was 17-years-old at the time of last assessment. He began walking at age 15 months, established bowel control at age 42 months, used first words at age 48 months and phrases at 60 months. Adaptive behavior was found to be within the average range based on parent interview conducted at age 9 years.

Subject appears to be high functioning with results of results of a gold-standard autism observation measure demonstrating limited features of ASD during an appointment conducted at age 10.14 years. At that time, the subject demonstrated unusual eye contact and used limited

gestures; however, he engaged in reciprocal conversations and report events, directed facial expressions to the examiner, shared enjoyment in his interactions with the examiner, responded appropriately to the examiner and initiated interactions well, and was rated as having good rapport. There were no restricted or repetitive behaviors or unusual sensory interests observed.

Based on parent report, subject uses repetitive language and has difficulty with reciprocal conversations. Gesture use is limited as are his imitation skills. He has a history of using others' bodies as a tool, and had delayed language development (receptive and expressive delays). Social difficulties were described (e.g., deficits in sharing, play skills, friendship behavior). Subject has unusual preoccupations that are interfering as well as circumscribed interests. He displays sensory sensitivities and has a history of stereotyped body movements. Medical information is not known.

Individual 5 in Family 5:
Event: inheritance unknown *NAA15* nonsense variant c.232A>T (p.Lys78*)

No further information provided other than what is in table.

Individual 6 in Family 6:
Event: *de novo NAA15* frameshift variant c.239_240delAT (p.His80Argfs*17)
This is a six-year-old African-American girl, with abnormal behaviors characterized by spinning early in life, echolalic speech, clapping, irregular respiration and teeth grinding, consistent with autism. Rett syndrome was excluded. Dysmorphisms include a small up turned nose, tented upper lip and frontal bossing. She was able to sit and roll and walk by 14 months, she is now able to climb but walks on her toes. She is able to feed herself with a spoon, she has some stereotypies including clapping. The major concern for her is speech which consisted of 5-10 single words before 18 moths when she had a regression and was non-verbal by 2 years of age. She has regained some language, but they are still single words, or two words put together. She recognizes numbers and sings the alphabet. Her neurologic exam is grossly normal except for noted hypotonia. She has multiple food allergies, including dairy, wheat, eggs, walnuts, shellfish, peanuts, bees, and sunflower seeds. The mother had placenta previa with failure to progress and fetal distress, necessitating C-section delivery. However, the clinical features do not resemble cerebral palsy secondary to fetal ischemic brain insult from placenta previa and bleeding. Head circumference is at 50th percentile, and she had a normal MRI and

normal EEG. Cardiac exam is normal, with no evidence for any cardiac disease. She is not on any medication. The proband has a sibling with speech delay, but who does not have the mutation in *NAA15*.

Individual 7 in Family 7:
Event: *de novo NAA15* frameshift variant c.239_240delAT (p.His80Argfs*17)

No further information provided other than what is in table.

Individual 8 in Family 8: Subject ID: 04147-8645 (Troina) – was listed as Subject 2 in Stessman et al. 2017 Nature Genetics paper
Event: inheritance unknown *NAA15* frameshift variant c.239_240delAT (p.His80Argfs*17)

Girl born in 2002 seen for the first time at 8 years and diagnosed as moderate ID (ICD-9-CM Diagnosis Code 318.0) and Other Specified Pervasive Developmental Disorder, Current or Active State (ICD-9-CM Diagnosis Code 299.80). ID and psychiatric illness are reported in her pedigree, but further detail is unavailable. She was born by dystocic term delivery, and neonatal cyanosis was reported. Phenotype at 8 years showed growth retardation (weight and height on the 3^{rd}-10^{th} percentile ranges), microcephaly (OFC $<2^{nd}$ percentile), facial asymmetry, hypertelorism, prominent antehelix 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. EEG and brain MRI were normal. Agilent 60K array CGH was normal, and MLPA test for 11p11.2 chromosomal band didn't show any deletion/duplication. Family history is notable for intellectual disability in a paternal uncle, and behavioral issues in the paternal grandfather.

Individuals 10-1, 10-2 and 10-3 in Family 10: Sydney Family with *NAA15* variant

Event: Maternally inherited *NAA15* frameshift variant c.239_240delAT (p.His80Argfs*17)

Individual 10-1 last assessment aged 17 years:
Pregnancy and birth: The pregnancy was relatively uncomplicated, although there was a two-day period of unexplained reduced fetal movements at 20 weeks gestation. He was delivered at 36 weeks with a birth-weight of 2930 g (75^{th} percentile), length 47cm (5^{th} percentile) and head circumference 34cm (50^{th}-75^{th} percentile). He had reduced tone and a slow suck, and required

head-box oxygen for 24 hours, and orogastric feeding for two weeks. He was discharged home breast-feeding after 2 weeks.

Development: He was delayed in all developmental milestones: smiling at 3 months, rolling at 8-9 months, sitting and crawling at a year and walking at 2 years. He only had a few words at 2 years of age. As a baby he rocked himself to sleep. He had no imaginary play, and would tend to line up his cars in rows. He had many sensory symptoms and spent a lot of time looking at fans. He was also attracted to car headlights and would notice small details. He had difficulties as an infant with transitions or changes to the routine and had frequent 'meltdowns' several times a day. He was very sensitive to lights and certain noises. He had a formal developmental assessment at 3 years of age and was diagnosed with global developmental delay (most skills in low mild to moderate delay range). He did not meet diagnostic criteria for autism at that assessment, although he was later diagnosed with Autism by his pediatrician in early childhood. He was also diagnosed with attention deficit hyperactivity disorder aged 3 and trialled on Ritalin: however, his parents felt his behavior deteriorated on Ritalin and this was ceased around the age of 4.

Schooling: He attends a support class (IM) at a mainstream school. His parents report no ongoing significant concerns with his concentration.

Mental health: He has had episodes of anxiety and depression, with poor sleep, low mood and low appetite. He has recently been commenced on sertraline. He has perfectionistic tendencies. He is still sensitive to noise and can only tolerate certain types of music- in particular calm, soothing classical music. He finds it easier to relate to adults than his peers, and finds it difficult to maintain peer friendships.

General health: He has been generally healthy. He was diagnosed with atypical absence seizures aged 8, treated with Sodium Valproate. He never had an EEG demonstrating epileptiform activity. His development reportedly improved after instigation of the Epilim. This medication was ceased around the age of 10 and he has had no further episodes.

Non diagnostic investigations included Fragile X PCR (normal, triplet repeat size 30), chromosomal microarray (normal result – no clinically significant change detected (ISCA v2.0 Bluegnome mean effective resolution 200kb), thyroid function (free T4, free T3 and TSH

normal), CK (121, normal), B12 (389, normal), normal full blood count, normal ferritin, urine metabolic screen (organic acid, amino acids (homocysteine not detected) and GAG screen), urine creatine metabolites (normal) *ARX* testing for the two most common mutations in exon 2 (normal).

On examination: He has some facial similarities to his mother, including a long and narrow head, mild retrognathia and malar hypoplasia). He has a high narrow palate and broad single uvula. He is tall and slim with a marfanoid habitus. Aged 17 his height was 189 cm tall (1 cm > 97th centile), with an arm span: height ratio of 1.03 and upper/lower segment of 0.85 (just within normal limits). His spine is straight, and he has a normal chest appearance and cardiorespiratory examination. He has long slender fingers and toes. He has a positive wrist but negative thumb sign. His whole hand length is 21 cm (>97th percentile) and mid palm length 11.5cm (97th percentile). He is not hypermobile (Beighton score 2/9). He has bilateral increased carrying angle of his elbows, but he does not have reduced elbow extension. He has pes planus with hindfoot deformity. His skin is of normal texture and elasticity without striae. He has no neurocutaneous marks. He has no abnormal or focal neurological signs.

Individual 10-2 last assessment aged 5 years:

The pregnancy was relatively uncomplicated, and this individual was delivered at 38 weeks: birth weight 3578 g (50th centile), length 53cm (50th-75th centile) and head circumference 35.5cm (50th-75th centile). Brief oxygen resuscitation was required immediately post-delivery, to which he responded quickly. He was monitored in special care nursery for the first 12 hours. He developed jaundice on the second day of life requiring two days of phototherapy, and was discharged home breast-feeding on his third day. There were some difficulties maintaining breast feeding due to a weak suck, and he was changed to formula feeds at 3 months.

Development: His parents were not concerned with development in the first year of life: he was a placid baby. They were increasingly concerned in the second year with delayed language development, difficulties with peer relationships and some distinctive behaviors, including hand flapping when excited and some unusual interests and fixations. He had delayed expressive and receptive language and variable eye contact. There were no concerns with developmental regression. His first formal psychometric testing aged 3 years 1 month was on the Griffiths Mental Developmental Scales-Extended revised. He was assessed as having moderate global

developmental delay with severe language delay. He did not meet diagnostic criteria for autism, although some autistic features were noted. He was also noted to have deficits in attention. On follow up assessment aged five years he received a formal diagnosis of autism and intellectual disability (Stanford Binet Intelligence Scales-Vth edition consistent with mild intellectual disability; ABAS-2nd edition consistent with moderate deficit). He still had severe language delay and attention issues.

In general, he is described as a happy 'easy going' child, he can play in parallel with peers, and occasionally co-operatively, especially if play is initiated by a friend. He is fascinated by cars, and has started playing some 'role play' scenarios with the cars. He likes playing car games on an iPad. He is able to speak in simple sentences, although with some articulation difficulties. He has difficulties following more complex requests, and sequencing tasks. He dislikes big changes in routine. He tends to flap his hands when excited. He is affectionate and has good eye contact.

