



NAA10 p.(N101K) disrupts N-terminal acetyltransferase complex NatA and is associated with developmental delay and hemihypertrophy

Nina McTiernan¹ · Harinder Gill² · Carlos E. Prada^{3,4,5} · Harry Pachajoa^{6,7} · Juliana Lores^{6,7} · CAUSES study⁸ · Thomas Arnesen^{1,9,10}

Received: 20 November 2019 / Revised: 31 July 2020 / Accepted: 8 September 2020 / Published online: 24 September 2020

© The Author(s) 2020. This article is published with open access

Abstract

Nearly half of all human proteins are acetylated at their N-termini by the NatA N-terminal acetyltransferase complex. *NAA10* is evolutionarily conserved as the catalytic subunit of NatA in complex with *NAA15*, but may also have NatA-independent functions. Several *NAA10* variants are associated with genetic disorders. The phenotypic spectrum includes developmental delay, intellectual disability, and cardiac abnormalities. Here, we have identified the previously undescribed *NAA10* c.303C>A and c.303C>G p.(N101K) variants in two unrelated girls. These girls have developmental delay, but they both also display hemihypertrophy a feature normally not observed or registered among these cases. Functional studies revealed that *NAA10* p.(N101K) is completely impaired in its ability to bind *NAA15* and to form an enzymatically active NatA complex. In contrast, the integrity of *NAA10* p.(N101K) as a monomeric acetyltransferase is intact. Thus, this *NAA10* variant may represent the best example of the impact of NatA mediated N-terminal acetylation, isolated from other potential *NAA10*-mediated cellular functions and may provide important insights into the phenotypes observed in individuals expressing pathogenic *NAA10* variants.

Introduction

N-terminal (Nt) acetylation is a ubiquitous protein modification that pertains to ~80% of the human proteome [1]. Eight N-terminal acetyltransferases (NATs), named NatA to NatH, have been identified to date, whereof all except NatG

are expressed in humans [2]. The cellular roles of Nt-acetylation are manifold and not fully understood, but some reported functions include regulation of protein complex formation, folding, degradation, subcellular localization, and membrane interactions [2–6]. NatA is the major NAT accounting for almost half of the Nt-acetylome due to its broad substrate specificity [1]. NatA is comprised of the catalytic subunit *NAA10* and its binding partners *NAA15*, *HYPK*, and *NAA50* (NatE) [7–11]. Binding of *NAA10* to *NAA15* ensures ribosomal anchoring and alters the substrate specificity of *NAA10* to NatA specific substrates

Supplementary information The online version of this article (<https://doi.org/10.1038/s41431-020-00728-2>) contains supplementary material, which is available to authorized users.

✉ Thomas Arnesen
Thomas.Arnese@uib.no

¹ Department of Biomedicine, University of Bergen, N-5020 Bergen, Norway

² Department of Medical Genetics, Children's and Women's Health Centre of British Columbia, Vancouver, BC V6H 3N1, Canada

³ Division of Human Genetics, Cincinnati Children's Hospital Medical Center, 45229 Cincinnati, OH, USA

⁴ Department of Pediatrics, University of Cincinnati College of Medicine, 45229 Cincinnati, OH, USA

⁵ Centro de Medicina Genomica y Metabolismo, Fundacion

Cardiovascular de Colombia, Floridablanca, Colombia

⁶ Centro de Investigaciones en Anomalías Congénitas y Enfermedades Raras Universidad Icesi, Cali, Colombia

⁷ Fundación Clínica Valle del Lili, Cali, Colombia

⁸ Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada

⁹ Department of Biological Sciences, University of Bergen, N-5020 Bergen, Norway

¹⁰ Department of Surgery, Haukeland University Hospital, N-5021 Bergen, Norway

including small, polar amino acids [10, 12–14]. Moreover, NAA10 also exists as a monomer in the cell and is suggested to independently act as a lysine acetyltransferase (KAT) and noncatalytic regulator of diverse target proteins [2, 15–19]. *NAA10* is an essential gene and loss of function is lethal in model organisms such as *T. brucei*, *D. rerio*, *D. melanogaster*, and *C. elegans* [20–23]. In humans, NAA10 has been implicated in cancer signalling pathways both as a tumour suppressor and an oncoprotein, and is believed to have a regulatory role in cell proliferation and survival [24]. Furthermore, *NAA10* missense variants have in recent years emerged as causative of genetic disease, collectively known as *NAA10*-related syndrome [25]. This X-linked condition is associated with a broad spectrum of phenotypes including developmental delay (DD), intellectual disability (ID), and cardiac abnormalities [26]. This was first discovered in 2011, when a *NAA10* c.109T>C p.(S37P) variant was detected as the cause of Ogden syndrome (OMIM #300855) [27]. Affected boys had severe global developmental delay, craniofacial abnormalities, hypotonia and cardiac arrhythmia, and died within 16 months of age. Studies revealed that the *NAA10* c.109T>C p.(S37P) variant led to impaired NatA complex formation as well as decreased Nt-acetylation of NatA substrates in patient cells [27–29]. A *NAA10* splice-site variant c.471+2T>A was identified as one cause of Lenz Microphthalmia Syndrome (OMIM #309800) in four males presented with anophthalmia, ID, developmental delay, and other malformations [30]. Popp et al. reported a boy and a girl carrying the *NAA10* variants c.319G>T p.(V107F), and c.346C>T p.(R116W) respectively, with severe ID, postnatal growth retardation, hypotonia, and behavioural anomalies [31]. Casey et al. described two brothers with ID, facial dysmorphism, scoliosis, and long QT who harboured a c.128A>C p.(Y43S) variant inherited from their mildly affected mother [32]. A recurrent missense variant, c.247C>T p.(R83C), has been identified in seven females with ID and developmental delay [26]. In all of these cases, this variant arose de novo, except for one case of maternal inheritance in which the female also had an affected brother who suffered a neonatal death. Furthermore, a missense variant affecting the same amino acid, c.248G>A p.(R83H), was recently detected in two unrelated boys with ID, developmental delay, and hypertrophic cardiomyopathy [33]. Missense variants affecting Arg83 are believed to impair Ac-CoA binding and cause reduced Nt-acetylation, resulting in the observed phenotypes [26, 33]. Three variants, c.384T>A p.(F128I), c.382T>A p.(F128L), and c.332T>G p.(V111G), have been reported in four females with ID and functional studies showed reduced Nt-acetylation caused by destabilisation of the NAA10 structure [26, 34]. Another three males displaying global DD, ID, and hypertrophic cardiomyopathy were found to harbour a c.215T>C p.(I72T) variant [35]. Finally,

