THUMPD1 bi-allelic variants cause loss of tRNA acetylation and a syndromic neurodevelopmental disorder

Authors

Martin Broly, Bogdan V. Polevoda, Kamel M. Awayda, ..., Bertrand Isidor, Benjamin Cogné, Mitchell R. O'Connell

Correspondence

benjamin.cogne@chu-nantes.fr **(B.C.)**,
mitchell_oconnell@urmc.rochester.edu **(M.R.O.)**

Broly et al., 2022, The American Journal of Human Genetics *109*, 587–600
April 7, 2022
https://doi.org/10.1016/j.ajhg.2022.02.001

CellPress

THUMPD1 bi-allelic variants cause loss of tRNA acetylation and a syndromic neurodevelopmental disorder

Martin Broly,[1,29,30] Bogdan V. Polevoda,[2,30] Kamel M. Awayda,[2] Ning Tong,[2] Jenna Lentini,[3] Thomas Besnard,[1,4] Wallid Deb,[1,4] Declan O'Rourke,[5] Julia Baptista,[6,7] Sian Ellard,[6,7] Mohammed Almannai,[8] Mais Hashem,[9] Ferdous Abdulwahab,[9] Hanan Shamseldin,[9] Saeed Al-Tala,[10] Fowzan S. Alkuraya,[9,11] Alberta Leon,[12] Rosa L.E. van Loon,[13] Alessandra Ferlini,[14] Mariabeatrice Sanchini,[14] Stefania Bigoni,[14] Andrea Ciorba,[15] Hans van Bokhoven,[16,17] Zafar Iqbal,[18] Almundher Al-Maawali,[19,20] Fathiya Al-Murshedi,[19,20] Anuradha Ganesh,[21] Watfa Al-Mamari,[22] Sze Chern Lim,[23,24] Lynn S. Pais,[25] Natasha Brown,[23,24,26] Saima Riazuddin,[27,28] Stéphane Bézieau,[1,4] Dragony Fu,[3] Bertrand Isidor,[1,4] Benjamin Cogné,[1,4,*] and Mitchell R. O'Connell[2,*]

Summary

Covalent tRNA modifications play multi-faceted roles in tRNA stability, folding, and recognition, as well as the rate and fidelity of translation, and other cellular processes such as growth, development, and stress responses. Mutations in genes that are known to regulate tRNA modifications lead to a wide array of phenotypes and diseases including numerous cognitive and neurodevelopmental disorders, highlighting the critical role of tRNA modification in human disease. One such gene, *THUMPD1,* is involved in regulating tRNA N4-acetylcytidine modification (ac4C), and recently was proposed as a candidate gene for autosomal-recessive intellectual disability. Here, we present 13 individuals from 8 families who harbor rare loss-of-function variants in *THUMPD1*. Common phenotypic findings included global developmental delay, speech delay, moderate to severe intellectual deficiency, behavioral abnormalities such as angry outbursts, facial dysmorphism, and ophthalmological abnormalities. We demonstrate that the bi-allelic variants identified cause loss of function of THUMPD1 and that this defect results in a loss of ac4C modification in small RNAs, and of individually purified tRNA-Ser-CGA. We further corroborate this effect by showing a loss of tRNA acetylation in two CRISPR-Cas9-generated *THUMPD1* KO cell lines. In addition, we also show the resultant amino acid substitution that occurs in a missense *THUMPD1* allele identified in an individual with compound heterozygous variants results in a marked decrease in THUMPD1 stability and RNA-binding capacity. Taken together, these results suggest that the lack of tRNA acetylation due to THUMPD1 loss of function results in a syndromic form of intellectual disability associated with developmental delay, behavioral abnormalities, hearing loss, and facial dysmorphism.

Introduction

Across all organisms, well more than 100 types of transfer RNA (tRNA) modifications exist and tRNAs contain on average 13 modifications per molecule.[1–6] Post-transcriptional modifications of tRNAs play multi-faceted roles in tRNA stability, folding, and recognition, as well as the rate and fidelity of translation, and other cellular processes such as growth, development, and stress responses.[7–10]

[1]Service de Génétique Médicale, CHU de Nantes, 44000 Nantes, France; [2]Department of Biochemistry and Biophysics, Center for RNA Biology, University of Rochester, Rochester, NY 14642, USA; [3]Department of Biology, Center for RNA Biology, University of Rochester, Rochester, NY 14642, USA; [4]Université de Nantes, CNRS, INSERM, l'institut du thorax, 44093 Nantes, France; [5]Department of Neurology, Children's Health Ireland at Temple Street, Dublin, D01 XD99, Ireland; [6]Exeter Genomics Laboratory, Royal Devon and Exeter NHS Foundation Trust, Exeter EX2 5DW, UK; [7]Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter EX1 2LU, UK; [8]Section of Medical Genetics, Children's Hospital, King Fahad Medical City, Riyadh 12231, Saudi Arabia; [9]Department of Translational Genomics, Center for Genomics Medicine, King Faisal Specialist Hospital and Research Center, Riyadh 11564, Saudi Arabia; [10]Pediatrics Department, Armed Forces Hospital, Khamis Mushait 62413, Saudi Arabia; [11]College of Medicine, Alfaisal University, Riyadh 11533, Saudi Arabia; [12]Research & Innovation (R&I Genetics) Srl, Genetic Laboratory, 35127 Padua, Italy; [13]Department of Genetics, University of Utrecht, University Medical Center Utrecht, 3584 CX Utrecht, the Netherlands; [14]Medical Genetics Unit, Department of Medical Sciences, University of Ferrara, Ferrara 44121, Italy; [15]ENT & Audiology Unit, Department of Neurosciences, University Hospital of Ferrara, 44124 Cona FE, Italy; [16]Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, 6525 HR Nijmegen, the Netherlands; [17]Department of Human Genetics, Radboud University Medical Center, 6525 HR Nijmegen, the Netherlands; [18]Department of Neurology, Oslo University Hospital, 0188 Oslo, Norway; [19]Department of Genetics, College of Medicine and Health Sciences, Sultan Qaboos University, Muscat 123, Oman; [20]Genetic and Developmental Medicine Clinic, Sultan Qaboos University Hospital, Muscat 123, Oman; [21]Department of Ophthalmology, Pediatric Ophthalmology and Ocular Genetics Unit, Sultan Qaboos University Hospital, Muscat 123, Oman; [22]Department of Child Health, Sultan Qaboos University Hospital, Muscat 123, Oman; [23]Victorian Clinical Genetics Services, Parkville, VIC 3052, Australia; [24]Murdoch Children's Research Institute, Parkville, VIC 3052, Australia; [25]Broad Center for Mendelian Genomics, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA; [26]Department of Paediatrics, University of Melbourne, Parkville, VIC 3052, Australia; [27]Laboratory of Molecular Genetics, Department of Otorhinolaryngology Head and Neck Surgery, School of Medicine, University of Maryland, Baltimore, MD 21201, USA; [28]Pakistan Institute of Medical Sciences, Shaheed Zulfiqar Ali Bhutto Medical University, Sector G-8/3, Islamabad, Pakistan
[29]Present address: Swiss Institute of Genomic Medicine, Medigenome, 1207 Geneva, Switzerland
[30]These authors contributed equally
*Correspondence: benjamin.cogne@chu-nantes.fr (B.C.), mitchell_oconnell@urmc.rochester.edu (M.R.O.)
https://doi.org/10.1016/j.ajhg.2022.02.001.

An increasing number of genes involved in tRNA modifications are implicated in intellectual disability. Variants in at least ten different genes[11] involved in tRNA modifications, including NOP2/Sun RNA Methyltransferase 2 (*NSUN2* [MIM: 610916]),[12] pseudouridine synthase 3 (*PUS3* [MIM: 616283]),[13,14] adenosine deaminase tRNA specific 3 (*ADAT3* [MIM: 615302]),[15,16] DALR anticodon binding domain containing 3 (*DALRD3* [MIM: 618904]),[17] tRNA Methyltransferase 1 (*TRMT1* [MIM: 611669]),[18] Cytosolic Thiouridylase Subunit 2 (*CTU2* [MIM: 617057]),[19] AlkB Homolog 8, tRNA Methyltransferase (*ALKBH8* [MIM: 613306]),[20] Pseudouridine Synthase 7 (*PUS7* [MIM: 616261]),[21,22] and WD Repeat Domain 4 (*WDR4* [MIM: 605924]),[23] have now been shown to be associated with neurological and neurodevelopmental disorders, emphasizing an increasingly apparent role for the importance of tRNA modifications in human neurodevelopment.

N4-acetylcytidine (ac4C) is a modification present in the wobble position of *E. coli* tRNA-Met,[24,25] in rRNAs in a range of organisms[26,27] and specifically at two conserved positions in human 18S rRNA (helix 34 and 45),[28–31] and at position 12 in eukaryotic serine and leucine tRNAs, with tRNA acetylation being conserved from yeast to mammals.[31–35] ac4C modification is catalyzed by the enzyme N-acetyltransferase 10 (NAT10) or its orthologs, and in humans, acetylation of serine and leucine tRNAs requires an additional adaptor protein known as THUMP-domain containing protein 1 (THUMPD1, or Tan1p in yeast).[31,35] This protein contains a single THUMP domain, named for its presence in thiouridine synthases, RNA methylases, and pseudouridine synthases across all domains of life.[36] The THUMP domain is an RNA-binding domain[35] and serves to recruit RNA molecules for subsequent modification by catalytic domains within the same polypeptide or by partner proteins.[31,37–39] In yeast, loss of the THUMPD1 ortholog, Tan1p, results in hypoacetylation of serine and leucine tRNAs,[31,35,40] and a decrease in tRNA stability that leads to a decrease in tRNA serine and leucine levels.[30,40–42] This results in a temperature-sensitive growth phenotype,[35,40] but it is unclear whether this effect is conserved in humans.

Recently, Maddirevula et al. proposed *THUMPD1* (MIM: 616662) as a candidate gene for syndromic autosomal-recessive intellectual disability based on findings from a single individual.[43] Here, we report the clinical phenotypes associated with bi-allelic variants in *THUMPD1* in 13 individuals from 8 families showing syndromic intellectual disability. We demonstrate that the variants identified cause loss of function of THUMPD1 and that this defect results in a loss of ac4C modification in small RNAs and of individually purified tRNA-Ser-CGA. We further corroborate this effect by showing a loss of tRNA acetylation in two CRISPR-Cas9-generated *THUMPD1* KO cell lines. Taken together, these results suggest that the lack of tRNA acetylation due to THUMPD1 loss of function results in syndromic intellectual disability.

Material and methods

Study participants

Individuals were recruited via an international, multi-center collaboration with research and diagnostic sequencing laboratories and medical genetics departments. Individuals were clinically evaluated in separate centers and contact between researchers was facilitated by use of the web-based tool GeneMatcher.[44] Inclusion criteria were rare bi-allelic variants in *THUMPD1* classified as pathogenic or likely pathogenic according to the American College of Medical Genetics and Genomics (ACMG) guidelines for the interpretation of sequence variants carried by individuals with compatible phenotypes.[45] Photographs and clinical data reports were provided by each attending physician and/or a detailed review of medical records.

Ethics statement

Informed consent to participate in the study was obtained from all individuals, or their parents in case of minor individuals, in accordance with local institutional review boards. Samples and clinical information were obtained for each individual following local procedures and ethical standards. When needed, additional written consent to use photographs in this publication was also obtained.

Sequencing and bioinformatics analysis

Whole-exome sequencing (WES) was performed on blood-derived DNA samples using each center's analysis platforms and next-generation sequencing (NGS) pipeline (see supplemental methods). Exome data were interpreted in agreement with local practices. cDNA and protein sequence variants are described in accordance with the recommendations of the Human Genome Variation Society using NCBI reference sequence GenBank: NM_001304550.1 and genomic coordinates from GRCh37/hg19. Priority was given to rare exonic or donor/acceptor splicing variants in accordance with the pedigree and phenotype, fitting a recessive (homozygous or compound heterozygous) model and/or variants in genes previously linked to developmental delay, intellectual disability, and other neurological disorders. Upon whole-exome sequencing analysis, *THUMPD1* variants were selected and all the other candidate variants were further verified using Sanger sequencing.