General health: His general health has been good: there are no current concerns with his hearing or vision, and no ongoing seizures (he previously had one febrile convulsion). He remains a very picky eater and tends to alternate between constipation and diarrhea. He does not sleep well: he has difficulties in sleep initiation and maintenance and his parents will need to settle him in his own bed several times a night.

Investigations: Fragile X PCR and chromosomal microarray (Bluegnome ISCA v2- effective resolution 200Kb) were normal.

On examination, he has a pointed chin, slightly fleshy upturned ear lobes, a normal palate and uvula and a normal hairline. He looks very similar to his sister who does not carry a *NAA15* variant. He has normal skin with no neurocutaneous marks. He has a normal appearance to his hands and feet, chest and spine and no abnormalities with his external genitalia. Cardiorespiratory examination was normal. There were no focal or abnormal features on neurological examination. His head circumference was between the 25th-50th percentile. At 3 years 1 month his height was between the 75th-90th percentile, weight on the 90th percentile and head circumference on the 50th percentile.

Individual 10-3 (Mother):

Subject had delayed developmental milestones and received a diagnosis of developmental delay prior to school entry. She had significant social anxiety and shyness in early childhood, and although was not diagnosed on the autistic spectrum, she and her mother feel, in retrospect, that she may have fulfilled diagnostic criteria for a broader autistic phenotype in childhood. She had learning difficulties at school, particularly with literacy and left school early, aged 16. After school she has completed a basic certificate in childcare and has worked as a child-care assistant. Her general health is good, and she has no history of seizures. She has a history of depression. She is a full-time mother to three children, and wonderful advocate. There is a history of learning difficulties and depression in her maternal family.

Individual 11 in Family 11:
Event: inheritance unknown *NAA15* frameshift variant c.239_240delAT (p.His80Argfs*17)

Intellectual disability, anxiety and speech delay. No further information provided other than what is in table.

Individual 11A in Family 11A (added during final revision of manuscript, so given the number 11A, as this is another recurrent variant):
This 12 years old boy had a normal pregnancy and birth parameters. At neonatal age there were no remarks. Both his speech and motor delay appeared delayed for which he attended special education from the age of 5 years. An intelligence test at the age of 8 years showed VIQ 87 and PIQ 67 (WISC-III).
At the age of 9 years and 9 months, his TIQ was 54 (VIQ 67 and PIQ<55).
His medical history is unremarkable. He does not show somatic abnormalities or significant facial dysmorphisms.
At the age of 5 years a DNA test of FMR1 and a 2050kSNP array gave normal results.

Individual 12 in Family 12:
Event: *de novo NAA15* nonsense variant c.248G>A (p.Trp83*)

This is a 7 year and 6-month-old ambidextrous handed female with history of Autism Spectrum Disorder, also diagnosed with Tuberous Sclerosis Complex (TSC) and incontinentia pigmenti (IP). Prior sequencing revealed a Arg256Term c.766C>T p.R256* mutation in *IKBKG*, possibly

associated with incontinentia pigmenti. Exome sequencing with GeneDx also revealed a *de novo LOF* variant in *NAA15*. No pathogenic variant in *TSC1/2* was detected.

Other clinical features include avascularization of retina from IP and strabismus. She has hypopigmented lesions, constipation, and food sensory issues. Renal ultrasound in 2011 was normal. During embryogenesis, echocardiography showed abnormality echogenicity in the womb, not felt to be rhabdomyoma. She has a seizure history as well, with 3 lifetime events all within 1.5 hours in March 2015, although EEG telemetry monitoring was normal. MRI brain imaging in April 2017 showed: Stable left periventricular subependydimal nodule and left parietal subcortical/cortical lesion. These findings most likely represent a subcortical hamartomatous lesion as seen in tuberous sclerosis. However, association with other neurocutaneous syndromes such as incontinentia pigmenti is possible. She has regular sensory issues. Some current medications include guanfacine, risperidone, trileptal, and methylphenidate. There is no family history of seizures, mental retardation, hearing loss, still birth, multiple miscarriages, migraine, neurocutaneous syndrome stigmatas, or metabolic disease. Her developmental history included that she turned supine to prone at 6 months, sat at 10 months; crawled at 14 months and walked at 19 months. She is not toilet trained. She spoke her first words at 5 yrs of age. She is repeating or at least attempting to imitate words and pointing. She has approximately 100 words that she uses specifically. She does not relate with peers, but does have fair eye contact. She does not have sleep issues. She is impulsive imresearch subject, short attention span, cannot focus, and some self-injuries. Her birth history is that she was born to a 43-year-old woman via in vitro fertilization- donated egg (father's sperm) born full term at 39 weeks' gestation via c section due to fetal distress and sudden loss of amniotic fluid. Birth weight 2.49 kg. Mother denied any teratogenic exposures to alcohol, tobacco, and nonprescription drugs except for prenatal vitamins and ondansetron. Mother denied any prenatal or neonatal complications. General examination revealed the following: head size- 2/6/15- 50 cms – 40%, Ht 4/26/17 129 cms- 57.8 %, Wt 11/15/16- 52.8 lbs – 38%, Was alert in no distress. but when approached was kicking and yelling. Appears thin. No facial dysmorphism., normocephalic. Gait was noted to be wide base. Extremities showed no cyanosis, no syndactyly, or polydactyly. Skin exam showed large hypopigmented spot on back, no other lesions seen, no swirling pigmentation seen. Nails normal. Abdomen was soft, no mass, no hepatosplenomegaly. Tongue midline. Adequate muscle bulk, tone and strength. And comprehension markedly delayed.

Individual 13 in Family 13: (Children's Hospital of Eastern Ontario, Ottawa, Canada)

Event*: de novo NAA15* nonsense variant c.287dupA (p.Tyr96*)

Proband is a 4-year-old boy referred for assessment in Medical Genetics for global developmental delay, autism, small size, eczema, and mildly low zinc level. Family history is significant for acrodermatitis enteropathica on the paternal side (non-consanguineous) and the proband is a carrier of a known pathogenic mutation in SLC39A4. There is no family history of developmental concerns or autism.

The proband has profound global developmental delay. He crawled at 18 months, walked at 2 years and jumped at 4 years. At 4 years he had no pincer grasp, was not feeding himself, was babbling, but no words (could indicate his wants with few gestures), and was not yet toilet trained. He was diagnosed with an autism spectrum disorder and was attending SK at a school for children with special needs. Behaviorally, he is very active and has issues with safety (following strangers).

He is generally healthy but had onset of severe eczema at 4 months, involving face, arms, legs, and trunk. Onset occurred with introduction of formula, and was treated with hydrocortisone and lubricants which improved symptoms. Subject was admitted to the hospital at 4 months for infection of atopic dermatitis requiring IV antibiotics. He was admitted at 20 months for respiratory syncytial virus (RSV) infection and failure to thrive (poor weight gain, 3rd centile). Solid food was delayed until 2 years because of his allergies which include dairy, eggs, peas, bananas, and lentils. He was born after induction at 37 weeks for growth concerns (IUGR, birth weight was 2210 g, small placenta noted) and admitted to NICU for 3 days for hypoglycemia, following an uncomplicated pregnancy.

Investigations included chromosomal microarray and Fragile X testing, both of which were normal. Head ultrasound indicated that the left lateral ventricle was slightly larger than the right, but both were within normal limits, and head MRI was unremarkable. Subject was examined by audiology and hearing was reported as normal.

At last examination at 4 years of age his height was on the 3rd centile (99 cm), weight on the 25th centile (15.3 kg), and OFC was on the 2nd centile (48 cm). He was nondysmorphic. His general examination was unremarkable. His skin was normal.

Individual 14 in Family 14:

Event: *de novo NAA15* frameshift variant c.420dupA (p.Leu141Thrfs*25)

Subject was 4 years old at last evaluation including physical examination. Delivery was induced

at 42 weeks' gestation after an uneventful pregnancy. Subject had a good start and no congenital anomalies were observed. There were no problems in the neonatal period. Motor development was normal with walking at age 12. Speech was severely delayed and based on tests he was diagnosed with moderate ID and autism. His length is normal and he has no facial dysmorphisms. Seizures have not been observed.

Individual 15 in Family 15 : Subject ID: AU031003 (AGRE, NIMH) - was listed as Subject 3 in Stessman et al. 2017 Nature Genetics paper

Event: *de novo NAA15* frameshift variant c.532_533delCA (p.Gln178Thrfs*5)

Female subject has been diagnosed with ASD. Subject was 12 years at the time of the last assessment. At that time, her adaptive skills were found to be very low, with an age equivalent of 1 year 10 months on the Vineland (composite score = 21). A valid nonverbal IQ score was not generated, as the participant was considered "untestable" on a measure of IQ as well as on a measure of receptive language (PPVT). She shows a relative strength in the area of motor skills, which were are estimated to be in the low average range, though she has fine motor difficulties. Subject walked at age 14 months, said her first word at 24 months, and first combined words at age 48 months. She was toilet trained at age 5.5 years.

The subject demonstrates limited verbal skills, uses stereotyped language, and has unusual intonation. She has significant nonverbal communication deficits including uses limited gestures and facial expressions to communicate. She engages in stereotyped motor behavior including hand flapping, jumping, and toe walking. The subject has sensory differences including sensitivity to sounds and seeking of tactile stimulation (e.g., pressure). Behavioral difficulties began at age 2 years, and the subject has been on several types of medication to address irritability (e.g., SSRI, Risperdal, amantadine, Depakote). Sleep problems first manifest at age 2 years with difficulty falling asleep and interrupted sleep. Subject also has a history of Pica since age 2 years.