a recent international cohort presented 23 individuals with ten different *NAA10* variants [36]. Three of the variants have previously been described including the recurrent c.247C>T p.(R83C) variant, which in this study was found in 11 more individuals. Novel *NAA10* variants presented in the cohort include c.29A>G p.(D10G), c.32T>G p.(L11R), c.259G>T p.(A87S), c.311C>A p.(A104D), c.361C>G p.(L121V), c.440T>C p.(M147T), and a frameshift variant c.455_458del p.(Thr152fs) [36]. An overview of previously described *NAA10* variants is available in the Supplementary information (Supplementary Table S1). As the clinical spectrum associated with *NAA10* deficiency is expanding and new variants continue to emerge, there is currently limited overall understanding of the underlying disease mechanisms involved. Here we present two unrelated females harbouring two different genetic *NAA10* variants, c.303C>A and c.303C>G, which both encode the same *NAA10* p.(N101K) variant. The females display overlapping phenotypes including developmental delay, dysmorphic features, hemihypertrophy, and hearing loss. Functional studies suggest that this variant only impairs NatA activity and not monomeric NAA10 function.

Materials and methods

Construction of plasmid

A plasmid encoding the NM_003491.3 *NAA10* c.303C>A p.(N101K) missense variant was generated using the Q5 Site Directed Mutagenesis Kit (New England Biolabs, MA, USA) and pcDNA3.1/*NAA10*-V5-His vector as template. The forward and reverse primers used were 5'-TGATAG AGAAATTC AATGCCAAATATGTCTCCC-3' and 5'-TG GCTCGAGAGGCCTGGT-3' respectively, with an annealing temperature of 68 °C. The variant was confirmed by sequencing.

Immunoprecipitation and Western blot analysis

Transfection of HeLa cells (ATTC, CCL-2) was performed using X-tremeGENE 9 DNA Transfection Reagent (Roche, Basel, Switzerland) as described in the instruction manual. The amount of DNA used for the different vectors were as follows: 10–12 µg of pcDNA3.1/*NAA10*-N101K-V5, 2.5–4 µg of pcDNA3.1/*NAA10*-WT-V5, 7.5–10 µg of pcDNA3.1/*NAA15*-myc-His or 10 µg of pcDNA3.1/*LacZ*-V5 (control-vector). 5–10 µg of empty pcDNA3.1/V5 plasmid was co-transfected with pcDNA3.1/*NAA10*-WT-V5 to attain equal amounts of DNA in the transfection mix. Transfected cells were harvested after two days of growth and lysed in 0.5–1 ml IPH lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA, 0.5% NP-40, 1×

complete EDTA-free protease inhibitor cocktail) for 15 min at 4 °C on a rotator. The cell lysates were centrifuged at 17,000 × *g* and 4 °C for 5 min and the supernatants collected. For immunoprecipitation (IP), the lysates were incubated with 1.5–4 µg of V5-tag antibody (Invitrogen #R960-25, CA, USA) or 3 µg NAA15 antibody [7] (BioGenes, Berlin, Germany) at 4 °C for 2 h on a rotator before adding 20–40 µl of washed Dynabeads Protein G (Invitrogen, CA, USA). After overnight incubation, the beads were washed three times in IPH lysis buffer. Then, the beads were either resuspended in 90 µl acetylation buffer (50 mM Tris-HCl pH 8.5, 1 mM EDTA, 10% Glycerol) to be used in a Nt-acetylation assay or in ×1 sample buffer to be analysed by Western blot. Samples were analysed using standard Western blot technique, and immunoblots were probed with V5-tag antibody (1:5000 dilution, Invitrogen #R960-25), NAA15 antibody (1:2000 dilution, BioGenes (named anti-NATH) [7]) and myc-tag antibody (1:2000 dilution, Invitrogen #R950-25). The ChemiDoc XRS+ system (Bio-Rad, CA, USA) coupled with Imagemol Software (Bio-Rad, CA, USA) was used to image and quantify protein bands.

[¹⁴C]-Ac-CoA-based Nt-acetylation assay

To compare the intrinsic enzyme activity of V5-immunoprecipitated NAA10 WT-V5 and NAA10 N101K-V5, Nt-acetylation assays were performed as described previously [37]. In short, we used triplicate reactions containing 10 µl immunoprecipitated enzyme, 50 µM [¹⁴C]-Ac-CoA (PerkinElmer, MA, USA), 200 µM oligopeptide SESS₂₄ (SESSKSRWGRPVGRRRRPVRVYP) or EEEI₂₄ (EEEIAALRWGRPVGRRRRPVRVYP) (BioGenes, Berlin, Germany), and acetylation buffer with a total volume of 25 µl. Oligopeptide was omitted in negative control reactions. Reactions were run for 30 min at 37 °C and 1400 rpm in a thermomixer. Afterwards, 23 µl of the supernatant was transferred onto P81 phosphocellulose filter discs (Millipore, MA, USA) which were subsequently washed three times in 10 mM HEPES buffer (pH 7.4) and air dried. Finally, the filter discs were submerged in 5 ml Ultima Gold F scintillation mixture (PerkinElmer, MA, USA) and the Nt-acetylated product was measured by a TriCarb 2900TR Liquid Scintillation Analyzer (PerkinElmer, MA, USA).