Lymphoblast cell culture

Lymphoblast cells lines were generated by EBV immortalization of primary human lymphocytes isolated from whole blood of a 23-year-old male with intellectual disability, dysmorphic features, hypertension, and diabetes (from Saudi Arabia) with bi-allelic (homozygous) variant in *THUMPD1* containing a GenBank: NM_001304550.1: c.706C>T (p.Gln236*) nonsense mutation in the corresponding mRNA[43] (case ID 15DG1395, individual F2:II.1 herein) as well as from whole blood of two independent control unrelated healthy Saudi individuals. Lymphoblast cells were grown in RPMI medium (Lonza) containing 25 mM glucose, 1× GlutaMAX Supplement (GIBCO), and 15% bovine calf serum (HyClone), in presence of ampicillin and streptomycin.

HeLa and HEK293T cell culture and knockout generation

HeLa cervical carcinoma cells and HEK293T human embryonic kidney cell lines were purchased from ATCC. *NAT10*$^{-/-}$ HeLa cells were obtained from Dr. Shalini Oberdoerffer (Laboratory of Receptor Biology and Gene Expression, National Cancer Institute, NIH, Bethesda, MD).[46] HeLa and HEK293T *THUMPD1* knockout (KO) cells were generated using a CRISPR-Cas9 Gene Knockout kit

(Synthego). Briefly, purified SpCas9 and synthetic gRNA targeting exon 2 were complexed and electroporated into cells using the Neon Transfection System (Invitrogen) according to the manufacturer's instructions. At 48 h post-transfection, cells were seeded in dilutions into 10-cm dishes. Single colonies were tested for depletion of THUMPD1 by western blot (see western blot analysis for more details). To determine the presence of insertions or deletions in *THUMPD1* targeted clones, genomic DNA was isolated using a Wizard SV genomic DNA purification System (Promega) and exon 2 of *THUMPD1* was PCR amplified using the primers listed in Table S1. PCR products were sequenced by Sanger sequencing and analyzed by ICE (Synthego). Clones with mutations in both *THUMPD1* alleles were identified and one clone was selected for further experiments. Additional confirmation of *THUMPD1* KO status was determined by sub-cloning PCR products into pCR2.1-TOPO using the TOPO TA cloning Kit (Thermo Fisher Scientific) and subsequent Sanger sequencing of individual clones. THUMPD1 protein in knockout cell lines was evaluated by western blots with corresponding cell extracts probed by anti-THUMPD1 antibodies (Bethyl Labs and/or Thermo Fisher Scientific). GAPDH and Tubulin antibodies served as a loading control for all cell protein extracts (Abcam). All other transfection experiments were performed using JetPRIME transfection reagent (PolyPlus) according to the manufacturer's instructions.

Construction of THUMPD1 WT and p.Pro164Ser *E. coli* expression plasmids

Wild-type full-length human *THUMPD1* in pDonr221 was obtained from the CCSB Human ORFeome collection and subcloned into pHGWA (which contains an N-terminal hexa-histidine tag followed by a TEV protease site) using standard Gateway Cloning (Thermo Fisher Scientific). Using this resultant plasmid as a template, THUMPD1 p.Pro164Ser construct was generated using standard site-directed mutagenesis procedure using the primers listed in Table S1. The resultant constructs were verified by Sanger sequencing. These plasmids have been deposited to Addgene.

Recombinant THUMPD1 *E. coli* overexpression and purification

To purify THUMPD1 WT GenBank: NP_001291479.1 and THUMPD1 p.Pro164Ser proteins, the corresponding 6× His-THUMPD1 constructs were transformed into *E. coli* Rosetta 2(DE3) strain (Novagen). Cells were grown in LB media supplemented with 0.5% w/v glucose at 37°C and protein expression was induced by addition of 0.5 mM IPTG followed by incubation at 16°C overnight. Cell lysis was conducted by sonication in lysis buffer (50 mM HEPES [pH 7.0], 1 M NaCl, 30 mM imidazole, 5% (v/v) glycerol, 1 mM DTT, 0.5 mM PMSF, EDTA-free protease inhibitor [Roche], 400 μg of RNase A [Sigma] per gram of cell pellet). The cell lysate was clarified by centrifugation and metal ion affinity purification was carried out on the soluble fraction using an AKTA Pure FPLC instrument (Cytiva) with a HisTrap 1 mL column (Cytiva). THUMPD1-containing fractions were dialyzed overnight in gel filtration buffer (20 mM HEPES [pH 7.0], 200 mM KCl, 5% glycerol, 1 mM DTT) with 30 mM imidazole and TEV protease. The next day, a HisTrap column (Cytiva) was used to capture the cleaved 6× His-tag, and the THUMPD1-containing flowthrough was concentrated. Any residual impurities in the samples were removed by size exclusion chromatography on AKTA Pure FPLC instrument (Cytiva) with a Superdex 200 10/300 GL column (Cytiva), using gel filtration buffer. Fractions containing only THUMPD1 were collected, concentrated, flash-frozen in liquid nitrogen, and stored at −80°C.

Thermo-fluorometry (thermal shift assays)

THUMPD1 WT and THUMPD1 p.Pro164Ser protein thermal shift assays were performed using a QuantStudio 5 Real-Time PCR System instrument (Applied Biosystems). 10 μL reactions were carried out in MicroAmp Optical 384-well reaction plates (Thermo) containing 2 μM protein in unfolding buffer (20 mM HEPES [pH 7.5], 100 mM KCl, 5 mM $MgCl_2$, 1 mM DTT) and 2.5 μL of 1:50 diluted SYPRO Orange dye (Sigma-Aldrich). The assays were performed by heating from 25°C to 95°C at 1°C/min with a 2 min hold at each temperature, and a fluorescence read at each temperature increment. All protein samples were quantified in triplicates. Protein dissociation curves were plotted from the data using GraphPad Prism 9 software.

Preparation of 5′-labeled RNAs

RNA used for electrophoretic mobility shift assays (EMSA) were produced in the T7 RNA polymerase system with DNA template listed in Table S1. Recombinant T7 RNA polymerase protein was expressed and purified as described in Rio.[47] *In vitro* transcription reactions contained 1–2 μg of linearized DNA template, transcription buffer (30 mM Tris-Cl [pH 8.1], 25 mM $MgCl_2$, 0.01% (v/v) Triton X-100, 2 mM spermidine), 5 mM of each rNTPs, 10 mM DTT, and 20 μg of purified T7 RNA polymerase. After 3 h, 2U RQ1 RNase-Free DNase (Promega) and 10 mM $MgCl_2$ were added to the transcription reactions for 1 h, then reactions were terminated by addition of one volume of 2× RNA Gel-Loading Buffer (95% deionized formamide, 5 mM EDTA [pH 8.0], 0.025% SDS, 0.025% bromophenol blue, 0.025% xylene cyanol). Transcription reactions were resolved by 15% denaturing PAGE and corresponding bands containing RNA were extracted from the gel and precipitated with ethanol. The 5′-ends of the RNA were radiolabeled with γ-[^{32}P]ATP (3,000 Ci/mmol, PerkinElmer) using T4 polynucleotide kinase (NEB) and purified from unincorporated label using G-25 spin columns (Cytiva).

EMSA assays

Complexes containing THUMPD1 bound to RNA were prepared by incubating 0 to 1 μM of THUMPD1 and constant amount (0.2 to 1 nM) of 5′-labeled RNA in binding buffer (20 mM HEPES [pH 7.5], 100 mM KCl, 5% glycerol, 0.01% v/v Igepal CA-630, 5 mM $MgCl_2$, 30 μg/mL heparin, 1 mM DTT) for 60 min at 37°C. Complexes assembled in each reaction condition were resolved on a 8% native PAGE gel in a cold room and visualized using a Typhoon 9410 Trio Phosphoimager (Cytiva) after drying and exposing the gel. Radioactive signals were quantitated with ImageQuant TL (Cytiva).

RNA preparation

Total RNA was extracted from human cell lines using TRIzol Reagent (Thermo Fisher Scientific) according to the manufacturer's instructions. For experiments requiring the separation of small (<100 nt) and large (>100 nt) RNA fractions, mirVana miRNA Isolation Kit (Thermo Scientific) was used according to the manufacturer's instructions. RNA concentrations were determined by Nanodrop 1000 (Thermo Scientific).

qRT-PCR

Total or large RNA samples were reverse transcribed using a High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and random hexamer priming, according to manufacturer's instructions. To quantify levels of *THUMPD1* mRNA expression, the PCR reaction was carried out using SsoAdvanced Universal SYBR Green Supermix (Biorad) with the primers listed in Table S1. The reactions were monitored using a QuantStudio 5 Real-Time PCR System (Applied Biosystems). The ΔΔCT method was used to quantify relative *THUMPD1* mRNA expression changes between control cells and *THUMPD1* KO / variant cells, using *GAPDH* as a housekeeping control. We designed the experiment with three biological and three technical replicates to validate the results (n = 3).

Northern blot analysis

A standard northern blot procedure was used to evaluate tRNA or U6 RNA levels from small RNA samples. Briefly, equivalents of 200 ng of RNA in triplicates for each sample were denatured in 1× RNA formamide loading buffer (Thermo Fisher Scientific) at 95°C for 10 min and separated on 10% PAGE gel containing 7M Urea for 45 min at constant 150V, then transferred onto Amersham Hybond N+ nylon membrane (Cytiva) in 1× TBE for 1 h at constant 400 mA. Membranes were air-dried, UV-crosslinked in Stratalinker 1800 (Stratagene/Agilent), and probed with γ-[^{32}P]ATP oligonucleotide specific to tRNA-Ser-CGA (Table S1) overnight at 42°C in ULTRAhyb Ultrasensitive Hybridization Buffer (Ambion). After washing membranes in 2× SSC followed by 0.2 × SSC, the bound probe bands were detected using a Typhoon Prosphoimager. Before re-probing membranes with U6 RNA- or tRNA-Ala-specific γ-[^{32}P]ATP labeled oligonucleotides, the membranes were stripped in 0.5% (w/v) SDS buffer at 60°C. tRNA-probed results were normalized to U6 RNA loading control signal to calculate the relative tRNA levels in control versus *THUMPD1* or *NAT10* KO RNA samples.

Western blot analysis

Protein cell extracts were prepared by lysing the cells on ice for 30 min in 1× PBS containing 1% NP-40 detergent, 1 mM DTT, and protease inhibitor cocktail (Thermo Fisher Scientific) with the addition of DNase I and RNase A. SDS-sample denaturing buffer was added to 1× concentration and samples heated at 95°C for 10 min. 30 μg of total protein was loaded onto 10% SDS-PAGE gel and separated at 200V for 45 min. Protein transfer to Amersham Protran 0.45 μm nitrocellulose membrane (Cytiva) was performed in Trans-Blot Turbo instrument (BioRad) using the turbo-option. Membranes were treated according to antibody manufacturer recommendations. We used the following antibodies: mouse monoclonal α-THUMPD1 (4A11, Thermo Fisher Scientific); rabbit polyclonal α-THUMPD1 (A304-643A, Bethyl Labs); rabbit polyclonal α-GAPDH (G9545, Sigma-Aldrich); and rat monoclonal α-tubulin (ab6160, Abcam) at suggested dilutions. Secondary HRP-conjugated goat anti-mouse and anti-rabbit antibodies were obtained BioRad and goat anti-rat from KPL. To visualize the protein bands on membranes, we used SuperSignal West Pico PLUS Chemiluminescent Substrate reagents from Thermo Fisher Scientific and Molecular Imager GelDoc XR+ (BioRad).