The subject was born at 40 weeks' gestation via vaginal delivery following an uncomplicated pregnancy. Subject had hyperbilirubinemia. Mother was age 38 years and father was age 42 years at time of conception. Subject weighed 8.5 pounds at birth. Seizure history was denied. Subject had strabismus that resolved spontaneously. Skin abnormalities were noted (unknown type) as well as frequent ear infections.

Individual 16 in Family 16: Subject ID: 115149 (Antwerp) – was listed as Subject 4 in Stessman et al. 2017 Nature Genetics paper

Event: *NAA15* nonsense mutation c.868G>T (p.Gly290*)

Not present in mother, father unavailable for genetic testing, thus inheritance unknown.

Subject is a 7-year-old boy with global developmental delay. Speech and language are also delayed (level of 2-2.5 years old at age 4) with inadequate communication and contact and autistic traits. He also has cognitive delay: SON-IQ 72, SON-RS 79, SON-PS 72: harmonic IQ profile.

Subject was delivered at 41 gestational weeks with a birth weight of 3280 g and length of 52 cm. He began crawling at 6.5 months and walking independently at 13 months. Toilet training was delayed. He experiences behavioral problems at school, and had a transient period of hair-pulling at night. He has normal hearing (BERA). Mother followed special education.

Individual 17 in Family 17:

Event: *de novo* intronic splice variant, c.908-2A>G

No further information provided other than what is in table.

Individual 18 in Family 18:

Event: *de novo NAA15* nonsense variant c.913A>T(p.Lys305*)

No further information provided other than what is in table.

Individual 19 in Family 19; Subject ID: 1-00455; PMID: 23665959

Event: *de novo NAA15* frameshift variant c.1009_1012delGAAA(p.Glu337Argfs*5)

Subject is a 7-year-old male in 2016 with complex congenital heart disease. He is diagnosed with developmental delays, and requires special education. He has more difficulty learning reading compared with math. He is hypotonic and has a speech delay.

The proband's major cardiac abnormalities include heterotaxy, asplenia syndrome, single ventricle, and congenital heart defect. Subject also has pulmonary artery stenosis, dextrocardia, tricuspid atresia, double outlet right ventricle, hypoplastic right ventricle, total anomalous pulmonary venous return to confluence behind right-sided left atrium, left superior vena cava, and intestinal malrotation. He has had the LADD procedure for the intestinal malrotation.

He also presents with hepatic veins, a right aortic arch, and dextrocardia. In April 2015, he was diagnosed with plastic bronchitis (2-3 episodes in 2016 and hospitalized for 1 of the 3 episodes). He also has a midline liver. Subject has a history of femoral vein thrombosis. He is also asplenic. Family history is benign.

Individual 20 in Family 20:(contributed by Michael Parker), DDD study, DECIPHER ID: 260821
Event: *de novo NAA15* frameshift variant c.1009_1012delGAAA(p.Glu337Argfs*5)

12-year-old girl, parents first-cousins; two brothers (one with degree of ID). Delivered via caesarean at term, with a birth weight of 2.41kg. Feeding difficulties initially. Admission for eczema at age two years and remains problematic. First walked at age 18 months and requires extra help in mainstream school. Somewhat obese and no particular dysmorphisms.

Individual 21, in Family 21:(Baylor)
Event: *de novo NAA15* c.1087+2T>C (intronic) variant.

Subject is an 8 year 3-month-old female with long-standing and significant developmental delay. She was delivered at 34 gestational weeks weighing 4 pounds with a fraternal twin sister through repeat C-section. At 3 months of age parents noted that she did not wake up to feed and that she was floppier compared to her twin sister. She got RSV at 10 months of age and at this time she quit taking sufficient oral intake. She had emesis and poor oral intake and required NG-tube placement, and subsequently a gastrostomy tube due to poor oral intake. She rolled over at 10 months, sat up at 11-12 months and walked at 17 months. She spoke her first word at 2.5 years old and talked with sentences at 4 years old. At 13 months she developed seizures with no clear etiology, and she has been seizure-free for the past 2.5 years on Keppra. Subject has esotropia, wears glasses and is concerned for worsening vision. Subject also had atypical migraines, mild hypoglycemia and mild growth hormone deficiency. Subject has intellectual disability and she receives PT, OT & ST regularly. On her last visit at 8y3mo old, she was able to read, knew some sight words and wrote her name. New symptoms found in the las visit are abdominal migraines, manifesting as constipation, emesis, abdominal distention and intestinal dysmotility. The fraternal twin sister is completely on-target for her development. There is an older brother who had some dysphagia and poor growth that resolved; he is cognitively ahead of his peers. Father had retinal detachment and hypotonia in infancy.

Individuals 22-1 and 22-2, in Family 22: (contributed by Mathilde Nizon)
Event: inherited *NAA15* stop gain variant c.1134C>A(p.Tyr378*)

Subject is a 4 years old 1/2 male with moderate global development delay. He was delivered at 38 gestational weeks with a birth weight of 3250g, height of 51 cm and OFC of 34 cm. He had hypotonia and began sitting at 10 months, walking at 24 months. He also experienced speech and language delay, saying 5-10 words at 3 years old, and associating words at 4 years 1/2. Subject has probably mild intellectual deficiency as well as behavioral problems including angry behavior, mild hyperactivity and attention deficit, and mild autistic features. Subject has small mouth and low set ears. Subject's father who harbored the same mutation was mildly affected. He walked at 2 years old and experienced speech delay with sentences at 4 years. He has very mild autistic features and shyness. Subject's father does not have intellectual disability but probably has a low IQ. He also has hypermetropia and strabismus. His weight is 80 Kg, height 175 cm and OFC 60,5 cm. Dysmorphic features are mild and similar to his son. There is no evidence of any cardiac phenotype in either the father or his son.

Individual 23 in Family 23: Subject ID: 105005 (Antwerp) – was listed as Subject 7 in Stessman et al. 2017 Nature Genetics paper.
Event: de novo NAA15 nonsense variant c.1695T>A(p.Tyr565*)

Subject is a 6 year old boy delivered at 38w1d weighing 2905 g(L 47 cm and OFC 35 cm). The child was sitting at 8.5 months, spoke his first words at 14 months, and was walking at 19 months. MRI shows no brain abnormalities. Subject has unilateral cryptorchidia, ptosis of the left eye, mild hypertelorism, flat philtrum, mild pectus carinatum upper side and excavatum below, and low-set ears. He had relative macrocephaly at age 2y4m: L84cm (P3), Weight 12,4kg (P10-25), OFC 51cm (P75-90). Subject has global developmental delay, speech and language delay, and possible autism spectrum disorder including aggressive behavior and traits of Noonan syndrome. He has healthy parents and two healthy older brothers. One younger brother has autism spectrum disorder.

Individual 24 in Family 24:
No further information provided other than what is in table.

Individual 25, in Family 25: (Mayo)
Event: de novo *NAA15* LOF variant c.1798_1801delAGAA (p.Arg600Glufs*8).

The Subject is a 12.5 years old male born at 39 weeks' gestation at 7 pounds 2 ounces via vacuum extraction. He walked at 12 months and had delayed fine motor milestones. He had delayed speech and language development, having about a 20-word vocabulary at 4 years old. He received a diagnosis of autism when he was about 3 years old. The subject developed seizures at age 4 which have been attempted to be controlled by multiple medications and he has a deep brain stimulator. This Subject has additional features which include microcephaly, constipation and extra skin between third and fourth fingers. Development is normal in Subject's brother and sister (full siblings), who has anxiety and autism. Father has ADHD/ADD and reported to have Asperger like tendencies.

Individual 26 in Family 26: Subject ID : 426396 (Leiden) – WES in diagnostics
No further information provided other than what is in table.

Individual 27 in Family 27: DECIPHER ID: 282455, DDD study, contributed by Ruth Armstrong
Event: A *de novo NAA15* variant c.1906G>T(p.Gly636*) was identified through the DDD study.

Referred at 4.5 years with global developmental delay, anteriorly placed anus, bilateral talipes, sacral dimple and dislocation of her right patella. Second child of healthy, unrelated Caucasian parents. Talipes were noted at 20 weeks' gestation. She was born at term by emergency caesarean section for meconium stained liquor. Her birth weight was 2353g (< 0.4th centile). She required some resuscitation but did not need admission to special care. There were initial difficulties with feeding, growth and persistent diarrhea. She required physiotherapy, casting and splints for her talipes and surgery for her dislocated patellae.
She walked between 18 months and 2 years. She required speech and language therapy for speech delay. She exhibited some unusual mannerisms including rocking and behavioral difficulties.
At 8 years she had joint hypermobility, constipation, difficulty sleeping and premature adrenarche. She was able to hold a simple conversation. She had initially attended mainstream primary school with 20 hours of additional support but later transferred to a specialist school. On examination her head circumference was 51.2 cm (2-9Th centile).

Dysmorphic features included a prominent nasal root, straight eyebrows and a squared chin. She also had irregular teeth, a sacral dimple and hypoplastic fifth toenails.
Echocardiography, microarray and testing for Fragile X syndrome was normal.

Individuals 28-1, 28-2, and 28-3 in Family 28: Subject ID: D2.11.05195 (Leiden) – was listed as Subject 8 in Stessman et al. 2017 Nature Genetics paper (Eichler panel).
Event: maternal inheritance *NAA15* frameshift variant c.1988delC (p.Pro663Argfs*2)

Female subject was born in 2005 and presented with ASD, IQ 50, and hyperactivity, yet no obvious dysmorphisms. Mother has followed special education, and both the mother and the sister of the proband are affected, with this variant being maternally inherited, as confirmed by Sanger sequencing.
Individual 28-2 in Family 28: sister of proband
Individual 28-3 in Family 28: mother of proband. The mother has partial syndactyly in one hand.