Multiple sequence alignment and structural modelling

A multiple sequence alignment was generated using Clustal Omega [38] and illustrated by ESPrpt 3.0 [39]. The protein sequences are available in Supplementary Table S2. The structural analysis of human NatA (PDB ID: 6C9M) [40] was performed using PyMOL [41]. Acetyl-CoA was

inserted in the hNatA structure through superimposition with the *S. pombe* NAA10 structure (PDB ID: 4KVX) [13] solved with acetyl-CoA bound.

Exome sequencing

The family 1 proband participated in a translational research study (CAUSES Study) with institutional ethical approval from the University of British Columbia [H15-00092]. The family 2 proband participated in a rare diseases registry, with institutional ethical approval from Fundación Valle del Lili [01504]. Genomic DNA was isolated using standard techniques from the peripheral blood of family trios (proband, mother, and father). Each member of the trio had hybridization-based exome capture sequencing performed on an Illumina platform at Ambry Genetics (Aliso Viejo, CA, USA) for family 1 and at Agilent Technologies (Santa Clara, CA, USA) for family 2. A bioinformatics pipeline identified functional rare variants that were consistent with Mendelian patterns of inheritance as determined with the trio family structure. Sequencing reads were aligned to a reference human genome based on hg19 (<ftp://hgdownload.cse.ucsc.edu/goldenPath/hg19/>) using Bowtie2–2.2.6 [42]. Once mapped, Picard-tools-1.139 (<http://broadinstitute.github.io/picard/>) marked duplicate reads, and GATK-3.5-0 was applied for indel realignment [43]. Samtools-0.1.19 and bcftools-0.2 were used to call and filter single-nucleotide variants and indels [44]. SnpEff-4.1L [45] annotated Ensembl 75 transcripts [46] and identified functional variants, which were then annotated and filtered against in-house and public databases: dbSNP 144 [47], 1000G [48], ESP6500SI-V2 [49], and ExAC v0.3 [50]. A list of rare functional variants consistent with Mendelian inheritance patterns was annotated and prioritised using custom scripts and ANNOVAR 2015–06–17 [51] with diagnostically relevant information (e.g., OMIM [52], ClinVar [53], DECIPHER [54], LOVD 3.0 [55], NCBI gene summaries (ftp://ftp.ncbi.nlm.nih.gov/refseq/H_sapiens/RefSeqGene/), RVIS [56], and dbNSFP v3.0 [57]). Functional studies were undertaken after both clinical teams independently approached the functional analysis lab at the University of Bergen. The NAA10 c.303C>G p.(N101K) and NAA10 c.303C>A p.(N101K) variants were submitted to ClinVar (SCV001193440 and SCV001335306).

Results

Clinical report: individual 1

Individual 1 is a female with a de novo NAA10 c.303C>A p.(N101K) (NG_031987.1 (NM_003491.3):c.303C>A, hg19: g.153197807G>T) variant who was referred to genetic

Fig. 1 Patient photographs.

Photographs of individual 1 showing dysmorphism (a), cup-shaped ears (b), hemihypertrophy (L>R) (c). Photographs of individual 2 showing dysmorphism (d), hemihypertrophy (R>L) (e) and broad hallux (f).



evaluation at 5 years of age for global developmental delay, dysmorphic features, and left sided hemihypertrophy (Fig. 1a–c). She is the first child of her non-consanguineous parents and has a healthy younger sibling. The pregnancy was complicated by light per vaginam bleeding in the second trimester. She was born with a birth weight of 3.15 kg. Asymmetry of her body was noted in the neonatal period with the entire left side being larger than the right. She fed very slowly initially but did gain weight. Left sided hip subluxation was noted and surgically corrected with good result.

She failed her initial hearing test and on follow up was noted to have mild bilateral sensorineural hearing loss and wears hearing aids. Ophthalmic assessment revealed right posterior embryotoxon and anomalous optic nerves. An MRI of the brain was performed abroad and thought to be normal and we are awaiting repeat locally. Parents note delays in all areas of development. She has been diagnosed with moderate intellectual disability. Chromosome microarray, cardiac echo, and renal ultrasound were all normal. She did not meet the clinical criteria for Beckwith–Wiedemann syndrome (BWS) in that there was no evidence of neonatal hypoglycaemia, birthweight was on the 25th centile and she had no dysmorphic features of this condition. In addition, molecular testing for BWS returned negative. Renal ultrasound surveillance has remained normal. At the time of our first meeting she was noted to have significant micrognathia, short palpebral fissures and simple

cup shaped ears (Fig. 1b). The hemihypertrophy is growing along with her, no sandal gap, thickening of the sole of the foot, vascular abnormalities, and lipomata. The left foot is 1.5 cm longer than the right foot (Fig. 1c). No other diagnosis for her asymmetry was diagnosed clinically. She has continued to grow along the 3rd centile for height and weight, and OFC has been $-2SD$ below the mean. Exome sequencing identified a de novo variant of uncertain significance (VUS) in the *NAA10* gene: *NAA10* c.303C>A p.(N101K).