Detection of ac4C modifications by ImmunoNorthern blot

10 μg total RNA and large RNA or 2 μg of small RNA samples were denatured in formaldehyde loading buffer, heated to 65°C for 15 min, and separated on 1% or 1.5% (for small RNA) agarose denaturing gel for 3 h at constant 70V. Gels were stained in 0.5 μg/L ethidium bromide and documented by UV imaging before transfer. RNA was transferred onto Amersham Hybond-N+ membranes (Cytiva) by capillary transfer using 20× SSC buffer overnight. Membranes were crosslinked twice in Stratalinker, blocked with Western Blot Blocking Reagent (BioRad) in 1× PBS containing 0.05% Tween-20 (PBS-T) for 30 min at room temperature, and probed with α-ac4C antibody (ab25215, Abcam) at a 1:1,000 dilution in 5% BSA at 4°C overnight. Then membranes were washed three times with PBS-T and incubated with HRP-conjugated secondary goat anti-rabbit antibody (BioRad) at a 1:5,000 dilution in 1× PBS-T at 4°C overnight followed by four membrane washes with 1× PBS-T and treatment with SuperSignal West Pico PLUS Chemiluminescent Substrate reagents. Chemiluminescence was detected using Molecular Imager GelDoc XR+.

Purification of individual tRNA-Ser-CGA1-1 and tRNA-Ala-AGC9-1 species

Biotin-conjugated single-stranded DNA oligonucleotides containing sequences complementary to tRNA-Ser-CGA1-1 or tRNA-Ala-AGC9-1 (Table S1) were immobilized onto Pierce streptavidin magnetic beads (Thermo Fisher Scientific) as suggested by manufacturer and after a final wash were resuspended in 2.4 M tetraethylammonium chloride (TEAC). 2 mg of total RNA prepared from control or variant-containing lymphoblasts in 2.4 M TEAC was heated at 75°C for 5 min and quickly chilled on ice before addition of 50 μL aliquot of immobilized magnetic beads. The mixture was heated 55°C for 3 min and hybridization was set at 20°C for 90 min. Then magnetic beads were separated using Invitrogen DynaMag-2 Magnet, washed 3 times with 2.4 M TEAC, and bound RNA was eluted by incubation with 100 μL of 2.4M TEAC at 20°C followed by standard precipitation with ethanol. To evaluate the quality of purified tRNAs 5 ng aliquots were ran on 2100 Bioanalyzer (Agilent) at the Genomics Research Center University of Rochester or 20 ng aliquots were separated onto 10% denaturing acrylamide gel and stained with Diamond Nucleic Acid Dye (Promega).

RNA hydrolysis

RNA sample hydrolysis for liquid chromatography-mass spectrometry (LC-MS) was conducted as described in Su et al.[48] Briefly, 10 μg of total RNA or 2 μg of small RNA was treated with a mixture of benzonase nuclease (Sigma-Aldrich), phosphodiesterase (US Biological), and alkaline phosphatase (Sigma-Aldrich) for 3 h at 37°C and then spin-filtered using Amicon 10K MWCO membranes (Millipore) at 16,000 × *g* for 15 min in cold microfuge. Resultant filtrates were analyzed by LC-MS.

LC-MS/MS analysis of RNA modifications

RNA modification analysis was carried out by adapting the method developed by Su et al.,[48] using a Dionex Ultimate 3000 UHPLC coupled to a Q Exactive Plus mass spectrometer (Thermo Scientific). After purification, analytes were separated on a Hypersil Gold 2.1 × 150 mm column, protected by a 2.1 × 10 mm Hypersil Gold guard column (Thermo Scientific). The mobile phases were A: 0.1% formic acid in water, and B: 0.1% formic acid in acetonitrile. The flow rate was set to 400 μL/min, and the column oven was set to 36°C. 8 μL of each sample was injected, and the analytes were eluted using a multi-step gradient as follows: 0.00 min, 0% B; 6.00 min, 0% B; 7.65 min, 1% B; 9.35 min, 6% B; 10.00 min, 6% B; 12.00 min, 50% B; 14.00 min, 75% B; 17.00 min, 75%; 17.50 min, 0%; and

20.00 min, 0%.The Q Exactive Plus was operated in positive mode with a heated electrospray ionization (HESI) source. The spray voltage was set to 3.5 kV, the sheath gas flow rate was set to 40, and the auxiliary gas flow rate set to 7, while the capillary temperature was set to 320°C. A parallel reaction monitoring (PRM) method was used to quantify the modifications. Precursor ions were isolated using a 1.5 m/z isolation width, and then fragmented in the collision cell with an NCE of 40. The maximum injection time was set to 80 ms, while the AGC target was 2e5. Retention time scheduling was employed to ensure enough scans across the peak for each analyte. The resulting fragment ions were detected in the Orbitrap with a resolution of 17,500 at m/z 200. Peak areas from the fragment ions were extracted with a 10 ppm mass tolerance using the LC Quan node of the XCalibur software (Thermo Scientific). All detected unmodified and modified nucleosides, m/z values, and their respective retention time is presented in Table S2.

Statistical analysis

The statistical analyses (Welch's t test, Brown-Forsythe, and Welch one-way ANOVA with Benjamini Hochberg correction, and two-way ANOVA with Benjamini Hochberg correction) were carried out using GraphPad Prism 9 (GraphPad, La Jolla, CA).

Results

Bi-allelic variants of THUMPD1 result in syndromic intellectual disability

We collected 13 individuals from 8 families originating from Europe, and the Middle East (8 males, 5 females) with rare bi-allelic variants in *THUMPD1* (F1–F8). Nine individuals were born from consanguineous parents. Age at last visit ranged from 2 to 23 years old (median: 7 years old). Table 1 summarizes the core clinical features of affected individuals. All evaluated individuals had global developmental delay (1 individual could not be evaluated) and intellectual disability (1 individual could not be evaluated), which varied in severity from mild to severe. All showed a delayed language development. Motor development was less impaired, with a median age of walking of 1.5 years old. Microcephaly was observed in 4/10 of the individuals and individual F5:II.1 had a head circumference of 46 cm at 3.5 years (−1.94 SD). Mild to severe hearing loss was detected in 4 families representing 6/8 of the individuals. Ophthalmological anomalies were detected in all individuals except two, represented by a variety of anomalies (eye position anomalies, strabismus, hypermetropia, esotropia, nystagmus, myopia, bilateral zonular cataract, pale optic discs, patchy pigmented fundus, hyperopia, and/or astigmatism). Seven individuals had brain MRIs with one individual presenting agenesis of the corpus callosum. Individual F8:II.1 brain MRI was compatible with white matter injury of prematurity while her sister (F8:II.3) had a thick and foreshortened corpus callosum, with a subtle deficiency of the rostral end, bulky caudate heads, and unusually deep sella (V shaped). Facial dysmorphism included up/down-slanting palpebral fissures, epicanthus, ptosis, hypo/hypertelorism, wide sparse eyebrows, high nasal root, hypoplastic columella, broad and flat philtrum, thin upper lip, prognathism with mild limitation of jaw opening, wide mouth, low-set ears, sparse hair, scaphocephaly, frontal bossing, and discolored dentition (Figure 1). Behavioral traits associated with the phenotype included hyperactivity, aggressive behavior, and angry outbursts. Other physical features included clinodactyly of both hands, bifid and short uvula, high arched palate, cubitus and genu valgum, mild joint laxity, pes-cavus bilateral, pes-planus, and hypoplastic scrotum.

All affected individuals harbor bi-allelic variants in *THUMPD1* (GenBank: NM_001304550.1) (Figure 2). 12 individuals are homozygous for the following variants: two brothers (F1:II.1, F1:II.2) are homozygous for a frameshift variant, c.495dup (p.Ser166Leufs*24), three (F2:II.1, F2:II.3, F3:II.1) are homozygous for variant c.706C>T (p.Gln236*), one (F4:II.1) is homozygous for variant c.341T>G (p.Leu114*), one (F5:II.1) is homozygous for variant c.303_306del (p.Glu102Leufs*10), three (F7:II.1, F7:II.2, F7:II.3) are homozygous for variant c.774_776del (p.Leu258del), two sisters (F8:II.1, F8.II.3) are homozygous for variant c.634dup (p.Glu212Glyfs*18), and lastly one individual (F6:II.1) has two variants: c.469C>T (p.Arg157*) and c.490C>T (p.Pro164Ser) resulting in a compound heterozygous state. Individual F2:II.6 has a different disease (Rahman syndrome [MIM: 617537]), which is caused by mutations in *HISTH1E,* but is able to attend school with borderline performance, good behavior, and normal development for his age. It should be noted that F2:II.1 and F2:II do not share the Rahman syndrome-associated variant. These *THUMPD1* variants were absent or very rare in the gnomAD Exome database (maximum allele frequency of 0.00001998). All variants lie either within the THUMP domain or the predicted α-helical N-terminal region of THUMPD1 (Figure 2B). Additionally, we identified a family with three affected brothers carrying the homozygous variant c.133G>A (p.Gly45Ser) (Table S3, family 9). Although their phenotype is in line with the other identified individuals, the missense variant is outside the known functional domains of THUMPD1, so in absence of functional data we still consider this variant of unknown significance.

Lymphoblasts from an individual homozygous for a c.706C>T variant exhibit a loss of THUMPD1 mRNA and protein expression, and a loss of tRNA acetylation

We obtained the lymphoblasts from one affected individual (F2:II.1; 15DG1395 from Maddirevula et al.[43]) who is homozygous for *THUMPD1* variant c.706C>T (p.Gln236*). Western blot analysis of the cell extracts prepared from these lymphoblasts showed that no full-length THUMPD1 was produced compared to two unrelated control individuals (Figure 3A). This is expected due to the p.Gln236* premature stop codon and because we probed with an anti-THUMPD1 antibody that was raised against C-terminal epitope (amino acids 303–353) of THUMPD1. Moreover, RT-qPCR analysis showed a significant decrease of *THUMPD1* mRNA levels in *THUMPD1* variant cells,

Table 1. Summary of phenotype information for NDD-affected individuals with *THUMPD1* bi-allelic pathogenic variants

Case index	F1:II.1	F1:II.2	F2:II.1	F2:II.2	F3:II:1	F4:II.1	F5:II.1	F6:II.1	F7:II.1	F7:II.2	F7:II.3	F8:II.1	F8:II.3	Summary
Age at last examination (y)	14	7.5	23	16	8	2	3.5	4	10	16	19	7	14	
Sex	M	M	M	M	F	M	F	M	M	F	M	F	F	
Consanguinity	–	–	+	+	+	N/A	+	–	+	+	+	+	+	9/12
Neurological problems														
Developmental delay	+	+	+	+	+	+	+	+	+	+	N/A	+	+	12/12
Microcephaly	–	–	–	N/A	+	+	–	–	+	–	N/A	+	N/A	4/10
Delayed sentences	+	+	+	+	+	N/A	+	+	+	N/A	N/A	+	+	10/10
Age at first words (y)	N/A	2	N/A	N/A	1	N/A	non verbal	2	2	N/A	N/A	N/A	2.5	
Age of walking (m)	17	16	24	N/A	18	N/A	36	18	36	N/A	N/A	19	15	
Intellectual disability	+	+	+	+	+	N/A	+	+	+	+	+	+	+	12/12
Behavior abnormality	N/A	N/A	+	+	+	N/A	+	+	+	+	N/A	–	–	7/9
Febrile seizures	N/A	+	–	–	+	N/A	–	–	+	+	+	–	+	6/11
Brain MRI abnormality	N/A	N/A	N/A	N/A	–	+	–	–	–	N/A	N/A	+	+	3/7
Systemic problems														
Short stature	–	–	–	N/A	+	N/A	–	–	+	–	N/A	+	–	3/10
Ophtalmological abnormality	+	+	+	–	+	+	+	+	+	+	+	+	–	11/13
Congenital heart defect	–	–	N/A	N/A	N/A	+	N/A	N/A	N/A	N/A	N/A	+	N/A	2/4
Hearing impairment	+	+	N/A	N/A	N/A	N/A	N/A	+	–	+	–	+	+	6/8

Symbols and abbreviations: AR, autosomal recessive; CH, compound heterozygous; N/A, no data or not reported; ADHD, attention deficit hyperactivity disorder

expressing ~20% of control #1 *THUMPD1* mRNA levels (Figure 3B). Using the most abundant *THUMPD1* transcript based on GTEx (GenBank: NM_001304550.1), we note that the premature terminal codon (PTC) for this variant resides in the second-last exon with the final intron within the 3′ UTR, and these kind of PTC-containing transcripts have been generally shown in most cases (but not all) to escape nonsense-mediated decay.[49] Thus, to more thoroughly test whether this variant (p.Gln236*) is an exception to this general rule, we carried out additional qPCR analysis comparing the expression levels of *THUMPD1* pre-mRNA to mature mRNA in control and p.Gln236* variant samples and found no significant difference in pre-mRNA levels but again a reduction in mature mRNA levels in the affected individual (Figure S1). This is highly indicative that this transcript is unable to escape NMD. This also suggests that other PTC or frameshift variants that reside in this second to last exon (e.g., variant p.Glu212Glyfs*18 found in F8:II.1 and F8.II.3) may also be sensitive to NMD.