Individual 29 in Family 29: Subject ID: 106663 (Antwerp) – was listed as Subject 9 in Stessman et al. 2017 Nature Genetics paper.
Event: *de novo NAA15* nonsense variant c.2086A>T(p.Lys696*)

Subject is a 31 year old male with intellectual disability, ASD, and no facial dysmorphism except 3 pre-auricular tags on the right side. His current biometrics are: L 165 cm, W 75-80 kg, and OFC 54.3cm. His parents are consanguineous.

Individual 30 in Family 30: contributed by DDD study and Derek Lim
DECIPHER ID: 267517
Event: de novo *NAA15* nonsense variant c.2300C>A(p.Ser767*)

Female
Twin 1 of 2 – non-identical
Born at 36 weeks gestation, birth weight 2.16kg. Developed severe eczema in first few weeks of life and subsequently developed multiple food allergies. No family history of atopy. Developmental delay noted in comparison to her twin sister. Sat unaided at 20 months. Walked unaided at 3 years. Hypotonia in first two years of life. Speech delay with babbling at 2.5 years.

First words after age 3 years. Now speaking in 2-3-word sentences age 7 years. Had stereotypical behavior – rocking and hitting of chest which resolved just before age 7 years. Growth all on the 9th centile which is similar to unaffected twin sister. No history of seizures and normal brain MRI scan. Making developmental progress in special educational needs school with moderate intellectual disability. Assessed as equivalent developmental age of a 3-4 year old. No regression. Mild dysmorphic features with straight eyebrows and depressed nasal bridge (different from other family members). No family history of note. Three unaffected siblings including twin sister.
Normal karyotype and aCGH.
Recruited into DDD study. De-novo heterozygous NAA15 nonsense mutation identified. In addition a MYH6 de-novo heterozygous missense substitution identified thought to be benign.

Individual 31 in Family 31
Event: de novo *NAA15* nonsense variant c.2322C>G(p.Tyr774*)
Negative family history. Contributed by HudsonAlpha Institute for Biotechnology, Alabama. No further information provided other than what is in table.

Individual 32 in Family 32: Subject ID: 3211 (Troina) – was listed as Subject 13 in Stessman et al. 2017 Nature Genetics paper.
Event: inheritance unknown NAA15 stop-gained variant c.2389A>T(p.Arg797*)

Boy born in 2001 and seen for a single time at 6 years, diagnosed as NOS ID (ICD-9-CM Diagnosis Code 319) and NOS Pervasive Developmental Disorder (ICD-9-CM Diagnosis Code 299.9). Epilepsy, delayed speech, and ID are shown by some relatives (one maternal uncle, two 1st degree cousins on the paternal line, one 2nd degree cousin on the maternal line). On the paternal line both grandparents showed endocrine disorders (Addison disease in the grandmother and Thyroid disease in the grandfather). He was delivered by ceasarean for early placental detachment after 36 weeks of uneventful pregnancy. Birth asphyxia and cyanosis required oxygen supplementation. The Apgar score was 6 at 1' and 8 at 5'. Weight at birth was 2600 grams. Weight and length thrived normally in the first months of life. Psychomotor development was delayed (independent sitting at 8 months, walking at 18 months, first words at 30-36, bladder control at 3 years, bowel control at 5 years). Weight and height at 6 years were

on the 97th and 75th percentile, respectively (28 kg and 122 cm). Brain MRI and EEG were normal. Agilent 180K Array-CGH was normal.

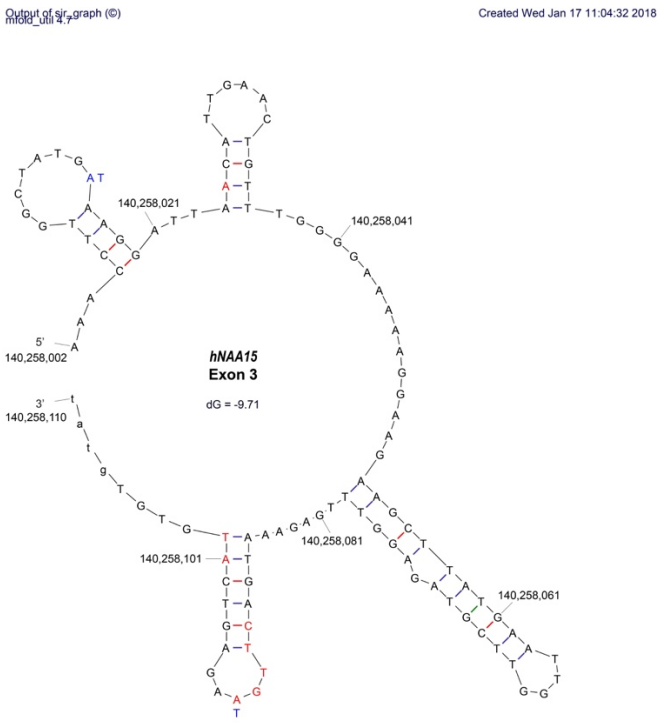
CNV *de novo* deletion case, 5.2 Mb, chr4:137504418 - 142696249, Decipher number: 332240

Eldest child of unrelated parents. No significant family history.
Presented at age 3 with developmental delay. He had walked at 18 months, and had only a single word by age two. The pregnancy and birth history was unremarkable, apart from difficulty feeding. Routine investigations were normal, including CT brain scan. Brain MRI done as an adult was also normal. He attended special school for children with moderate learning difficulties from the beginning.
As an adult, he was able to live alone with support. He had periods of employment in a supermarket. He enjoyed voluntary work for charities. He developed testicular carcinoma at age 25, treated with chemotherapy and resection. The pathology was choriocarcinoma with teratoma.
He developed type II diabetes at age 31, treated with metformin. He was also hypertensive.
Micro array undertaken on review at age 34 demonstrated a de novo 4q deletion; 4q28.3-q31.1.
At age 35, he was found dead in bed, the day after completing a run for charity. A full autopsy was performed, which did not find any gross abnormalities in the brain, heart, or other organs. Specialist examination of the heart found no morphological abnormality, excluding myocardial infarction as the cause of death.
A molecular autopsy is awaited.

Figure S1

Figure S1. Facial features of individual 20 and individual 29. No obvious and common patterns of dysmorphology were noted.

Figure S2
A)



B)

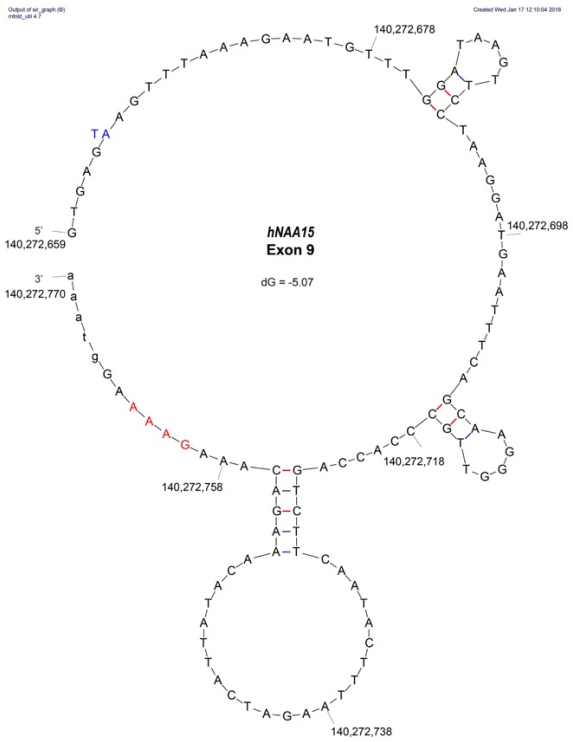


Figure S2. Computational prediction of quasi-palindromic structures. Deleted bases are highlighted in red and SNV bases in blue. A) Exon 3. All variant bases are found within

predicted secondary palindromic structures. The two recurrent deletions are both in the fourth quasipalindromic structure, which is distal to a larger quasipalindromic structure. B) Exon 9. The recurrent mutation c.1009_1012delGAAA in exon 9 lies just distal to a quasipalindromic structure.