Clinical report: individual 2

Individual 2 with a de novo *NAA10* c.303C>G p.(N101K) (NG_031987.1 (NM_003491.3):c.303C>G, hg19:g.153197807G>C) variant is a 4-year-old female referred to genetics evaluation for global developmental delay, dysmorphism, short stature, and right sided hemihypertrophy (Fig. 1d–f). Proband is the first child of non-consanguineous healthy parents of Colombian descent. She has one healthy sibling and no family history of congenital anomalies or intellectual disability. Mother was a healthy 29-year-old G2P1. Pregnancy was complicated by intrauterine growth restriction and prenatal imaging concerning for Dandy–Walker malformation. Individual was born at 36 weeks of gestational age via c-section. Birth weight was 1.53 kg (<3rd centile), length 40 cm (<3rd centile), and head circumference 29.5

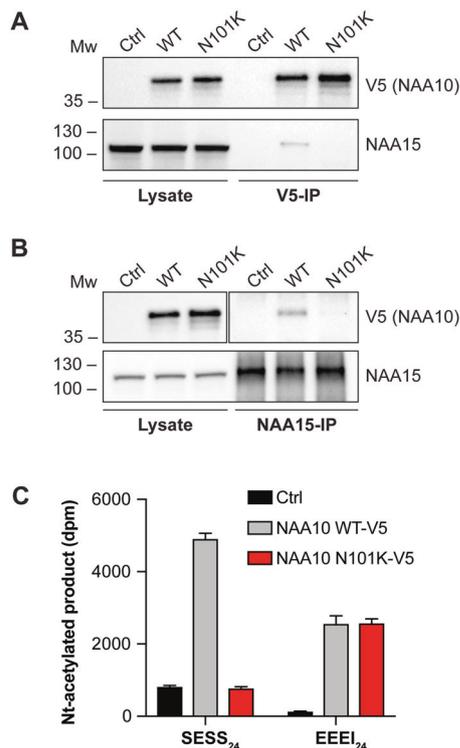


Fig. 2 NatA complex formation and catalytic activity of immunoprecipitated NAA10 WT-V5 and NAA10 N101K-V5. NAA10 WT-V5 and NAA10 N101K-V5 were overexpressed in HeLa cells, immunoprecipitated by V5-tag antibody (a) or NAA15 antibody (b) and analysed by Western blotting. Densitometry analysis was performed to quantify NAA10-V5 and NAA15 bands. c Nt-acetylation assay displaying catalytic activity of immunoprecipitated NAA10 WT-V5 and NAA10 N101K-V5. The measured catalytic activity toward NatA substrate SESS₂₄ and monomeric NAA10 substrate EEEL₂₄ was normalised to the amount of immunoprecipitated NAA15 and NAA10-V5, respectively. Reaction mixtures either with immunoprecipitated β -gal-V5 or without peptide were used as negative controls to account for background signal. The IP and activity measurements were performed in three independent setups, each with three technical replicates per assay. One representative setup is shown.

cm (<3rd centile). The individual had poor respiratory effort requiring admission to the neonatal intensive care unit for 10 days. No additional hospitalizations or respiratory concerns. Hearing evaluation revealed severe bilateral hearing loss. Developmental history is remarkable for delayed walking (age 3) and speech (first words at age 4). BWS testing has not been performed as she does not meet the clinical criteria. Dysmorphology evaluation at age 4 years was remarkable for broad forehead, arched eyebrows, esotropia, broad columella, and full lips. She had joint hypermobility, short fingers with trident appearance, broad hallux, and the right foot was longer than the left foot (0.8 cm at last evaluation) (Fig. 1e, f). Growth parameters measured at last evaluation were weight was 17.15 kg (50th centile, 0.1 SD), height 90.8 cm (<3rd centile, -3.19 SD),

and head circumference 49.5 cm (30th centile, 0.6 SD). Imaging evaluation included brain MRI that confirmed Dandy–Walker malformation and agenesis of the corpus callosum. Echocardiogram and renal ultrasound were normal. The girl has stereotypies and severe aggressive behaviour including biting and kicking caregivers. At last visit she was communicating using few words, her walk was more stable, and family reported that she needed help with feeding, dressing, and was also not yet potty trained. Genetic workup included normal chromosomes (46, XX) and chromosomal microarray. Exome trio analysis was performed and this identified a de novo VUS in the *NAA10* gene: *NAA10* c.303C>G p.(N101K). This variant was not found in GnomAD exomes or genomes.

Functional assessment of *NAA10* p.(N101K)

In order to investigate the catalytic activity of NAA10 N101K in comparison to NAA10 WT, V5-tagged NAA10 was overexpressed in HeLa cells and immunoprecipitated using V5-tag antibody. The immunoprecipitates were used in Nt-acetylation assays and the amount of NAA10-V5 and co-immunoprecipitated NAA15 in the samples was determined by Western blot analysis (Fig. 2a). Interestingly, there was not detected any co-immunoprecipitation of NAA15 with the NAA10 N101K-V5 variant in contrast to NAA10 WT-V5, indicating that the missense variant hinders NatA complex formation. This observation was further supported by a reciprocal IP using NAA15 antibody, where only NAA10 WT-V5 co-immunoprecipitated with NAA15 (Fig. 2b). To exclude the possibility that NAA10 N101K-V5 does not bind NAA15 because it cannot compete with endogenous NAA10, we also simultaneously overexpressed both NAA15-myc and NAA10-V5 and performed V5-IP (Fig. S1). In agreement with the other IP experiments, NAA15-myc did not co-immunoprecipitate with NAA10 N101K-V5 while NAA15-myc readily formed a complex with NAA10 WT-V5.