Given THUMPD1's role in tRNA acetylation, we next examined whether the lack of full-length THUMPD1 may affect acetylation of RNA in THUMPD1 GenBank: NP_001291479.1 (p.Gln236*) variant lymphoblasts. We extracted and digested small RNAs (predominantly containing tRNAs and small nuclear RNAs) from THUMPD1 p.Gln236* variant and control individual lymphoblasts and subjected the extracts to liquid chromatography coupled with mass spectrometry analysis (LC-MS) (see Material and methods). This analysis showed that compared to control individual samples, the THUMPD1 p.Gln236* variant individual small RNA sample contained no detectable level of ac4C (Figures 3C and S2) while most other normal and modified nucleosides remained at similar levels to control sample levels (Figure 3C). A complete list of modified nucleosides identified by LC-MS in our experiments is presented in Table S2. To further confirm that the loss of ac4C in this sample was specific to a known NAT10/THUMPD1 tRNA substrate, we specifically purified tRNA-Ser-CGA and tRNA-Ala-AGC (as a control tRNA species known not to contain ac4C) from control and THUMPD1 p.Gln236* variant lymphoblasts, performed LC-MS analysis of hydrolyzed RNA, and observed a >100-fold loss of ac4C signal in the tRNA-Ser-CGA sample from THUMPD1 p.Gln236* variant cells when compared a tRNA-Ser-CGA sample from control cells (Figure 3D). As expected, control tRNA-Ala-AGC samples had similar, background-level ac4C values to the THUMPD1 p.Gln236* variant cells (Figure S3). Taken together, our observations suggest that a loss of THUMPD1 protein expression results in a loss of ac4C modification of small RNAs.

Targeted THUMPD1 knockouts in human cell lines results in a loss of ac4C modification of small RNAs

To ensure that the loss of THUMPD1 in individuals with bi-allelic *THUMPD1* variants specifically affects small RNA

Figure 1. Clinical features of individuals with bi-allelic THUMPD1 variants
Photographs of five individuals from our cohort, showing facial dysmorphisms. Each individual is referenced with the corresponding identifier used throughout the manuscript: Individuals F1:II:1 (A, B) and F1:II:2 (C) are brothers and show downslanting palpebral fissures, broad and flat philtrum, hypoplastic columella, and low-set ears. Individual F6:II:1 (F) has a wide mouth, epicanthus, ptosis, and narrow palpebral fissures. Individual F8:II.1 (G, H) shows scaphocephaly, frontal bossing, deep-set ears, hypotelorism, and intermittent bilateral esotropia.

ac4C modification and was not due to any indirect secondary effects, we generated *THUMPD1* knockouts (KO) in HEK293T and HeLa cell lines. *THUMPD1* KO cell lines were generated using a gRNA targeting the coding sequence in exon 2 (HEK293T, Figure 4A; HeLa, Figure S4A), cloned cell colonies were subjected to sequence analysis to detect premature termination codon (PTC)-inducing indels (HEK293T, Figure 4B; HeLa, Figure S4B) and appropriate clones in each background were selected for further experiments (HEK293T, Figure 4C; HeLa, Figure S4C). *THUMPD1* KO cells were tested for THUMPD1 abundance via western blot and we observed a clear absence of the full-length THUMPD1 protein in *THUMPD1* KO cells compared to WT HEK293T and HeLa cells (HEK293T, Figure 4D; HeLa, Figure S4D). In addition, RT-qPCR showed that *THUMPD1* mRNA levels in both KO cells were substantially reduced, exhibiting ~20% of the expression levels observed in WT cell lines (HEK293T, Figure 4E; HeLa, Figure S4E), very similar to what was observed in THUMPD1 p.Gln236* variant lymphoblasts.

Importantly, much like what was observed in THUMPD1 p.Gln236* variant lymphoblasts, *THUMPD1* disruption in HEK293T and HeLa cells led to a complete loss of ac4C modification of small RNAs. Here we used two independent procedures to confirm this observation. First, immuno-northern blots of large and small RNA samples prepared from HEK293T WT and KO cells (Figure S5) probed with an anti-ac4C antibody demonstrated a loss of ac4C modification signal of small RNAs from *THUMPD1* KO cells (Figure 4F). In contrast, *THUMPD1* KO did not affect the ac4C status of 18S rRNA species present in the large RNA sample, similar to what is observed in yeast.[31,35] As a control, we also showed that in NAT10 KO HeLa cells there is a complete loss of ac4C modification in both rRNA and tRNA (Figure S4F). Second, LC-MS analysis of hydrolyzed small RNA samples prepared from WT and KO cells revealed that *THUMPD1* KO cells exhibited an undetectable level of ac4C modification in small RNAs while most of the other nucleoside modifications remained unchanged as compared to WT cell samples (HEK293T, Figure 4G; HeLa, Figure S4G); this contrasts with the levels of ac4C in the large RNA fraction, which were much less affected (Figure S6). Next, to determine whether the loss of ac4C modification on tRNA-Ser-CGA results in a downstream reduction of tRNA-Ser-CGA levels, we carried out northern blot experiments on small RNA samples from HeLa WT, *THUMPD1* KO, and *NAT10* KO with specific probes that recognize tRNA-Ser-CGA, or tRNA-Ala-AGC or U6 snRNA as a control. Interestingly, we did not detect any large differences in steady-state levels of tRNA-Ser-CGA in either KO samples (Figure S7). It is possible that in HeLa cells under these growth conditions, the Rapid tRNA Decay (RTD) surveillance pathway is not measurably active against hypoacetylated tRNA-Ser-CGA.

Taken together, the specific knockout of THUMPD1 in HEK293T and HeLa cell lines recapitulates our observation of the substantial loss of ac4C modification of small RNAs

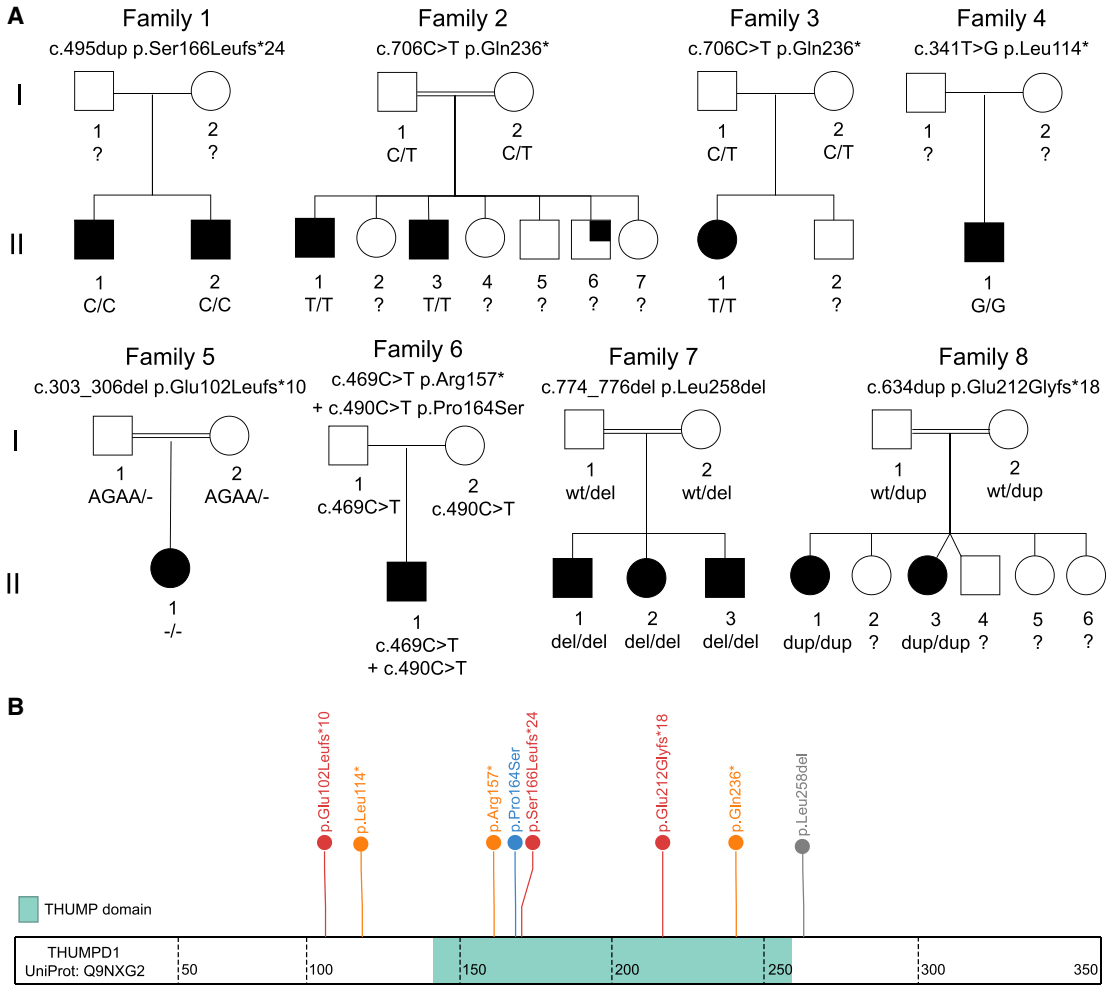

Figure 2. Family pedigrees of individuals with bi-allelic THUMPD1 variants and a schematic of THUMPD1 with positions of the corresponding mutated amino acid residues indicated

(A) Pedigrees of eight unrelated families with affected members indicated as filled circles (females) and squares (males) with denoted *THUMPD1* variants and the corresponding protein alterations. Double horizontal lines indicate consanguinity. Individual F2:II.6 has a different disease (Rahman syndrome) but is able to attend school with borderline performance, good behavior, and normal development for his age.

(B) Schematic representation of the human THUMPD1 protein sequence with THUMP-domain highlighted in green. The nature and positions of the variants are shown above the primary sequence diagram (colored blue for missense mutations, orange for nonsense mutations, and red for frameshift variants).

and tRNA-Ser-CGA present in THUMPD1 p.Gln236* variant lymphoblasts obtained from individual F2:II.1 (15DG1395) in our study.

Recombinant THUMPD1 p.Pro164Ser exhibits a marked decrease in thermal stability and RNA-binding capacity

A single missense variant was identified in our cohort, p.Pro164Ser in individual F5.II.1. Pro164 is strictly conserved within THUMPD1 orthologs across eukaryotes (Figure 5A) and in architecturally similar THUMP domain-containing proteins in archaea,[36] suggesting it is important for protein structure and function.