Figure S3

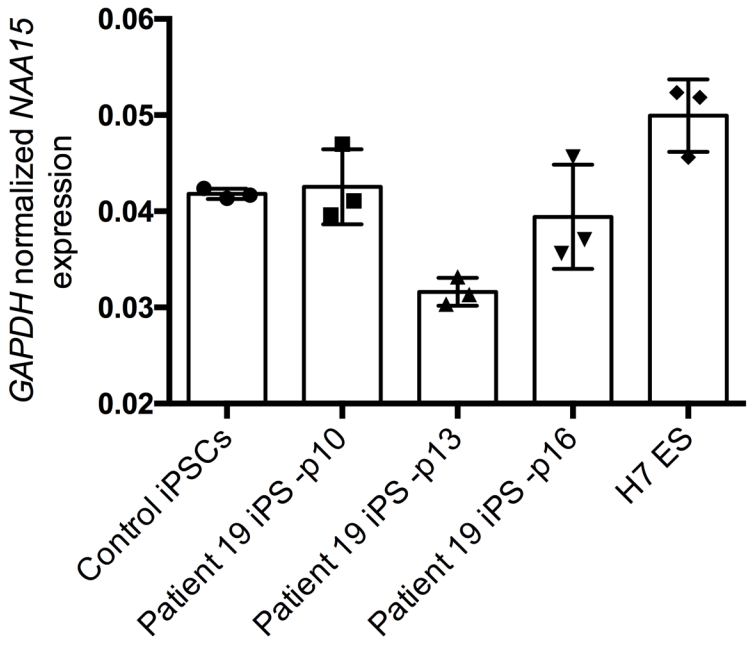
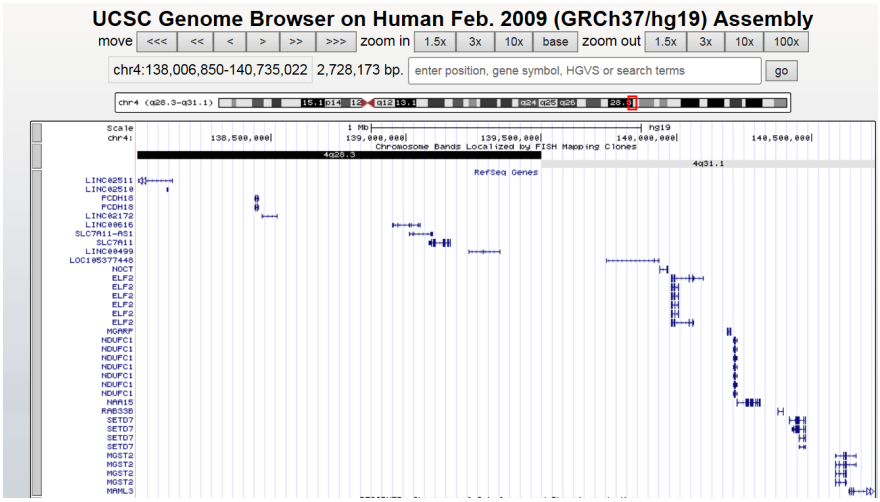


Figure S3. Taqman assay from RNA from induced pluripotent stem cell (iPS) line from Individual 19 (c.1009_1012delGAAA). This Taqman assay showed variability in *NAA15* expression (normalized against *GAPDH*) from 3 different passages of the iPS line from Individual 19, as compared to a control iPS line and a control human embryonic stem cell (hESC) line. Error bars are standard deviation.

Figure S4
A)



B)

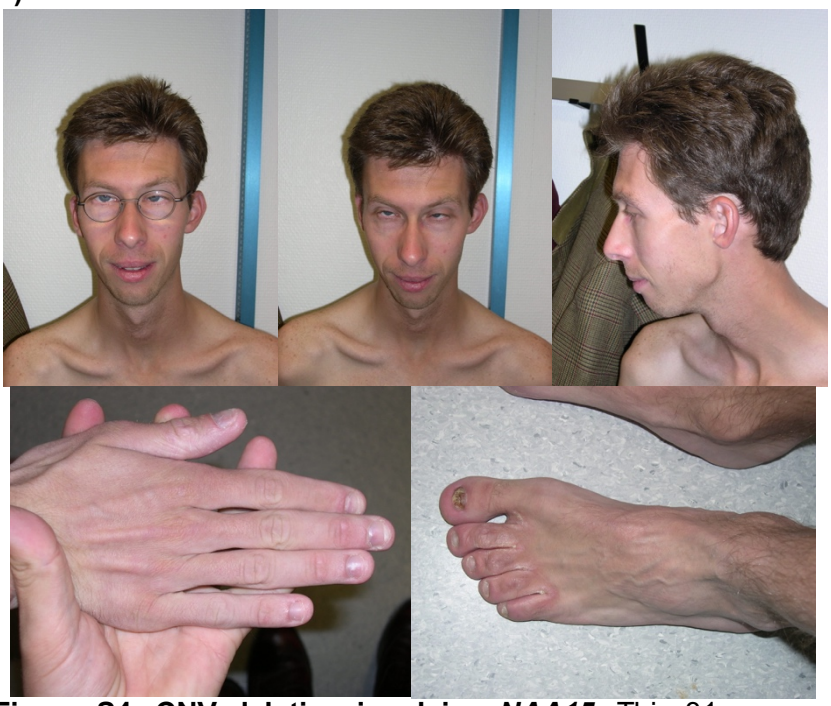


Figure S4: CNV deletion involving *NAA15*. This 31 years old man with a mild intellectual disability was born after 42 weeks of gestation with a birth weight of 2650 gr (- 2SD). In the pregnancy growth retardation was ascertained. At the age of 8 months a strabisum convergens was noted. His motor development was delayed, sitting at 8 months, walking at 16 months. From the age of 5 years he attended special education. At the age of 12 years a kidney stone was removed. He had poor vision due to a cerebral visual impairment. MRI of the brain revealed a small cavernous haemangioma at the left temporal in the radiation optica. At physical examination at 26 years he had a height of 173 cm (10th centile), head circumference of 55 cm (

10th centile) and weight of 52 kg (10th centile). He had midfacial hypoplasia, upward slanting palpebral fissures, strabismus convergens left, mild ptosis and small but normally shaped ears, retro/micrognathia and a long neck. He had long hands and fingers (>98th centile), narrow feet and a lumbal lordosis. Metabolic studies did not show any abnormalities. However, SNP array analysis revealed a de novo 2.73 Mb deletion on chromosome 4 including NAA15: 46,XY.arr 4q28.3q31.1(138,006,851-140,735,023)x1 dn (Hg19). This de novo microdeletion was regarded as likely contributing to the combination of his problems, so a decision was made to not proceed with additional testing, such as whole exome sequencing.

Figure S5

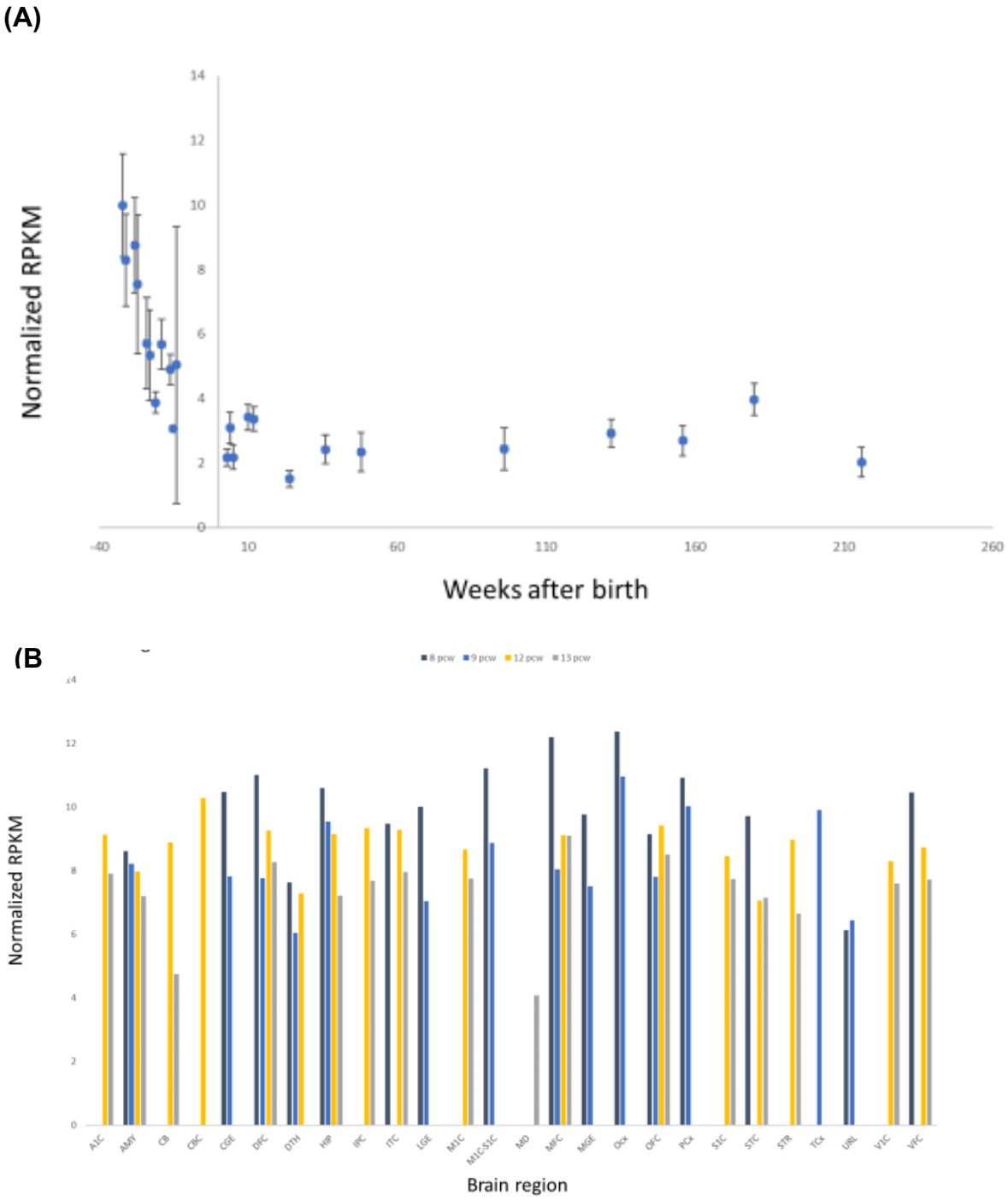


Figure S5. Patterns of *NAA15* expression in the developing human brain.