The catalytic activity of NAA10 N101K-V5 and NAA10 WT-V5 was tested in an Nt-acetylation assay using the oligopeptides SESS₂₄ and EEEL₂₄, representing a NatA substrate and an in vitro monomeric NAA10 substrate, respectively (Fig. 2c, Supplementary Table S3). Since SESS₂₄ is a NatA substrate, the catalytic activity toward this substrate was normalised to the amount of NAA15 in the immunoprecipitate, while the catalytic activity toward EEEL₂₄ was normalised to the amount of NAA10-V5. As shown in Fig. 2c, NAA10 N101K-V5 has an abolished NatA activity toward SESS₂₄ which is in accordance with the lack of co-immunoprecipitated NAA15 seen in the Western blot analysis (Fig. 2a). In contrast, the catalytic activity of NAA10 N101K-V5 toward EEEL₂₄ was equal to

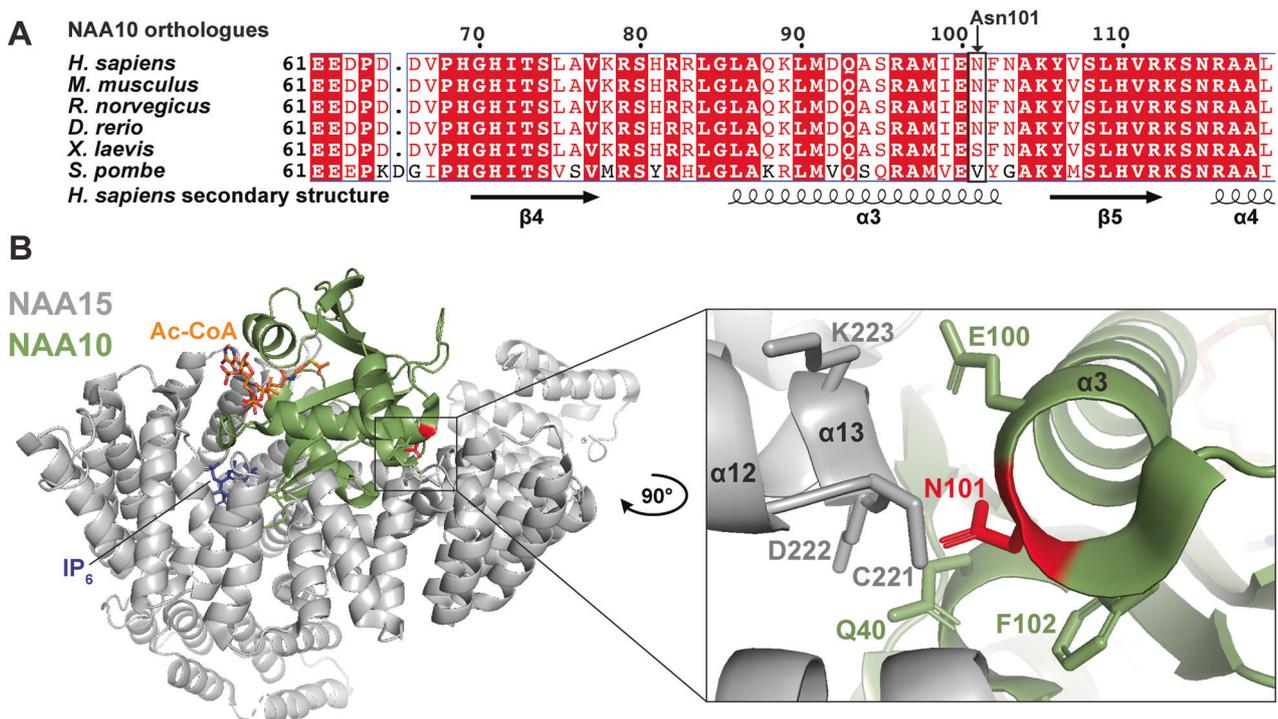


Fig. 3 NAA10 multiple sequence alignment and NatA structural analysis. **a** Multiple sequence alignment of NAA10 orthologues from human, mouse, rat, zebrafish, frog, and yeast. Secondary structure was determined from hNatA structure (PDB ID: 6C9M) [40] and amino acid conservation is indicated by red colour. **b** Human NatA structure (PDB ID: 6C9M) [40] with the auxiliary subunit NAA15 (grey), the

catalytic subunit NAA10 (green) and Ac-CoA and IP₆ shown as orange and blue sticks, respectively. The structure was superimposed on Ac-CoA from the *S. pombe* NAA10 structure (PDB ID: 4KVX) [13]. The variant site Asn101 is coloured red. Close-up of Asn101 shows that it is located in NAA10 $\alpha 3$ helix with its side chain protruding toward NAA15.

that of NAA10 WT-V5, suggesting that the monomeric NAA10 catalytic function is not affected by the variant. Taken together, these results indicate that the NAA10 c.303C>A and c.303C>G p.(N101K) variants are incapable of binding to NAA15, which results in abolished NatA catalytic activity, while monomeric NAA10 c.303C>A and c.303C>G p.(N101K) catalytic activity remains intact.

Multiple sequence alignment and structural analysis

NAA10 adapts the characteristic GCN5-related N-acetyltransferase (GNAT) fold common for many acetyltransferases. The GNAT fold is a highly conserved structural domain comprising an Ac-CoA binding region, six or seven β -strands and four α -helices [58]. Asn101 is located within the conserved GNAT fold of NAA10, but a multiple sequence alignment of NAA10 orthologues revealed that the Asn101 residue itself is only semi-conserved across the species presented in Fig. 3a. Structural investigations of the hNatA structure (PDB ID: 6C9M) [40] showed that Asn101 is located in the $\alpha 3$ helix in close proximity to NAA15, with its polar side chain protruding toward the $\alpha 12$ -loop- $\alpha 13$ segment of NAA15 (Fig. 3b). A structural analysis performed in PyMOL did not show any predicted interactions between the

side chain of Asn101 and surrounding amino acids in neither NAA15 nor NAA10 (Fig. 3b).