To this end, we overexpressed and purified recombinant WT and mutant p.Pro164Ser THUMPD1 to homogeneity (Figure 5B) and tested thermal stability and RNA-binding capacity, using fluorescent thermal melt assays and electrophoretic mobility shift assays (EMSAs), respectively. We observed that THUMPD1 p.Pro164Ser exhibited a marked decrease in protein thermal stability with an ~5°C decrease in melting temperature relative to WT THUMPD1 (Figures 5C and 5D), suggesting that the p.Pro164Ser substitution results in less stably folded THUMPD1. To test whether this loss of stability affects THUMPD1's capacity to bind RNA, we carried out RNA-binding EMSA assays against an *in vitro* transcribed, radiolabelled tRNA-Ser-CGA (with a CCA tail added during transcription). While WT THUMPD1 bound to tRNA-Ser-CGA with an approximate nanomolar affinity, we observed a significant loss in RNA binding capacity for the p.Pro164Ser mutant (Figures 5E and 5F), indicating that this variant may not be able to efficiently bind its tRNA targets *in vivo* for subsequent tRNA acetylation. It should also be noted that to our knowledge

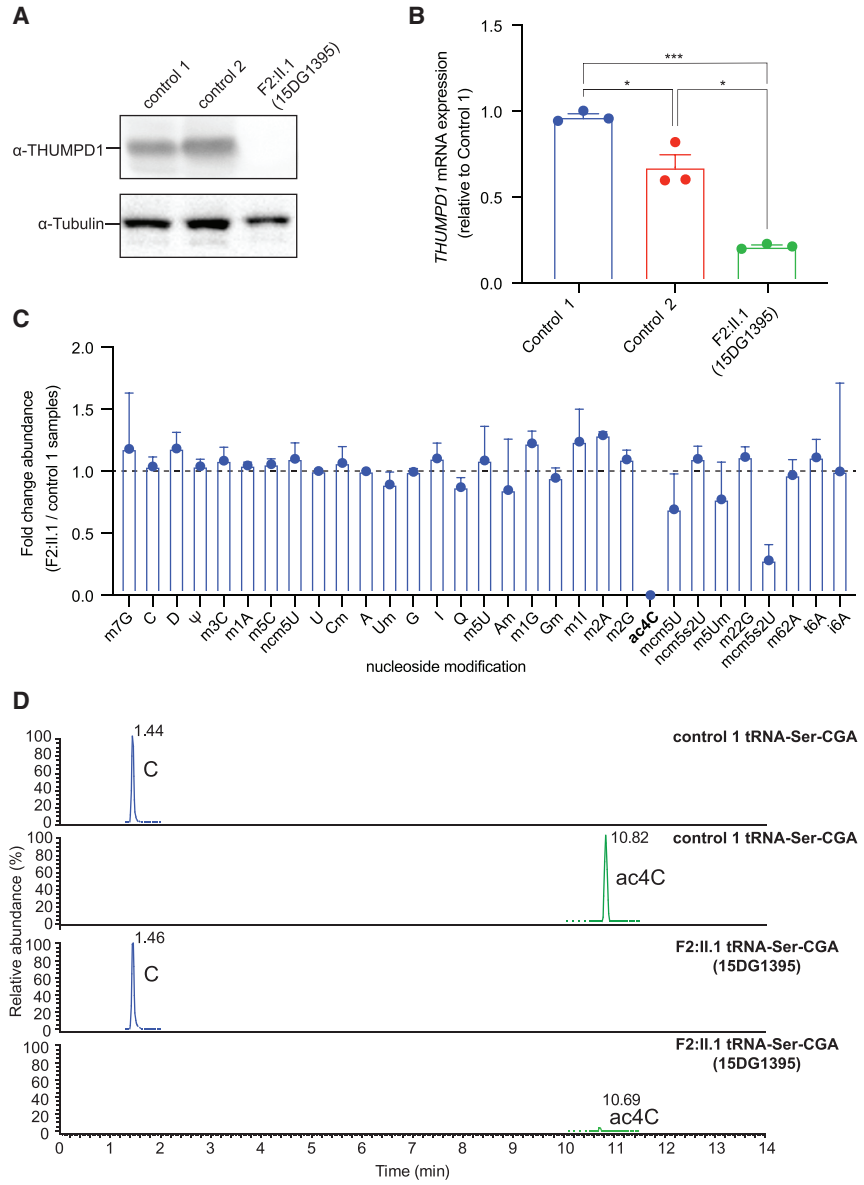

Figure 3. Bi-allelic c.706C>T THUMPD1 (p.Gln236*) causes a lack of ac4C modification in tRNA

(A) Western blot analysis of the protein cell extracts prepared from two control individuals and individual F2:II.1 (15DG1395) lymphoblasts probed with anti-THUMPD1 antibody and anti-tubulin antibody (as loading control).

(B) *THUMPD1* mRNA expression in two control individuals and individual F2:II.1 (15DG1395) lymphoblasts as determined by RT-qPCR of the corresponding RNA samples (relative to control #1). Samples were analyzed in triplicate, error bars represent ± SD. *p(adjusted) < 0.05, ***p(adjusted) < 0.001 as determined by a Brown-Forsythe and Welch one-way ANOVA with Benjamini-Hochberg correction (FDR < 5%).

(C) LC-MS nucleoside modification analysis of small RNA samples from individual F2:II.1 (15DG1395) lymphoblasts compared (normalized) to control #1 small RNA samples. Alterations (fold change) for each nucleoside modification was calculated as a ratio of the corresponding value in variant-containing RNA to control sample RNA normalized based on the abundance of all four standard non-modified nucleosides (A, C, G, and U). Samples were run in duplicate, error bars represent ± SD.

(D) LC-MS nucleoside modification analysis of the purified tRNA-Ser-CGA1-1 samples prepared from control #1 (two top panels) and individual F2:II.1 (15DG1395) (two bottom panels) lymphoblasts. The ac4C peaks are shown in green and unmodified cytidine peaks are in blue (as controls). LC elution times are shown above each peak.

this is the first demonstration of tRNA binding by WT human THUMPD1. Taken together, these data suggest that the variant p.Pro164Ser results in a loss of function of THUMPD1. Given that this variant is absent from gnomAD database and was identified in compound heterozygous state with p.Arg157* in individual F5.II.1, we consider it as pathogenic.

Discussion

In this study, we identified bi-allelic variants in *THUMPD1* in 13 individuals presenting with a neurodevelopmental disorder. Given the genotypic and phenotypic similarities in our case series, we propose that these bi-allelic variants in *THUMPD1* should be considered as pathogenic. In addition, we present functional studies that indicate *THUMPD1* loss of function in these individuals. We demonstrate that variants identified cause loss of function of THUMPD1 and that this defect results in a loss of ac4C modification in small RNAs and of individually purified tRNA-Ser-CGA. We further corroborate this effect by showing a loss of tRNA acetylation in two CRISPR-Cas9-generated *THUMPD1* KO cell lines. Taken together, these results suggest that the lack of tRNA acetylation due to THUMPD1 loss of function results in syndromic intellectual disability.

tRNA modifications play an important role in neurodevelopment. For example, loss of modifications in the anticodon loops of tRNA can result in perturbations at all stages of protein translation (initiation, decoding efficiency and fidelity, termination, and so on), while loss of a range of tRNA modifications in the body of tRNA appear to primarily result in changes in tRNA tertiary structure and stability, and thus also tend to result in similar, if more indirect, perturbations of protein translation (see Ramos and Fu[11] for an extensive review). We speculate that phenotypes observed in this study's individuals are a result of a similar phenomenon by which loss of tRNA modification may result in impaired protein synthesis and thus altered proteostasis at important stages of neurodevelopment. Furthermore, the individuals

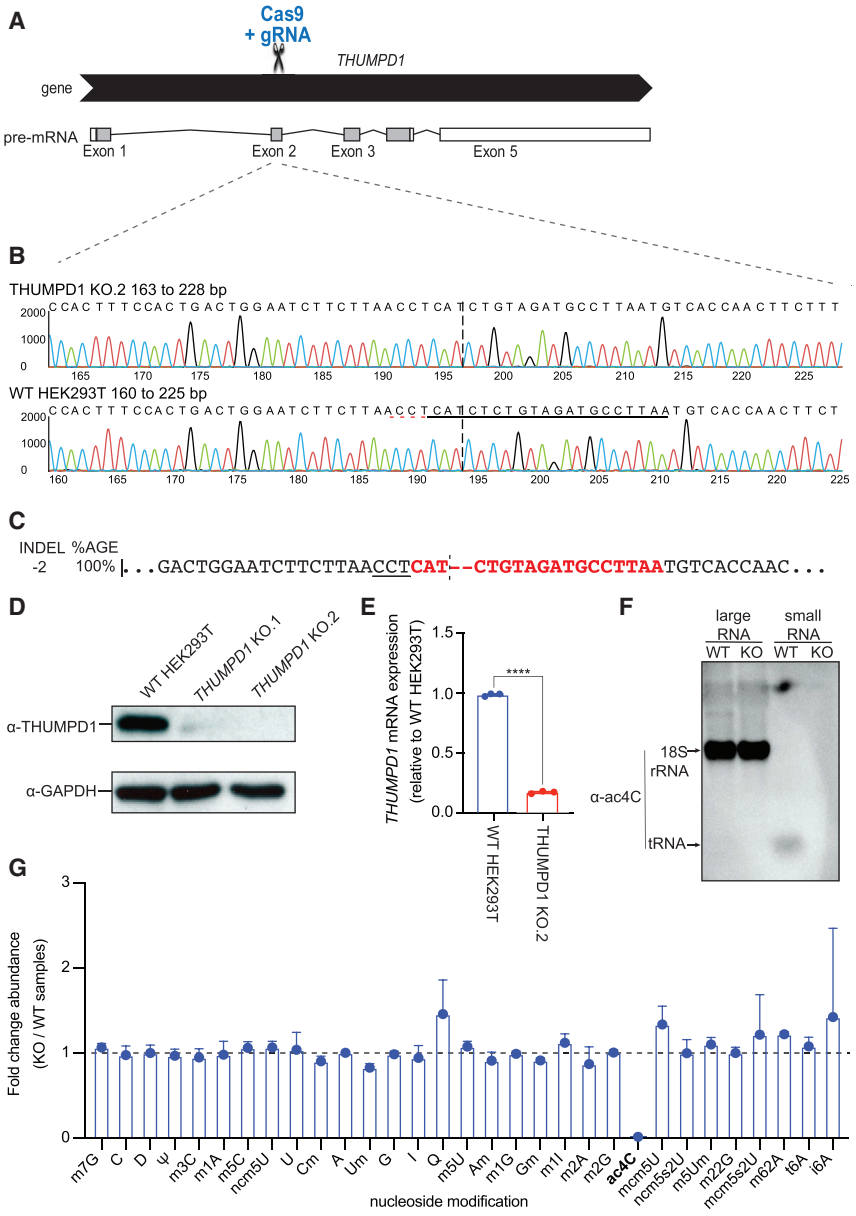

Figure 4. THUMPD1 KO in HEK293T cells causes loss of ac4C modification in small RNA samples

(A) Schematic representation of the *THUMPD1* KO in HEK293T cell line. sgRNA was designed to disrupt *THUMPD1* nucleotide sequence in exon 2 to produce an aberrant PTC-containing transcript (see Table S1 for gRNA sequence).

(B) Sanger DNA sequencing data of the *THUMPD1* exon 2 genomic region obtained from selected *THUMPD1* KO clone (top) and WT HEK293T cells (bottom). Underlined DNA sequence in WT HEK293T sample is complementary to the designed sgRNA. Dotted vertical line in both sequences indicates a Cas9 target cleavage site.

(C) ICE analysis (Synthego) results of the *THUMPD1* exon 2 genomic DNA region from a selected *THUMPD1* KO clone. Target nucleotide sequence complementary to sgRNA is in red font and dotted vertical line indicates a Cas9 target cleavage site. The indel type and percentage of detected sequence are shown on the left.