(A) Scatterplot shows the expression of *NAA15* by RNA-Seq analysis from the BrainSpan: Atlas of the Developing Human Brain over the course of development from eight post-conception weeks (pcw) to 18 years of age. Data points represent the mean expression and standard deviation across all brain tissues and samples collected at each time point. Normalized reads

per kilobase of exon model per million mapped reads (RPKM) is plotted on the y-axis and age in months on the x-axis assuming a 40-week gestational period. (B) Graph shows the mean expression of *NAA15* by brain region at 8 pcw (black), 9 pcw (blue), 12 pcw (gold), and 13 pcw (grey) which represent the highest levels of expression over the course of development. A1C = primary auditory cortex (core); AMY = amygdaloid complex; CB = cerebellum; CBC = cerebellar cortex; CGE = caudal ganglionic eminence; DFC = dorsolateral prefrontal cortex; DTH = dorsal thalamus; HIP = hippocampus (hippocampal formation); IPC = posteroventral (inferior) parietal cortex; ITC = inferolateral temporal cortex (area TEv, area 20); LGE = lateral ganglionic eminence; M1C = primary motor cortex (area M1, area 4); M1C-S1C = primary motor-sensory cortex (samples); MD = mediodorsal nucleus of thalamus; MFC = anterior (rostral) cingulate (medial prefrontal) cortex; MGE = medial ganglionic eminence; Ocx = occipital neocortex; OFC = orbital frontal cortex; PCx = parietal neocortex; S1C = primary somatosensory cortex (area S1, areas 3,1,2); STC = posterior (caudal) superior temporal cortex (area 22c); STR = striatum; TCx = temporal neocortex; URL = upper (rostral) rhombic lip; V1C = primary visual cortex (striate cortex, area V1/17); VFC = ventrolateral prefrontal cortex.

Table S1: Summary of clinical presentations and molecular findings in individuals with *NAA15* LGD variants

Individual ID	Age[a]	Sex	Variant	Inheritance	Main clinical presentations					Reference
					Motor abilities	ID	Verbal abilities	ASD/ behavioral issue	Others	
1	22	F	c.154A>T (p.Lys52*)	unknown	global developmental delay	moderate	N/A	ASD	mildly dysmorphic, cutis marmorata	1
2	17	M	c.163delA (p.Thr55Hisfs*2)	*de novo*	mild ataxia	moderate	difficulty speech	ASD	abnormal MRI, dysmorphic, hypertonia, atrial ectopic (multifocal) tachycardia	This study
3	6	M	c.228_232delCTTGA (p.Asp76Glufs*20)	*de novo*	abnormality of movement	moderate ID	N/A	N/A	VSD with cardiac repair in infancy	This study
4	17	M	c.228_232delCTTGA (p.Asp76Glufs*20)	*de novo*	normal	N/A	delayed	ASD		1
5	7	M	c.232A>T (p.Lys78*)	unknown	motor delay	N/A	sign language	ASD	dysmorphic, tall stature	1
6	6	F	c.239_240delAT (p.His80Argfs*17)	*de novo*	global developmental delay	moderate	absent	ASD	mildly dysmorphic	This study
7	1.5	F	c.239_240delAT (p.His80Argfs*17)	*de novo*	mild delay	mild delay	delayed	ASD	congenital diaphragmatic hernia, lung hypoplasia, plagiocephaly and mild torticollis (resolved)	3
8	8	F	c.239_240delAT (p.His80Argfs*17)	unknown	motor delay	moderate	delayed	ASD		1
9	13	F	c.239_240delAT (p.His80Argfs*17)	unknown	global developmental delay	global developmental delay	absent	ASD		1
10-1	17	M	c.239_240delAT	maternal	motor delay	mild	delayed	ASD	mild	This study

		(son)	(p.His80Argfs*17)						dysmorphic, tall and slim, marfanoid habitus	
10-2	5	M (son)	c.239_240delAT (p.His80Argfs*17)	maternal	motor delay	moderate	delayed	ASD	mild dysmorphic; attention deficit, one febrile convulsion	This study
10-3	40	F (Mother)	c.239_240delAT (p.His80Argfs*17)	unknown	global developmental delay	learning difficulties	delayed	social anxiety; depression	mild dysmorphic	This study
11	11	M	c.239_240delAT (p.His80Argfs*17)	unknown	N/A	ID	delayed	anxiety	marked hypersomnolence in early years, upper airway obstruction, central and obstructive sleep apnea	1
11A[b]	12	M	c.239_240delAT (p.His80Argfs*17)	*de novo*	motor delay	moderate ID	delayed	ADHD	none	This study
12	7.5	F	c.248G>A (p.Trp83*)	*de novo*	N/A	impulsive impatient, short attention span, cannot focus, some self-injuries	delayed	ASD	incontinentia pigmenti, broad-based gait, tuberous sclerosis complex, constipation	This study
13	4	M	c.287dupA (p.Tyr96*)	*de novo*	motor delay	profound global developmental delay	delay; no speech	ASD	eczema	This study
14	4	M	c.420dupA (p.Leu141Thrfs*25)	*de novo*	N/A	moderate	delayed	N/A		This study
15	12	F	c.532_533delCA (p.Gln178Thrfs*5)	*de novo*	fine motor difficulties	ID, adaptive functioning in the very low range	delayed	ASD, behavior problems	skin abnormalities, sleep problems, sensory	1

									differences and auditory sensitivities	
16	7	M	c.868G>T (p.Gly290*)	unknown	Gross and fine motor difficulties.	ID	delayed	ASD		1
17	18	M	c.908-2A>G	*de novo*	normal gross, poor fine motor	ID	basic speech	ASD	mild dysmorphic, marfanoid habitus and striae	1
18	10	F	c.913A>T (p.Lys305*)	*de novo*	motor delay, fine motor difficulties	Mild/moderate ID	Delayed expressive language	normal	mild dysmorphic	This study
19	8	M	c.1009_1012delGAAA (p.Glu337Argfs*5)	*de novo*	motor delay	learning issues	speech issues	normal	congenital heart disease, Midline liver, hypotonic	2
20	11	F	c.1009_1012delGAAA (p.Glu337Argfs*5)	*de novo*	motor delay	N/A	N/A	N/A	eczema, obesity	This study
21	8 1/4	F	c.1087+2T>C	*de novo*	motor delay	ID	delayed	normal	failure to thrive, history of seizures, hypotonia, esotropia	This study
22-1	4 1/2	M	c.1134C>A (p.Tyr378*)	paternal	motor delay	moderate ID	delayed	hetero-aggressiveness, hyperactivity, mild ASD	mild dysmorphic, hypotonia	This study
22-2	31	M (father)	c.1134C>A (p.Tyr378*)	unknown	N/A	low IQ	speech delay	mild autistic features and shyness	hypermetropia and strabismus, dysmorphic features are mild and similar to his son.	This study

23	2 1/2	M	c.1695T>A (p.Tyr565*)	*de novo*	global developmental delay	global developmental delay	delayed	aggressive behavior traits of Noonan syndrome, possible ASD	dysmorphic, mild pectus carinatum upper side and excavatum below	1
24	14	M	c.1795_1798delCAAA (p.Gln599Glufs*9)	unknown	fine motor coordination problems	ID	mild speech delay	N/A		1
25	12 1/2	M	c.1798_1801delAGAA (p.Arg600Glufs*8)	*de novo*	global developmental delay, motor delay	N/A	speech delay	ASD	microcephaly, seizures	This study
26	18	F	c.1841delA (p.Asn614Metfs*22)	*de novo*	global developmental delay	moderate ID	global developmental delay	N/A	hypertelorism, flat broad nose bridge	This study
27	8	F	c.1906G>T (p.Gly636*)	*de novo*	motor delay	moderate to severe ID	speech delay	concerns for ASD	patellar dislocation, sacral dimple, talipes	This study
28-1	15	F (daughter)	c.1988delC (p.Pro663Argfs*2)	maternal	N/A	moderate ID	speech delay	ASD	hypermobility	1
28-2	12	F (daughter)	c.1988delC (p.Pro663Argfs*2)	maternal	motor delay	moderate ID	N/A	ASD	N/A	This study
28-3	45	F (mother)	c.1988delC (p.Pro663Argfs*2)	unknown	N/A	mild ID	normal	normal	N/A	This study
29	31	M	c.2086A>T (p.Lys696*)	*de novo*	N/A	ID	N/A	ASD	mild dysmorphic	1
30	7	F	c.2300C>A (p.Ser767*)	*de novo*	motor delay	moderate ID	speech delay	stereotypy (stopped at age 7)	neonatal hypotonia, mild dysmorphic	This study
31	2	M	c.2322C>G (p.Tyr774*)	*de novo*	motor delay	moderate ID	speech delay	impulsive and obsessive	mild dysmorphic, prematurity	This study
32	6	M	c.2389A>T (p.Arg797*)	unknown	motor delay	ID	speech delay	ASD		1

[a]Age: age at last assessment (years)

[b]Added during manuscript final revision, as another recurrent variant.

ASD = autism spectrum disorder; ID = intellectual disability; N/A = not available

The accession number for the *NAA15* sequence reported in this paper is RefSeq: NM_057175.4 (on GRCH37/hg19 assembly).

Reference: [1]Stessman HA et al. Nat Genetics. 2017. [2]Zaidi S et al. Nature. 2013 and. [3]Longoni M. Hum Genet. 2017.

Table S2-S6: see excel file.