Discussion

In recent years, an increasing number of NAA10 variants have been identified in both male and female individuals with varying degrees of phenotype severity [25]. In this study, we report two novel de novo genetic variants NAA10 c.303C>A and c.303C>G p.(N101K) in two unrelated females with overlapping phenotypes including developmental delay, hemihypertrophy, hearing loss, and dysmorphic features. The significantly impaired function of NAA10 p.(N101K) defined in our biochemical assays combined with the other features clearly classify these variants as pathogenic (class 5) according to ACMG guidelines [59]. X-inactivation patterns have previously been suggested to influence phenotype severity in female carriers of NAA10 variants [26]. Due to the severe functional impairment of NatA activity of NAA10 p.(N101K), and the necessity of NatA mediated Nt-acetylation for life in multicellular eukaryotes [20–23, 60], we speculate that the females harbouring the NAA10 c.303C>A and c.303C>G

p.(N101K) variants have skewed X-inactivation. However, this has not been experimentally tested.

Most of the previously characterised *NAA10* missense variants have been shown to reduce monomeric NAA10 NAT-activity in vitro after being ectopically expressed and purified [26, 27, 31–34]. Interestingly, a recent cohort study demonstrated that ectopically purified NAA10 variants displayed different effects on catalytic activity depending on whether it was present in the core NatA complex (NAA10–NAA15) or the trimeric NatA/HYPK complex (HYPK is a stable interactor of the NatA complex in vivo) [36]. For this reason, we believe that testing NAT-activity using immunoprecipitated NAA10 or NatA complexes from human cells, as performed herein, presents a more reliable method for predicting the catalytic consequences of *NAA10* variants in vivo. Due to the unresolved biochemical complexity of NAA10 and the plethora of downstream cellular phenotypes [2], a standardised assay to comparatively assess the full impact of a larger number of *NAA10* variants is not currently available.

Both IP of NAA10 N101K-V5 and reciprocal IP of NAA15 showed that NAA10 N101K-V5 does not bind NAA15 (Fig. 2a, b). Since this could be due to an inability of NAA10 N101K-V5 to compete with endogenous NAA10, a scenario avoided in patient cells expressing *NAA10* p.(N101K), NatA complex formation was also assessed in cells overexpressing both NAA10-V5 and NAA15-myc. Interestingly, even with an excess of NAA15-myc, NAA10 N101K-V5 was not able to bind NAA15 (Fig. S1). This strongly suggests that the missense variant completely eradicates NatA complex formation.

The in vitro Nt-acetylation activity assay displayed an abolished NatA catalytic activity of NAA10 N101K-V5, whereas the monomeric NAA10 catalytic activity appeared unaffected (Fig. 2c). In contrast to NAA10 N101K-V5, a portion of the NAA10 WT-V5 is complexed with NAA15 which exert little if any catalytic activity toward EEEI₂₄ [14]. This could imply that the actual monomeric catalytic activity is slightly higher for NAA10 WT-V5 than the NAA10 N101K-V5 variant.

To understand why the *NAA10* c.303C>A and c.303C>G p.(N101K) variants hinder complex formation with NAA15, a sequence- and structural analysis was conducted. The multiple sequence alignment revealed that the Asn101 residue of NAA10 is not strictly conserved between orthologues (Fig. 3a). However, all the amino acids in position 101 are small and uncharged which suggest that lysine with its long, positively charged side chain is too dissimilar to be tolerated. Interestingly, Asn101 is located in the NAA10 α 3 helix which is part of the contact surface with NAA15 and the side chain of Asn101 is protruding toward the NAA15 α 12– α 13 loop (Fig. 3b). A previous study that delineated the NAA10–NAA15 interactions of *S. pombe* NatA [13] reported

that the NAA10 α 1-loop– α 2 region forms the most intimate interactions with NAA15, but the NAA10 α 3 helix was also found to make intermolecular interactions that supplements the NAA10–NAA15 interface. It is plausible that the longer side chain and positive charge of lysine can cause steric hindrance/and or charge repulsion. Consequently, potentially important intermolecular interactions mediated by Asn101 and/or other residues in the NAA10 α 3 helix could be disrupted and hinder optimal complex formation between NAA10 and NAA15.

Altogether, the data indicate that the *NAA10* c.303C>A and c.303C>G p.(N101K) variants abolish NatA complex formation and consequently also all NatA mediated N-terminal acetylation on the ribosome. Monomeric NAA10 has also been proposed to have NatA independent functions as a KAT catalysing lysine acetylation as well as a noncatalytic regulator of target substrates [2]. However, the NAA10 KAT activity toward some substrates has been disputed due to a lack of reproducibility [61]. The many cellular roles of NAA10 corroborates the complexity and challenge of defining the molecular mechanisms underlying clinical manifestations associated with *NAA10* deficiency. In the case of *NAA10* c.303C>A and c.303C>G p.(N101K), the variants do not seem to affect the monomeric functions of NAA10. Thus, the girls' phenotypes are most likely mediated via impaired NatA (NAA10–NAA15) Nt-acetylation activity, and not KAT, NAT, or noncatalytic roles of monomeric NAA10. Interestingly, the females harbouring the *NAA10* c.303C>A and c.303C>G p.(N101K) variants display hemihypertrophy, which has not previously been described for any individuals harbouring pathogenic *NAA10* variants. Thus this may be one such NatA-specific phenotype. In sum, *NAA10* p.(N101K) is the first variant reported to completely eradicate binding of NAA15 and it may uniquely reflect the functional impact of NatA.

Acknowledgements We are very grateful for the participation of the two families that are involved in this study. The investigators of the CAUSES Study include Shelin Adam, Christele Du Souich, Alison Elliott, Anna Lehman, Jill Mwenifumbo, Tanya Nelson, Clara Van Karnebeek, and Jan Friedman.

Funding The work has been supported by Mining for Miracles, British Columbia Children's Hospital Foundation and Genome British Columbia, the Research Council of Norway (project 249843), the Norwegian Health Authorities of Western Norway (project 912176), and the Norwegian Cancer Society. The funding bodies had no role in the design of the study, collection, analysis, or interpretation of data or in writing the manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests.