(D) Western blot analysis of the protein cell extracts prepared from HEK293T control and two independent *THUMPD1* KO cell lines probed with anti-THUMPD1 antibody and anti-GAPDH antibody (as loading control).

(E) *THUMPD1* mRNA expression in HEK293T control and *THUMPD1* KO cells determined by RT-qPCR of the corresponding RNA samples. Samples were analyzed in triplicate, error bars represent ± SD. ****$p < 0.0001$, as determined by a Welch's t test.

(F) Northern immunoblot analysis of the large and small RNA samples prepared from HEK293T control (WT) and *THUMPD1* KO cells. Positions of 18S rRNA and tRNA on the gel are shown by the arrows on the left.

(G) LC-MS nucleoside modification analysis of the small RNA samples prepared from HEK293T control and *THUMPD1* KO cells. Alterations (fold change) for each nucleoside modification was calculated as a ratio of the corresponding value in *THUMPD1* KO RNA to HEK293T control RNA normalized based on the abundance of all standard non-modified nucleosides (A, C, G, and U). Samples were run in duplicate, error bars represent ± SD.

in our study share a number of phenotypes common to diseases where mutations in aminoacyl-tRNA synthetases have been identified (see Ognjenović and Simonović [50] for an extensive review), further highlighting that perturbation of protein translation can result in global development delay and intellectual disability. It is also possible that the lack of tRNA-Ser or tRNA-Leu acetylation affects translation of only a limited subset of proteins that are important for neurodevelopment. Indeed, it is well understood that neurons are particularly sensitive to alterations in protein synthesis, and the multi-tissue effects we observe in the individuals in this study could be a result of the complex developmental interplay and feedback between multiple tissues and the brain that occur during development (for review see Kapur et al.[51]).

Specifically, given ac4C's position within the body of Ser- and Leu-tRNAs (in the D-stem of the tRNA) and its ability to increase the base pairing strength between G:C,[52,53] it is possible that ac4C helps tRNAs obtain the correct cloverleaf fold and tertiary contacts, and the lack of this modification likely results in a less stable tRNA that may be more prone to degradation. Indeed, in *S. cerevisiae,* the deletion of *TAN1* (*THUMPD1* ortholog) resulted in loss of ac4C modifications on Ser- and Leu-tRNAs, causing an increased sensitivity of tRNA-Ser-CGA, tRNA-Ser-UGA, and tRNA-Leu-GAG to the rapid tRNA decay pathway that ultimately led to reduced levels of these tRNA species.[35,40–42,54] This effect is mild in a single *tan1*Δ deletion strain and is significantly exacerbated when a *tan1*Δ deletion is coupled with either a mutant

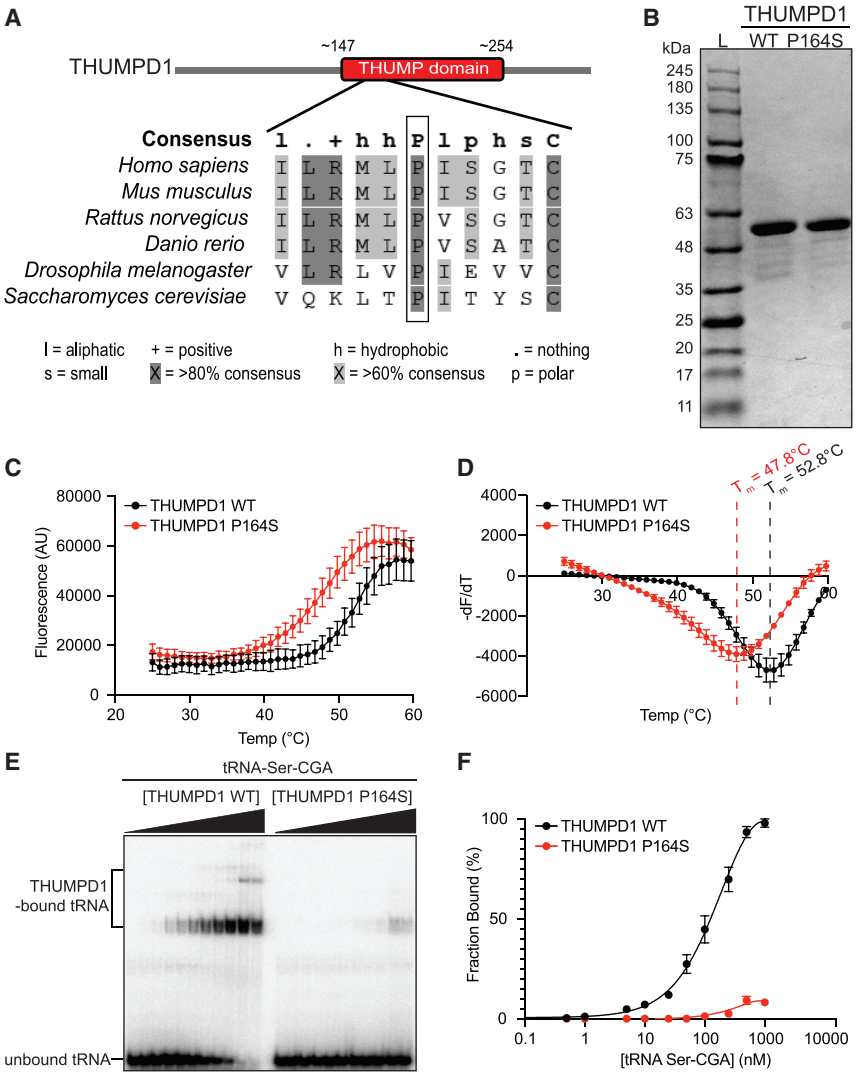

Figure 5. THUMPD1 p.Pro164Ser affects its stability and binding to RNA

(A) Proline at position 164 (outlined with a black box) within THUMP domain (positions 147–254 [in red] is highly conserved in primary sequences of THUMPD1 orthologs), including model organisms *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Danio rerio*, *Drosophila melanogaster,* and *Saccharomyces cerevisiae*.

(B) Coomassie blue-stained gel showing purified WT and p.Pro164Ser (P164S) mutant proteins. Protein molecular marker sizes are shown on the left side of the gel.

(C) Thermal denaturation profile from a protein thermal shift assay of purified THUMPD1 WT (black line) and P164S mutant (red line) proteins.

(D) A plot first derivative of the fluorescence emission as a function of temperature (−dF/dT) (from the data in C). The protein melting temperature (T_m) of the P164S mutant protein (red line) is ∼5°C lower than the WT protein (black line).

(E) EMSA gel image testing the binding of purified THUMPD1 WT protein (left) and P164S mutant protein (right) to tRNA-Ser-CGA. The ascending triangles in black show an increasing concentration of THUMPD1 in the EMSA assay samples from 0 nM to 1,000 nM.

(F) Quantitation of the EMSA of THUMPD1 binding to tRNA-Ser-CGA from (E) showed that purified THUMPD1 WT protein (black line) had substantially higher affinity/binding to tRNA than P164S mutant (red line). Samples were run in duplicate, error bars represent ± SD.

tRNA-Ser-CGA[35] or *trm44* deletion (which encodes a tRNA 2′-O-methyltransferase that methylating position U44).[40] Both of these double mutants possess a temperature-sensitive growth phenotype. In the case of *THUMPD1* and *NAT10* KO HeLa cell lines (Figure S5) and THUMPD1 p.Gln238* variant lymphoblasts (data not shown) used in this study, we did not observe any dramatic changes in tRNA-Ser-CGA expression relative to WT cells. We speculate that either changes in tRNA-Ser-CGA expression due to tRNA hypoacetylation in these cell lines are too minor to be detected, but could be more pronounced at certain neurodevelopmental stages and/or cell types in humans, or when cells are undergoing stress; or unlike the slight reduction in tRNA-Ser-CGA levels Δ*tan1* yeast, the loss of tRNA acetylation does not result in a reduction of corresponding tRNA levels in humans. With this in mind, we cannot disregard that ac4C may have additional roles beyond stabilizing the structure of certain tRNAs and may be required for additional tRNA processing, aminoacylation, recognition, or protein translation processes. Relatedly, it has been previously shown using large-scale RNA-binding protein mass spectrometry screens that THUMPD1 may broadly interact with mRNA[55] and specific pre-microRNAs,[56] although further work needs to be done to confirm whether these observations are physiologically relevant.

Taken together, we showed that bi-allelic loss-of-function variants in *THUMPD1* can be added to the growing list of gene variants known to affect tRNA modification and interfere with normal neurodevelopment. Additional research is required to understand the cellular response to the loss of tRNA acetylation (and indeed many other tRNA modifications) and how it results in altered neurodevelopment. Finally, our work also highlights that quantitative LC-MS is becoming an increasingly powerful tool to dissect changes in RNA modifications in both research and potentially clinical settings.[57] The ability to detect a large number of modified nucleosides in parallel enables a more systems-level understanding of RNA modification dynamics in health and disease, and may potentially assist in the development of novel biomarker approaches for research and diagnosis of genetic disorders involving RNA modification, and aid in connecting genotype to phenotype.

Data and code availability

Variants were submitted to Clinvar with accession numbers Clinvar: SCV002056133, Clinvar: SCV002056134, Clinvar: SCV002056135, Clinvar: SCV002056136, Clinvar: SCV002056137, Clinvar: SCV002056138, Clinvar: SCV002056139, Clinvar: SCV002056140, and Clinvar: SCV002056141.

Supplemental information

Supplemental information can be found online at https://doi.org/10.1016/j.ajhg.2022.02.001.

Acknowledgments

We thank participating families for their cooperation. We thank members of the O'Connell lab for helpful discussions. We thank Dr. Shalini Oberdoerffer for providing NAT10 KO HeLa cells. We thank Kevin Welle and the University of Rochester Mass Spectrometry Resource Laboratory for invaluable assistance with nucleoside mass spectrometry experiments. This work used the Typhoon RGB scanner at the Center for RNA Biology, University of Rochester, supported by the NIH S10 OD021489-01A1 instrumentation grant. M.R.O. is supported by the National Institute of General Medical Sciences grant R35GM133462. Sequencing and analysis for subjects from family 8 were provided by the Broad Institute of MIT and Harvard Center for Mendelian Genomics (Broad CMG) and was funded by the National Human Genome Research Institute, the National Eye Institute, and the National Heart, Lung and Blood Institute grant UM1 HG008900 and in part by National Human Genome Research Institute grant R01 HG009141.

Declaration of interests

The authors declare no competing interests.