Table S7: Primers used in yeast studies

ID	Primer	Oligo sequence
oTA816	ins HA-tag hNAA15 F	5'-gtgcccgactacgcatccaccatg**ATG**CCGGCCGTGAGCCTC-3'
oTA817	ins HA-tag hNAA15 R	5'-gtcgtaggggtaggccatggtctcCGTTGGATCCGTCGAAACTAAGTTC TTG-3'
oTA781	NAA15 c.163delA_p. T55fs F	5'-CATTGAACTGTTTGGGGAAAAAG-3'
oTA782	NAA15 c.163delA_p. T55fs R	5'-TAATCCTTTCATAGCCAAGG-3'
oTA800	NAA15 c.A913T_p.K 305X F	5-' ATCTGGTGAG**t**AGTTTAAAGAATG-3'
oTA801	NAA15 c.A913T_p.K 305X R	5-'AAAAAGTTTAACGGCAGC-3'
oyTA263	NAA15 hphNT1 F	5'CTTGTTTAAGACAAATACCATTGAGGAAGGCGATTGACCCTAACGA AGT**ATG**cgtacgctgcaggtcgac-3'
oyTA264	NAA15 hphNT1 R	5'-CATAAATTAAGTAAGAGTTAATTGACACATTGAGGAGTTGCA GG**CTA**atcgatgaattcgagctcg-3'

Table S8: Plasmids used in yeast studies

ID	Plasmid	Reference
pTA53	pBEVY-U	Miller et al (1998) Nucleic Acids Res[3]
pTA56	pBEVY-U-hNAA10-hNAA15	Arnesen et al (2009) PNAS[1]
pTA586	pBEVY-U-hNAA10-HA-hNAA15	This study
pTA635	pBEVY-U-hNAA10-HA-NAA15_c.163delA_p.T55fs	This study
pTA637	pBEVY-U-hNAA10-HA-NAA15_c.A913T_p.K305X	This study

Table S9: Yeast strains used in this study

yTA ID	Genotype	Description
	BY4742; MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; [pTA53(pBEVY-U)]	Control/Wt strain, pBEVU-U-empty
1022	BY4742; MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; *YHR013c::kanMX4*, *YDL040C::hphNT1;* [pTA53(pBEVY-U)]	*naa10*Δ*, naa15*Δ, pBEVU-U-empty
1042	BY4742; MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; *YHR013c::kanMX4*, *YDL040C::hphNT1;* [pTA586(pBEVY-U-hNAA10-HA-hNAA15)],	*naa10*Δ*, naa15*Δ*,* hNAA10, HA-hNAA15
1044	BY4742; MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; *YHR013c::kanMX4*, *YDL040C::hphNT1;* [pTA635(pBEVY-U-hNAA10-HA-NAA15_c.163delA_p.T55fs)],	*naa10*Δ*, naa15*Δ*,* hNAA10, HA-hNAA15 T55fs
1050	BY4742; MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0; *YHR013c::kanMX4*, *YDL040C::hphNT1;* [pTA637(pBEVY-U-hNAA10-HA-NAA15_c.A913T_p.K305X)]	*naa10*Δ*, naa15*Δ, hNAA10, HA-hNAA15 K305X

Supplemental Materials and Methods

Plasmid Construction

Plasmids and primers are listed in **Table S7** and **S8**. The yeast expression vector pBEVY-U-hNatA was generated by inserting hNAA10 after the ADH1 promoter using XmaI/EcoRI sites and hNAA15 after the GPD promoter using the BamHI/SalI sites[1]. A hemagglutinin (HA) tag was incorporated N-terminally of hNAA15 using the Q5 Site-Directed Mutagenesis Kit (NEB) with the primer pair oTA816 and oTA817. The resulting plasmid was used to construct the subject-specific c.163delA; p.T55_fs and c.913A>T; p.K305_stop NAA15 target mutations using site-directed mutagenesis.

Yeast Strains

All yeast strains used in this study are derivatives of BY4742 (MATα; ura3Δ0; leu2Δ0; his3Δ1; lys2Δ0) and are listed in **Table S9**. The naa10Δ, naa15Δ double deletion strain was constructed by replacing NAA15 with the HygR gene in a naa10Δ strain (EUROSCARF, accno. Y10976) via PCR-based homologous recombination. The NAA15-specific hphNT1 deletion cassette was amplified from pFA6a-hphNT1[2] using the primer pair oyTA263 and oyTA264. Deletion of NAA15 was confirmed by PCR. Human NatA variants were expressed from the bidirectional pBEVY-U expression vector carrying URA3 for selection[3].

Serial dilution spot assay

Yeast strains were grown in SD-URA medium [0.67% (w/v) yeast nitrogen base without amino acids and ammonium sulfate; 0.84 % (w/v) ammonium sulfate 0.17% (w/v) yeast synthetic drop out media supplements without uracil; and 3% (w/v) glucose] to early log phase (OD600 0.8-1.0), washed twice in MilliQ water, and adjusted to 10 OD600/ml. Ten-fold serial dilutions were manually spotted (2 µl) onto standard yeast extract-peptone-dextrose (YPD) agar plates (Sigma, Y1500). The plates were
incubated at 30°C or 37°C and imaged after 3 days and 6 days, respectively, using the Gel Doc EZ imaging system from Bio-Rad.

Protein extraction and immunoblot analysis

Protein extraction was performed according to Kushnirov with minor adjustments[4]. Cells in exponential phase were harvested at 5,000 x g for 10 min at 4°C. The cell pellet was washed twice in cold MilliQ water and resuspended in 0.1 M NaOH. Following 5 min alkaline treatment at RT, the cell suspension was pelleted at 10,000 x g for 1 min., resuspended in SDS sample

buffer (60 mM Tris-HCl, pH 6.8, 5% glycerol, 2% SDS, 100 mM DTT, and bromophenol blue), and incubated for 10 min at 90°C. To remove cell debris, the cell suspension was centrifuged at 10,000 x g for 30 sec. 0.25 mg of protein extracts were analyzed by immunoblotting. Primary antibodies used were rabbit pAbs anti-hNAA10[5], rabbit pAbs anti-HA (Abcam, ab9110), and mouse mAbs anti-beta-actin (Abcam, 8224). Secondary antibodies used were ECL anti-Rabbit. IgG HRP-linked whole Ab from donkey (GE Healthcare, NA934) and ECL anti-Mouse IgG, HRP-linked whole Ab from sheep (GE Healthcare, NA931).

Sequencing of the families

This study was performed in accordance with protocols approved by the institutional review board of Leiden University Medical Center (family 1, 26 and 28), ErasmusMC University Medical Center (family 2), Children's Hospital of Eastern Ontario Family 13), Baylor College of Medicine (family 17, 21, 22 and 25), Columbia University (family 19), GeneDx (family 6 and 12), Massachusetts General Hospital/Boston Children's Hospital (family 7), Adelaide and Sydney GoLD Service (family 5, 9, 10, 11, 17 and 24) and other participating institutions.

Generation and validation of research subject-derived cell lines

A LCL was established from Individual 10-1 using standard EBV transformation, and iPS lines was generated from Individual 19 as below.

Description of control cell lines used

H7 ES line: this is a commercial line derived form a female human embryo whose disease is not reported (https://www.wicell.org/home/stem-cell-lines/catalog-of-stem-cell-lines/wa07.cmsx). Passage number unknown.

The control iPS line was derived in separate experiments from a female, Asian, age between 30-45, collected for RNA at passage 16.

The control LCLs are as follows:

control 1, made from a blood sample from a 27 year old Caucasian male, passage number unknown, but the cell line has been repeatedly passaged over many years.

control 2, made from a blood sample from a 41 year old Caucasian male, passage number unknown, but the cell line has been repeatedly passaged over many years.

control 3, made from a blood sample from a 71 year old Finnish male, with this being the first thaw of the line, collected for RNA at around passage 3.

iPSCs Generation: The study was carried out under Stanford Institutional Review Board and Stem Cell Research Oversight Committee guidelines. Research subjects' peripheral blood mononuclear cells (PBMCs) were collected and reprogrammed using the Sendai virus-mediated non-integrating technique (CytoTune™-iPS 2.0 Sendai Reprogramming Kit, Thermo Fisher Scientific).

Cell Culture: iPSCs were maintained under feeder-free conditions in defined E8 media (Thermo Fisher Scientific) on tissue culture plates coated with hESC-qualified Matrigel (BD Biosciences).

Sanger sequencing validation of cell lines

Genomic DNA was isolated, followed by Sanger sequencing to validate that the lines corresponded still to the research subject from which they were derived. For Individual 19, the primers used were:

NAA15-delGAAA-F-5'-ccaagaaggctgccgttaaactt-3'

NAA15-delGAAA-R-5'-ggcttctgaccatgaacaagatga-3'

Quantitative PCR (qPCR)

Total RNAs were isolated from the iPS cell lines using the miRNeasy Mini kit (QIAGEN). 1 µg of RNA was used for synthesizing cDNA with the iScriptTM cDNA Synthesis kit (Bio-Rad). 0.25 µL of the cDNA was used to quantify gene expression in triplicate using TaqMan probes (ThermoFisher NAA15 catalog #4331182, assay ID Hx00228208_m1) and TaqMan Universal PCR Master Mix (Thermo Fisher Scientific). Quantitaive RT-PCR was also performed using the primers used for the LCL lines, as described below. Additionally, RT-PCR, followed by Sanger sequencing was also conducted for the iPS lines using the cDNA primers:

NAA15-RT-F-5'-AGA GAA ATC CTG AAA ACT GGG CCT-3'

NAA15-RT-R-5'-GGC CTC ATC CAT CCA CCT TGC A-3'

Total RNA was extracted from frozen LCL pellets using the RSC simplyRNA cells kit (Maxwell). cDNA was synthesized from 2µg of RNA using random hexamer primers 5x First Strand Buffer and Superscript III reverse transcriptase (Invitrogen). 2 µl of cDNA was used as a template for quantitative real time PCR, in triplicate for each sample. Fast SYBR green (Applied Biosystems) was used to perform qRT-PCR using Step one PCR machine (Thermofisher). GAPDH was used for gene normalisation. Relative mRNA expression was calculated using the 2-δCt method where Ct represents cycle threshold. Primer Sequences are as follows; GAPDH F-5'-TGCACCACCAACTGCTTAGC-3',R-5'-GGCATGGACTGTGGTCATGAG-3',NAA15 F-5'-TGTATGGAGGTATTGGAAGCC-3', R-5'-CTCTTCATATCCAGGAGGCAT-3'.