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Arnesen T, Van Damme P, Polevoda B, Helsens K, Evjenth R, Colaert N, et al. Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. *Proc Natl Acad Sci USA*. 2009;106:8157–62.
- Aksnes H, Ree R, Arnesen T. Co-translational, post-translational, and non-catalytic roles of N-terminal acetyltransferases. *Mol Cell*. 2019;73:1097–114.
- Shemorry A, Hwang CS, Varshavsky A. Control of protein quality and stoichiometries by N-terminal acetylation and the N-end rule pathway. *Mol Cell*. 2013;50:540–51.
- Holmes WM, Mannakee BK, Gutenkunst RN, Serio TR. Loss of amino-terminal acetylation suppresses a prion phenotype by modulating global protein folding. *Nat Commun*. 2014;5:4383.
- Dikiy I, Eliezer D. N-terminal acetylation stabilizes N-terminal helicity in lipid- and micelle-bound alpha-synuclein and increases its affinity for physiological membranes. *J Biol Chem*. 2014;289:3652–65.
- Scott DC, Monda JK, Bennett EJ, Harper JW, Schulman BA. N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. *Science*. 2011;334:674–8.
- Arnesen T, Anderson D, Baldersheim C, Lanotte M, Varhaug JE, Lillehaug JR. Identification and characterization of the human ARD1-NATH protein acetyltransferase complex. *Biochem J*. 2005;386(Pt 3):433–43.
- Arnesen T, Starheim KK, Van Damme P, Evjenth R, Dinh H, Betts MJ, et al. The chaperone-like protein HYPK acts together with NatA in cotranslational N-terminal acetylation and prevention of huntingtin aggregation. *Mol Cell Biol*. 2010;30:1898–909.
- Arnesen T, Anderson D, Torsvik J, Halseth HB, Varhaug JE, Lillehaug JR. Cloning and characterization of hNAT5/hSAN: an evolutionarily conserved component of the NatA protein N- α -acetyltransferase complex. *Gene*. 2006;371:291–5.
- Gautschi M, Just S, Mun A, Ross S, Rucknagel P, Dubaquié Y, et al. The yeast N(alpha)-acetyltransferase NatA is quantitatively anchored to the ribosome and interacts with nascent polypeptides. *Mol Cell Biol*. 2003;23:7403–14.
- Mullen JR, Kayne PS, Moerschell RP, Tsunasawa S, Gribskov M, Colavito-Shepanski M, et al. Identification and characterization of genes and mutants for an N-terminal acetyltransferase from yeast. *EMBO J*. 1989;8:2067–75.
- Magin RS, Deng S, Zhang H, Cooperman B, Marmorstein R. Probing the interaction between NatA and the ribosome for co-translational protein acetylation. *PLOS One*. 2017;12:e0186278.
- Liszczyk G, Goldberg JM, Foyen H, Petersson EJ, Arnesen T, Marmorstein R. Molecular basis for N-terminal acetylation by the heterodimeric NatA complex. *Nat Struct Mol Biol*. 2013;20:1098–105.
- Van Damme P, Evjenth R, Foyen H, Demeyer K, De Bock P-J, Lillehaug JR, et al. Proteome-derived peptide libraries allow detailed analysis of the substrate specificities of N(alpha)-acetyltransferases and point to hNaa10p as the post-translational actin N(alpha)-acetyltransferase. *Mol Cell Proteom*. 2011;10:M110.004580.
- Lim J-H, Chun Y-S, Park J-W. Hypoxia-inducible factor-1 α obstructs a Wnt signaling pathway by inhibiting the hARD1-mediated activation of β -catenin. *Cancer Res*. 2008;68:5177–84.
- Qian X, Li X, Cai Q, Zhang C, Yu Q, Jiang Y, et al. Phosphoglycerate kinase 1 phosphorylates Beclin1 to induce autophagy. *Mol Cell*. 2017;65:917–31.e6.
- Ree R, Varland S, Arnesen T. Spotlight on protein N-terminal acetylation. *Exp Mol Med*. 2018;50:90.
- Lee C-C, Peng S-H, Shen L, Lee C-F, Du T-H, Kang M-L, et al. The role of N- α -acetyltransferase 10 protein in DNA methylation and genomic imprinting. *Mol Cell*. 2017;68:89–103.e7.
- Seo JH, Park J-H, Lee EJ, Vo TTL, Choi H, Kim JY, et al. ARD1-mediated Hsp70 acetylation balances stress-induced protein refolding and degradation. *Nat Commun*. 2016;7:12882.
- Ingram AK, Cross GAM, Horn D. Genetic manipulation indicates that ARD1 is an essential N- α -acetyltransferase in *Trypanosoma brucei*. *Mol Biochem Parasit*. 2000;111:309–17.
- Wang Y, Mijares M, Gall MD, Turan T, Javier A, Bornemann DJ, et al. *Drosophila* variable nurse cells encodes arrest defective 1 (ARD1), the catalytic subunit of the major N-terminal acetyltransferase complex. *Dev Dyn*. 2010;239:2813–27.
- Sönnichsen B, Koski LB, Walsh A, Marschall P, Neumann B, Brehm M, et al. Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature*. 2005;434:462–9.
- Ree R, Myklebust LM, Thiel P, Foyen H, Fladmark KE, Arnesen T. The N-terminal acetyltransferase Naa10 is essential for zebrafish development. *Biosci Rep*. 2015;35:e00249.
- Kalvik TV, Arnesen T. Protein N-terminal acetyltransferases in cancer. *Oncogene*. 2013;32:269–76.
- Wu Y, Lyon GJ. NAA10-related syndrome. *Exp Mol Med*. 2018;50:85.
- Saunier C, Støve SI, Popp B, Gérard B, Blenski M, Ahmew N, et al. Expanding the phenotype associated with NAA10-related N-terminal acetylation deficiency. *Hum Mutat*. 2016;37:755–64.
- Rope Alan F, Wang K, Evjenth R, Xing J, Johnston Jennifer J, Swensen, Jeffrey J, et al. Using VAAST to identify an X-linked disorder resulting in lethality in male infants due to N-terminal acetyltransferase deficiency. *Am J Hum Genet*. 2011;89:28–43.
- Myklebust LM, Van Damme P, Støve SI, Dörfel MJ, Abboud A, Kalvik TV, et al. Biochemical and cellular analysis of Ogden syndrome reveals downstream N-terminal acetylation defects. *Hum Mol Genet*. 2015;24:1956–76.
- Van Damme P, Støve SI, Glomnes N, Gevaert K, Arnesen T. A *Saccharomyces cerevisiae* model reveals in vivo functional impairment of the ogden syndrome N-terminal acetyltransferase NAA10 Ser37Pro mutant. *Mol Cell Proteom*. 2014;13:2031–41.
- Esmailpour T, Riazifar H, Liu L, Donkervoort S, Huang VH, Madaan S, et al. A splice donor mutation in results in the dysregulation of the retinoic acid signalling pathway and causes Lenz microphthalmia syndrome. *J Med Genet*. 2014;51:185–96.
- Popp B, Støve SI, Ende S, Myklebust LM, Hoyer J, Sticht H, et al. De novo missense mutations in the NAA10 gene cause severe non-syndromic developmental delay in males and females. *Eur J Hum Genet*. 2015;23:602–9.
- Casey JP, Støve SI, McGorrian C, Galvin J, Blenski M, Dunne A, et al. NAA10 mutation causing a novel intellectual disability syndrome with Long QT due to N-terminal acetyltransferase impairment. *Sci Rep*. 2015;5:16022.