Received: October 7, 2021
Accepted: February 1, 2022
Published: February 22, 2022

Web resources

ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
GenBank, https://www.ncbi.nlm.nih.gov/genbank/
OMIM, https://www.omim.org/

References

1. El Yacoubi, B., Bailly, M., and de Crécy-Lagard, V. (2012). Biosynthesis and function of posttranscriptional modifications of transfer RNAs. Annu. Rev. Genet. *46*, 69–95.
2. Ontiveros, R.J., Stoute, J., and Liu, K.F. (2019). The chemical diversity of RNA modifications. Biochem. J. *476*, 1227–1245.
3. Pan, T. (2018). Modifications and functional genomics of human transfer RNA. Cell Res. *28*, 395–404.
4. de Crécy-Lagard, V., and Jaroch, M. (2021). Functions of Bacterial tRNA Modifications: From Ubiquity to Diversity. Trends Microbiol. *29*, 41–53.
5. Suzuki, T. (2021). The expanding world of tRNA modifications and their disease relevance. Nat. Rev. Mol. Cell Biol. *22*, 375–392.
6. Phizicky, E.M., and Hopper, A.K. (2015). tRNA processing, modification, and subcellular dynamics: past, present, and future. RNA *21*, 483–485.
7. Klassen, R., Bruch, A., and Schaffrath, R. (2020). Induction of protein aggregation and starvation response by tRNA modification defects. Curr. Genet. *66*, 1053–1057.
8. Ranjan, N., and Leidel, S.A. (2019). The epitranscriptome in translation regulation: mRNA and tRNA modifications as the two sides of the same coin? FEBS Lett. *593*, 1483–1493.
9. Agris, P.F., Narendran, A., Sarachan, K., Väre, V.Y.P., and Eruysal, E. (2017). The Importance of Being Modified. The Enzymes (Elsevier), pp. 1–50.
10. Väre, V.Y., Eruysal, E.R., Narendran, A., Sarachan, K.L., and Agris, P.F. (2017). Chemical and Conformational Diversity of Modified Nucleosides Affects tRNA Structure and Function. Biomolecules *7*, 29.
11. Ramos, J., and Fu, D. (2019). The emerging impact of tRNA modifications in the brain and nervous system. Biochim. Biophys. Acta. Gene Regul. Mech. *1862*, 412–428.
12. Khan, M.A., Rafiq, M.A., Noor, A., Hussain, S., Flores, J.V., Rupp, V., Vincent, A.K., Malli, R., Ali, G., Khan, F.S., et al. (2012). Mutation in NSUN2, which encodes an RNA methyltransferase, causes autosomal-recessive intellectual disability. Am. J. Hum. Genet. *90*, 856–863.
13. Abdelrahman, H.A., Al-Shamsi, A.M., Ali, B.R., and Al-Gazali, L. (2018). A null variant in *PUS3* confirms its involvement in intellectual disability and further delineates the associated neurodevelopmental disease. Clin. Genet. *94*, 586–587.
14. Shaheen, R., Han, L., Faqeih, E., Ewida, N., Alobeid, E., Phizicky, E.M., and Alkuraya, F.S. (2016). A homozygous truncating mutation in PUS3 expands the role of tRNA modification in normal cognition. Hum. Genet. *135*, 707–713.
15. Alazami, A.M., Hijazi, H., Al-Dosari, M.S., Shaheen, R., Hashem, A., Aldahmesh, M.A., Mohamed, J.Y., Kentab, A., Salih, M.A., Awaji, A., et al. (2013). Mutation in *ADAT3*, encoding adenosine deaminase acting on transfer RNA, causes intellectual disability and strabismus. J. Med. Genet. *50*, 425–430.
16. El-Hattab, A.W., Saleh, M.A., Hashem, A., Al-Owain, M., Asmari, A.A., Rabei, H., Abdelraouf, H., Hashem, M., Alazami, A.M., Patel, N., et al. (2016). *ADAT3*-related intellectual disability: Further delineation of the phenotype. Am. J. Med. Genet. A. *170A*, 1142–1147.
17. Lentini, J.M., Alsaif, H.S., Faqeih, E., Alkuraya, F.S., and Fu, D. (2020). DALRD3 encodes a protein mutated in epileptic encephalopathy that targets arginine tRNAs for 3-methylcytosine modification. Nat. Commun. *11*, 2510.
18. Zhang, K., Lentini, J.M., Prevost, C.T., Hashem, M.O., Alkuraya, F.S., and Fu, D. (2020). An intellectual disability-associated missense variant in TRMT1 impairs tRNA modification and reconstitution of enzymatic activity. Hum. Mutat. *41*, 600–607.
19. Shaheen, R., Mark, P., Prevost, C.T., AlKindi, A., Alhag, A., Estwani, F., Al-Sheddi, T., Alobeid, E., Alenazi, M.M., Ewida, N., et al. (2019). Biallelic variants in *CTU2* cause DREAM-PL syndrome and impair thiolation of tRNA wobble U34. Hum. Mutat. *40*, 2108–2120.
20. Monies, D., Vågbø, C.B., Al-Owain, M., Alhomaidi, S., and Alkuraya, F.S. (2019). Recessive Truncating Mutations in ALKBH8 Cause Intellectual Disability and Severe Impairment of Wobble Uridine Modification. Am. J. Hum. Genet. *104*, 1202–1209.
21. de Brouwer, A.P.M., Abou Jamra, R., Körtel, N., Soyris, C., Polla, D.L., Safra, M., Zisso, A., Powell, C.A., Rebelo-Guiomar,

P., Dinges, N., et al. (2018). Variants in PUS7 Cause Intellectual Disability with Speech Delay, Microcephaly, Short Stature, and Aggressive Behavior. Am. J. Hum. Genet. *103*, 1045–1052.
22. Shaheen, R., Tasak, M., Maddirevula, S., Abdel-Salam, G.M.H., Sayed, I.S.M., Alazami, A.M., Al-Sheddi, T., Alobeid, E., Phizicky, E.M., and Alkuraya, F.S. (2019). PUS7 mutations impair pseudouridylation in humans and cause intellectual disability and microcephaly. Hum. Genet. *138*, 231–239.
23. Shaheen, R., Abdel-Salam, G.M.H., Guy, M.P., Alomar, R., Abdel-Hamid, M.S., Afifi, H.H., Ismail, S.I., Emam, B.A., Phizicky, E.M., and Alkuraya, F.S. (2015). Mutation in WDR4 impairs tRNA m(7)G46 methylation and causes a distinct form of microcephalic primordial dwarfism. Genome Biol. *16*, 210.
24. Ohashi, Z., Murao, K., Yahagi, T., Von Minden, D.L., McCloskey, J.A., and Nishimura, S. (1972). Characterization of C+ located in the first position of the anticodon of Escherichia coli tRNA Met as N4 - acetylcytidine. Biochim. Biophys. Acta *262*, 214.
25. Gupta, R. (1984). Halobacterium volcanii tRNAs. Identification of 41 tRNAs covering all amino acids, and the sequences of 33 class I tRNAs. J. Biol. Chem. *259*, 9461–9471.
26. Bruenger, E., Kowalak, J.A., Kuchino, Y., McCloskey, J.A., Mizushima, H., Stetter, K.O., and Crain, P.F. (1993). 5S rRNA modification in the hyperthermophilic archaea Sulfolobus solfataricus and Pyrodictium occultum. FASEB J. *7*, 196–200.
27. McCarroll, R., Olsen, G.J., Stahl, Y.D., Woese, C.R., and Sogin, M.L. (1983). Nucleotide sequence of the Dictyostelium discoideum small-subunit ribosomal ribonucleic acid inferred from the gene sequence: evolutionary implications. Biochemistry *22*, 5858–5868.
28. Thomas, G., Gordon, J., and Rogg, H. (1978). N4-Acetylcytidine. A previously unidentified labile component of the small subunit of eukaryotic ribosomes. J. Biol. Chem. *253*, 1101–1105.
29. Ito, S., Horikawa, S., Suzuki, T., Kawauchi, H., Tanaka, Y., Suzuki, T., and Suzuki, T. (2014). Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). J. Biol. Chem. *289*, 35724–35730.
30. Ito, S., Akamatsu, Y., Noma, A., Kimura, S., Miyauchi, K., Ikeuchi, Y., Suzuki, T., and Suzuki, T. (2014). A single acetylation of 18 S rRNA is essential for biogenesis of the small ribosomal subunit in Saccharomyces cerevisiae. J. Biol. Chem. *289*, 26201–26212.
31. Sharma, S., Langhendries, J.-L., Watzinger, P., Kötter, P., Entian, K.-D., and Lafontaine, D.L.J. (2015). Yeast Kre33 and human NAT10 are conserved 18S rRNA cytosine acetyltransferases that modify tRNAs assisted by the adaptor Tan1/THUMPD1. Nucleic Acids Res. *43*, 2242–2258.
32. Zachau, H.G., Dütting, D., and Feldmann, H. (1966). The structures of two serine transfer ribonucleic acids. Hoppe Seylers Z. Physiol. Chem. *347*, 212–235.
33. Kowalski, S., Yamane, T., and Fresco, J.R. (1971). Nucleotide sequence of the "denaturable" leucine transfer RNA from yeast. Science *172*, 385–387.
34. Kruppa, J., and Zachau, H.G. (1972). Multiplicity of serine-specific transfer RNAs of brewer's and baker's yeast. Biochim. Biophys. Acta *277*, 499–512.
35. Johansson, M.J.O., and Byström, A.S. (2004). The Saccharomyces cerevisiae TAN1 gene is required for N4-acetylcytidine formation in tRNA. RNA *10*, 712–719.
36. Aravind, L., and Koonin, E.V. (2001). THUMP–a predicted RNA-binding domain shared by 4-thiouridine, pseudouridine synthases and RNA methylases. Trends Biochem. Sci. *26*, 215–217.
37. Kamalampeta, R., Keffer-Wilkes, L.C., and Kothe, U. (2013). tRNA binding, positioning, and modification by the pseudouridine synthase Pus10. J. Mol. Biol. *425*, 3863–3874.
38. Waterman, D.G., Ortiz-Lombardía, M., Fogg, M.J., Koonin, E.V., and Antson, A.A. (2006). Crystal structure of Bacillus anthracis ThiI, a tRNA-modifying enzyme containing the predicted RNA-binding THUMP domain. J. Mol. Biol. *356*, 97–110.
39. McCleverty, C.J., Hornsby, M., Spraggon, G., and Kreusch, A. (2007). Crystal structure of human Pus10, a novel pseudouridine synthase. J. Mol. Biol. *373*, 1243–1254.
40. Kotelawala, L., Grayhack, E.J., and Phizicky, E.M. (2008). Identification of yeast tRNA Um(44) 2¢-O-methyltransferase (Trm44) and demonstration of a Trm44 role in sustaining levels of specific tRNA(Ser) species. RNA *14*, 158–169.
41. Whipple, J.M., Lane, E.A., Chernyakov, I., D'Silva, S., and Phizicky, E.M. (2011). The yeast rapid tRNA decay pathway primarily monitors the structural integrity of the acceptor and T-stems of mature tRNA. Genes Dev. *25*, 1173–1184.
42. Dewe, J.M., Whipple, J.M., Chernyakov, I., Jaramillo, L.N., and Phizicky, E.M. (2012). The yeast rapid tRNA decay pathway competes with elongation factor 1A for substrate tRNAs and acts on tRNAs lacking one or more of several modifications. RNA *18*, 1886–1896.
43. Maddirevula, S., Alzahrani, F., Al-Owain, M., Al Muhaizea, M.A., Kayyali, H.R., AlHashem, A., Rahbeeni, Z., Al-Otaibi, M., Alzaidan, H.I., Balobaid, A., et al. (2019). Autozygome and high throughput confirmation of disease genes candidacy. Genet. Med. *21*, 736–742.
44. 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.
45. ACMG Board of Directors (2015). ACMG policy statement: updated recommendations regarding analysis and reporting of secondary findings in clinical genome-scale sequencing. Genet. Med. *17*, 68–69.
46. Arango, D., Sturgill, D., Alhusaini, N., Dillman, A.A., Sweet, T.J., Hanson, G., Hosogane, M., Sinclair, W.R., Nanan, K.K., Mandler, M.D., et al. (2018). Acetylation of Cytidine in mRNA Promotes Translation Efficiency. Cell *175*, 1872–1886.e24.
47. Rio, D.C. (2013). Expression and purification of active recombinant T7 RNA polymerase from E. coli. Cold Spring Harb. Protoc. *2013*. https://doi.org/10.1101/pdb.prot078527.
48. Su, D., Chan, C.T.Y., Gu, C., Lim, K.S., Chionh, Y.H., McBee, M.E., Russell, B.S., Babu, I.R., Begley, T.J., and Dedon, P.C. (2014). Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. Nat. Protoc. *9*, 828–841.
49. Lindeboom, R.G.H., Supek, F., and Lehner, B. (2016). The rules and impact of nonsense-mediated mRNA decay in human cancers. Nat. Genet. *48*, 1112–1118.
50. Ognjenović, J., and Simonović, M. (2018). Human aminoacyl-tRNA synthetases in diseases of the nervous system. RNA Biol. *15*, 623–634.
51. Kapur, M., Monaghan, C.E., and Ackerman, S.L. (2017). Regulation of mRNA Translation in Neurons-A Matter of Life and Death. Neuron *96*, 616–637.