RT-PCR with Sanger sequence quantitation

The PCR products were submitted for Sanger sequencing. Quantification of the ratio of wild-type cDNA to mutant cDNA was performed using TIDE[6].

DDD study participants

WES and variant analysis for individuals and their parents in the DDD cohort was performed using a previously described strategy [7].

Methods for Individual contributed by Rotterdam, listed as Individual 2/Family2

Trio (proband and parents) whole-exome sequencing was performed for clinical diagnostic purposes at the Erasmus Medical Center, Rotterdam. Exome-coding DNA was captured with the Agilent Sure Select Clinical Research Exome (CRE) kit and sequencing was performed on the HiSeq Illumina platform. Reads were aligned to GRCh37 (hg19) using BWA and variants were called by the Genome Analysis Toolkit (GATK, reference https:// software.broadinstitute. org/gatk). Detected variants were annotated, filtered (for de novo, autosomal recessive and X-linked variants) and prioritized using Bench lab NGS platform (Agilent Technologies Belgium). The variant was confirmed (the novo) by Sanger sequencing.

Methods for Individuals contributed by Antwerp, listed as Individual 16/Family 16, Individual 23/Family 23, and Individual 29/Family 29

Individuals were MIP sequenced as part of a research study as described by Stessman et al., 2017. Sanger confirmation of the variants detected in the Subject and parental testing was performed in Antwerp using standard protocols.

Methods for Individuals 10-1, 10-2 and 10-3 in Family 10: Sydney Family with *NAA15* variant

Whole Exome Sequencing (WES) was performed in Adelaide, Australia, on the mother both of her affected sons. Roche NimbleGen SeqCap EZ Exome v3 capture was used and the captured DNA sequenced on the HiSeq 2000 sequencing platform using 2x101 bp paired end sequencing. Reads were mapped with BWA-MEM 0.7.8-r455 and variant calling was performed using GATK 3.2-2-gec30cee software packages[8; 9], using default parameters and following GATK Best Practices v3. The human genome reference b37 with decoy sequences was used for read mapping. PCR duplicates were marked using the Picard 1.117 MarkDuplicates software. Annotation was performed using an in-house annotation pipeline. Sanger sequencing validation of the NAA15 variant was performed from genomic DNA using standard methods. Mapping achieved mean coverage of targets of 53x, 35.6x and 32.8x for individuals 10-1, 10-2 and 10-3 respectively and only 3% of target bases were not covered in each exome. Segregation analysis showed the presence of the variant in the mother and the two affected sons. The variant was not present in the unaffected father and unaffected sister.

Methods for Individual contributed by Ottawa, listed as Individual 13/Family 13

Trio-WES (proband and both his parents) was performed for Individual 36 at GeneDx for clinical diagnostic purposes. DNA was sequenced on the Illumina HiSeq platform with 100bp or greater paired-end reads and captured with the Agilent Sure Select Kit. Reads were aligned to reference sequence GRCh37/UCSC hg19 and analyzed for variants using a custom-developed analysis tool (Xome Analyzer). Analysis includes segregation and evaluation of variants identified to be de novo, compound heterozygous, homozygous, heterozygous and X-linked. Reporting was based on clinical information provided. Pathogenic variant confirmation was done by capillary sequencing.

Exome Sequencing and Analysis (Individual 21/Family 21 and Individual 25/Family 25, Baylor cases)

Proband whole-exome sequencing was performed, who were ascertained from around 10,000 individuals referred for clinical whole-exome sequencing test at Baylor Genetics. The test targets approximately 20,000 genes, including coding and untranslated region (UTR) exons. Library preparation, exome capture, HiSeq next-generation sequencing and data analyses were conducted as described[10]. Samples were also analyzed by a cSNP-array (Illumina HumanExome-12 v1 array) for quality-control assessment of exome data, as well as for

detecting large copy-number variants (CNVs) and regions of absence of heterozygosity (AOH). Candidate variants were further confirmed by Sanger sequencing in the proband and the unaffected parents for segregation studies.

Exome Sequencing and Analysis (Individual 18/Family 18 and Individual 22/Family 22, BHCMG cases)

Informed consent was obtained for participation in the Baylor Hopkins Center for Mendelian Genomics (BHCMG) study under Baylor College of Medicine Institutional Review Board protocol H-29697. Proband WES was performed in individual 18 and trio-WES (proband and both parents) was performed for individual 22, who were ascertained from ~6,700 individuals who have undergone WES in the BHCMG to date. Library preparation, exome capture using the VCRome 2.1 design (42 Mb NimbleGen, Cat. No. 06266380001), and sequencing on the HiSeq Illumina platform were performed as previously described[11; 12]. Variant data were aligned and mapped to the human reference genome (Hg19) using the Mercury in-house bioinformatics pipeline, variants were called using ATLAS, and annotated using an in-house pipeline (Cassandra) that incorporates Annotation of Genetic Variants (ANNOVAR)[13]. Orthogonal validation of identified SNPs was performed by Sanger sequencing of PCR amplicons. For family 18 the Sanger segregation studies were performed at Telemark Hospital, Norway.

Methods for Individual 31/Family 31 contributed by HudsonAlpha Institute for Biotechnology

Samples from the proband and his unaffected parents were used for whole genome sequencing as part of our Clinical Sequencing Exploratory Research (CSER) research study as described by Bowling et al., 2017[14]. Sanger confirmation of the variant detected in the Subject and parental testing was performed by Emory Genetics Laboratory.

Additional Acknowledgements

The study has UK Research Ethics Committee approval (10/H0305/83, granted by the Cambridge South REC, and GEN/284/12 granted by the Republic of Ireland REC). The research team acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network. S.G., E.S. and C.R. are supported by the Italian Ministry of Health and '5 per mille' funding. M.A.C. is supported by a Cerebral Palsy Alliance career development award. 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.

G.J.L serves on advisory boards for GenePeeks, Inc. and Seven Bridges Genomics, Inc. The Department of Molecular and Human Genetics at BCM derives revenue from molecular testing offered at Baylor Genetics Laboratories. J.R.L has stock ownership in 23 and Me, is a paid consultant for Regeneron Pharmaceuticals, has stock options in Lasergen, Inc., and is a co-inventor on multiple United States and European patents related to molecular diagnostics for inherited neuropathies, eye diseases, and bacterial genomic fingerprinting. E.E.E. is on the scientific advisory board (SAB) of DNAnexus, Inc. W.K.C. is on the scientific advisory board of the Regeneron Genetics Center. Richard Person, Rebecca Willaert are employees of GeneDx, Inc., a wholly owned subsidiary of OPKO Health, Inc. The remaining authors declare that they have no competing interests.

Supplemental References

1. Arnesen, T., Van Damme, P., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., Varhaug, J.E., Vandekerckhove, J., Lillehaug, J.R., Sherman, F., et al. (2009). Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. Proc Natl Acad Sci U S A 106, 8157-8162.
2. Janke, C., Magiera, M.M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno-Borchart, A., Doenges, G., Schwob, E., Schiebel, E., et al. (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. Yeast 21, 947-962.
3. Miller, C.A., 3rd, Martinat, M.A., and Hyman, L.E. (1998). Assessment of aryl hydrocarbon receptor complex interactions using pBEVY plasmids: expressionvectors with bi-directional promoters for use in Saccharomyces cerevisiae. Nucleic Acids Res 26, 3577-3583.
4. Kushnirov, V.V. (2000). Rapid and reliable protein extraction from yeast. Yeast 16, 857-860.
5. Arnesen, T., Anderson, D., Baldersheim, C., Lanotte, M., Varhaug, J.E., and Lillehaug, J.R. (2005). Identification and characterization of the human ARD1-NATH protein acetyltransferase complex. Biochem J 386, 433-443.
6. Brinkman, E.K., Chen, T., Amendola, M., and van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. Nucleic Acids Res 42, e168.
7. Deciphering Developmental Disorders, S. (2017). Prevalence and architecture of de novo mutations in developmental disorders. Nature 542, 433-438.
8. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760.
9. DePristo, M.A., Banks, E., Poplin, R., Garimella, K.V., Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nature genetics 43, 491-498.
10. Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 369, 1502-1511.
11. Lupski, J.R., Gonzaga-Jauregui, C., Yang, Y., Bainbridge, M.N., Jhangiani, S., Buhay, C.J., Kovar, C.L., Wang, M., Hawes, A.C., Reid, J.G., et al. (2013). Exome sequencing resolves apparent incidental findings and reveals further complexity of SH3TC2 variant alleles causing Charcot-Marie-Tooth neuropathy. Genome Med 5, 57.
12. Bainbridge, M.N., Wang, M., Wu, Y., Newsham, I., Muzny, D.M., Jefferies, J.L., Albert, T.J., Burgess, D.L., and Gibbs, R.A. (2011). Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. Genome Biol 12, R68.
13. Reid, J.G., Carroll, A., Veeraraghavan, N., Dahdouli, M., Sundquist, A., English, A., Bainbridge, M., White, S., Salerno, W., Buhay, C., et al. (2014). Launching genomics into the cloud: deployment of Mercury, a next generation sequence analysis pipeline. BMC Bioinformatics 15, 30.
14. Bowling, K.M., Thompson, M.L., Amaral, M.D., Finnila, C.R., Hiatt, S.M., Engel, K.L., Cochran, J.N., Brothers, K.B., East, K.M., Gray, D.E., et al. (2017). Genomic diagnosis for children with intellectual disability and/or developmental delay. Genome Med 9, 43.