33. Ree R, Geithus AS, Tørring PM, Sørensen KP, Damkjær M, Lynch SA, et al. A novel NAA10 p.(R83H) variant with impaired acetyltransferase activity identified in two boys with ID and microcephaly. *BMC Med Genet.* 2019;20:101.
34. McTiernan N, Støve SI, Aukrust I, Mårli MT, Myklebust LM, Houge G, et al. NAA10 dysfunction with normal NatA-complex activity in a girl with non-syndromic ID and a de novo NAA10 p.(V111G) variant – a case report. *BMC Med Genet.* 2018; 19:47.
35. Støve SI, Blenski M, Stray-Pedersen A, Wierenga KJ, Jhangiani SN, Akdemir ZC, et al. A novel NAA10 variant with impaired acetyltransferase activity causes developmental delay, intellectual disability, and hypertrophic cardiomyopathy. *Eur J Hum Genet.* 2018;26:1294–305.
36. Cheng H, Gottlieb L, Marchi E, Kleyner R, Bhardwaj P, Rope AF, et al. Phenotypic and biochemical analysis of an international cohort of individuals with variants in NAA10 and NAA15. *Hum Mol Genet.* 2019;28:2900–19.
37. Drazic A, Arnesen T. [14C]-acetyl-coenzyme A-based in vitro N-terminal acetylation assay. In: Schilling O, editor. *Protein terminal profiling: methods and protocols.* New York, NY: Springer New York; 2017. p. 1–8.
38. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 2011;7:539.
39. Robert X, Gouet P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 2014;42(W1): W320–W4.
40. Gottlieb L, Marmorstein R. Structure of human NatA and its regulation by the huntingtin interacting protein HYPK. *Structure.* 2018;26:925–35.e8.
41. Schrödinger, LLC The PyMOL Molecular Graphics System, Version 2.3. 2019.
42. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9:357–9.
43. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20:1297–303.
44. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence Alignment/Map format and SAMtools. *Bioinformatics.* 2009;25:2078–9.
45. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly.* 2012;6:80–92.
46. Flicek P, Amode MR, Barrell D, Beal K, Billis K, Brent S, et al. Ensembl 2014. *Nucleic Acids Res.* 2014;42(D1):D749–D55.
47. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001;29:308–11.
48. Auton A, Abecasis GR, Altshuler DM, Durbin RM, Abecasis GR, Bentley DR, et al. A global reference for human genetic variation. *Nature.* 2015;526:68–74.
49. Exome Variant Server. NHLBI GO Exome Sequencing Project (ESP) (Seattle, WA) [<http://evs.gs.washington.edu/EVS/>].
50. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature.* 2016;536:285–91.
51. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* 2010;38:e164.
52. Online Mendelian Inheritance in Man, OMIM. McKusick–Nathans Institute of Genetic Medicine, Johns Hopkins University (Baltimore, MD) [<https://omim.org/>].
53. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic Acids Res.* 2018;46:D1062–d7.
54. Firth HV, Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: database of chromosomal imbalance and phenotype in humans using Ensembl resources. *Am J Hum Genet.* 2009;84:524–33.
55. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. *Hum Mutat.* 2011;32:557–63.
56. Petrovski S, Wang Q, Heinzen EL, Allen AS, Goldstein DB. Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet.* 2013;9:e1003709.
57. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A one-stop database of functional predictions and annotations for human non-synonymous and splice-site SNVs. *Hum Mutat.* 2016;37:235–41.
58. Vetting MWS, de Carvalho LP, Yu M, Hegde SS, Magnet S, Roderick SL, et al. Structure and functions of the GNAT superfamily of acetyltransferases. *Arch Biochem Biophys.* 2005;433:212–26.
59. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17:405–24.
60. Linster E, Stephan I, Bienvenut WV, Maple-Grødem J, Myklebust LM, Huber M, et al. Downregulation of N-terminal acetylation triggers ABA-mediated drought responses in *Arabidopsis*. *Nat Commun.* 2015;6:7640.
61. Magin RS, March ZM, Marmorstein R. The N-terminal acetyltransferase Naa10/ARD1 does not acetylate lysine residues. *J Biol Chem.* 2016;291:5270–7.