52. Kumbhar, B.V., Kamble, A.D., and Sonawane, K.D. (2013). Conformational preferences of modified nucleoside N(4)-acetylcytidine, ac4C occur at "wobble" 34th position in the anticodon loop of tRNA. Cell Biochem. Biophys. *66*, 797–816.
53. Bartee, D., Nance, K.D., and Meier, J.L. (2021). Site-Specific Synthesis of N4-Acetylcytidine in RNA Reveals Physiological Duplex Stabilization. Preprint at bioRxiv. https://doi.org/10.1101/2021.11.12.468326.
54. Alexandrov, A., Chernyakov, I., Gu, W., Hiley, S.L., Hughes, T.R., Grayhack, E.J., and Phizicky, E.M. (2006). Rapid tRNA decay can result from lack of nonessential modifications. Mol. Cell *21*, 87–96.
55. Kwon, S.C., Yi, H., Eichelbaum, K., Föhr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. Nat. Struct. Mol. Biol. *20*, 1122–1130.
56. Lee, H.Y., Haurwitz, R.E., Apffel, A., Zhou, K., Smart, B., Wenger, C.D., Laderman, S., Bruhn, L., and Doudna, J.A. (2013). RNA-protein analysis using a conditional CRISPR nuclease. Proc. Natl. Acad. Sci. USA *110*, 5416–5421.
57. Jin, G., Xu, M., Zou, M., and Duan, S. (2020). The Processing, Gene Regulation, Biological Functions, and Clinical Relevance of N4-Acetylcytidine on RNA: A Systematic Review. Mol. Ther. Nucleic Acids *20*, 13–24.

The American Journal of Human Genetics, Volume *109*

Supplemental information

THUMPD1 bi-allelic variants cause loss of tRNA acetylation and a syndromic neurodevelopmental disorder

Martin Broly, Bogdan V. Polevoda, Kamel M. Awayda, Ning Tong, Jenna Lentini, Thomas Besnard, Wallid Deb, Declan O'Rourke, Julia Baptista, Sian Ellard, Mohammed Almannai, Mais Hashem, Ferdous Abdulwahab, Hanan Shamseldin, Saeed Al-Tala, Fowzan S. Alkuraya, Alberta Leon, Rosa L.E. van Loon, Alessandra Ferlini, Mariabeatrice Sanchini, Stefania Bigoni, Andrea Ciorba, Hans van Bokhoven, Zafar Iqbal, Almundher Al-Maawali, Fathiya Al-Murshedi, Anuradha Ganesh, Watfa Al-Mamari, Sze Chern Lim, Lynn S. Pais, Natasha Brown, Saima Riazuddin, Stéphane Bézieau, Dragony Fu, Bertrand Isidor, Benjamin Cogné, and Mitchell R. O'Connell

Figure S1

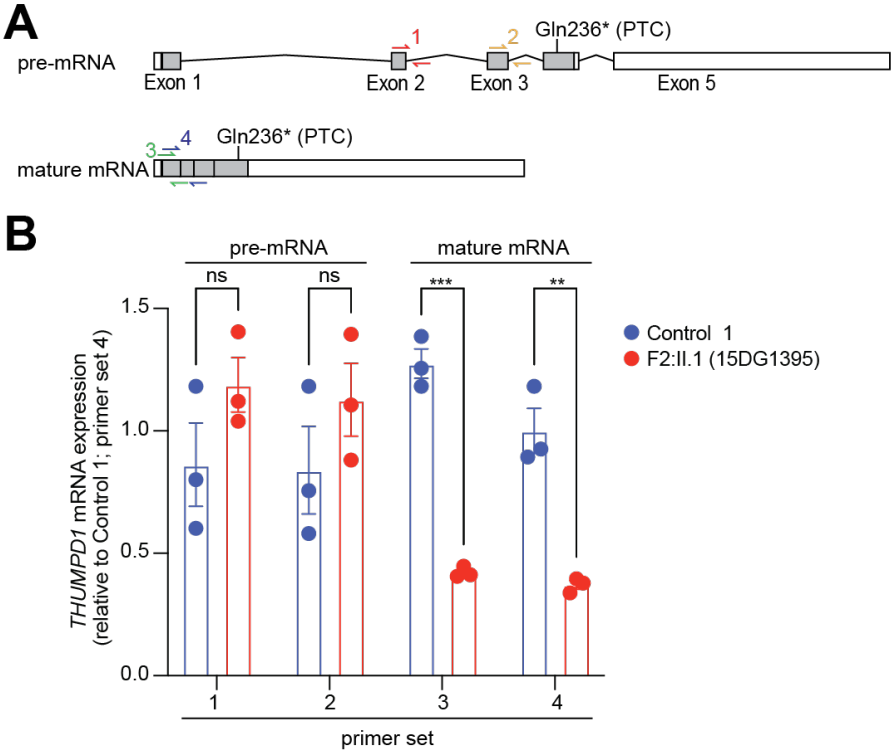

Figure S1. THUMPD1 pre-mRNA and mature mRNA expression control and individual F2:II.1 lymphoblasts. (A) A schematic of *THUMPD1* pre-mRNA and mature mRNA (NM_001304550.1) showing the location of primer pairs 1-4 on the transcript. (B) RT-qPCR of the corresponding RNA samples. Samples were analyzed in triplicate, error bars represent -/+ SD with the data normalized to Control 1, primer pair 4 (mature mRNA) *THUMPD1* expression levels. **: p(adjusted) < 0.005, ***: p(adjusted) < 0.0005 as determined by a two-way ANOVA with Benjamini Hochberg correction (FDR < 5%).

Figure S2

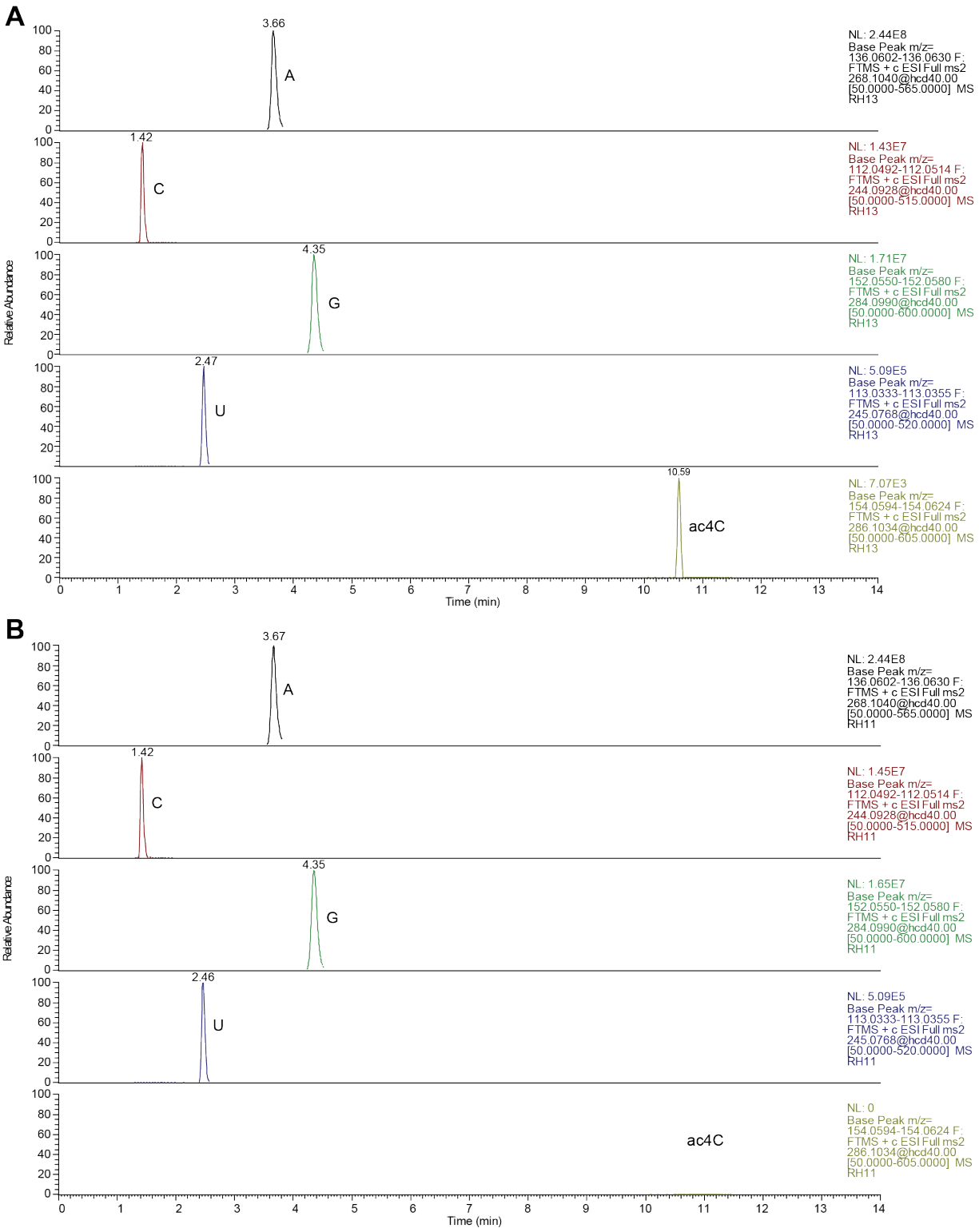
A
3.66
A
NL: 2.44E8
Base Peak m/z=
136.0602-136.0630 F:
FTMS + c ESI Full ms2
268.1040@hcd40.00
[50.0000-565.0000] MS
RH13
1.42
C
NL: 1.43E7
Base Peak m/z=
112.0492-112.0514 F:
FTMS + c ESI Full ms2
244.0928@hcd40.00
[50.0000-515.0000] MS
RH13
4.35
G
NL: 1.71E7
Base Peak m/z=
152.0550-152.0580 F:
FTMS + c ESI Full ms2
284.0990@hcd40.00
[50.0000-600.0000] MS
RH13
2.47
U
NL: 5.09E5
Base Peak m/z=
113.0333-113.0355 F:
FTMS + c ESI Full ms2
245.0768@hcd40.00
[50.0000-520.0000] MS
RH13
10.59
ac4C
NL: 7.07E3
Base Peak m/z=
154.0594-154.0624 F:
FTMS + c ESI Full ms2
286.1034@hcd40.00
[50.0000-605.0000] MS
RH13
Relative Abundance
Time (min)
B
3.67
A
NL: 2.44E8
Base Peak m/z=
136.0602-136.0630 F:
FTMS + c ESI Full ms2
268.1040@hcd40.00
[50.0000-565.0000] MS
RH11
1.42
C
NL: 1.45E7
Base Peak m/z=
112.0492-112.0514 F:
FTMS + c ESI Full ms2
244.0928@hcd40.00
[50.0000-515.0000] MS
RH11
4.35
G
NL: 1.65E7
Base Peak m/z=
152.0550-152.0580 F:
FTMS + c ESI Full ms2
284.0990@hcd40.00
[50.0000-600.0000] MS
RH11
2.46
U
NL: 5.09E5
Base Peak m/z=
113.0333-113.0355 F:
FTMS + c ESI Full ms2
245.0768@hcd40.00
[50.0000-520.0000] MS
RH11
ac4C
NL: 0
Base Peak m/z=
154.0594-154.0624 F:
FTMS + c ESI Full ms2
286.1034@hcd40.00
[50.0000-605.0000] MS
RH11
Relative Abundance
Time (min)

Figure S2. LC-MS nucleoside modification analysis of the small RNA samples prepared from control and individual F2:II.1 lymphoblasts. Representative LC-separation chromatograms for small RNA samples from control (A) and (B) individual F2:II.1 lymphoblasts showing the retention times (at the top of the peaks) and relative abundance (with absolute scale values on the right of the field) for all four normal nucleosides, A, C, G and U, and ac4C modified nucleoside.

Figure S3

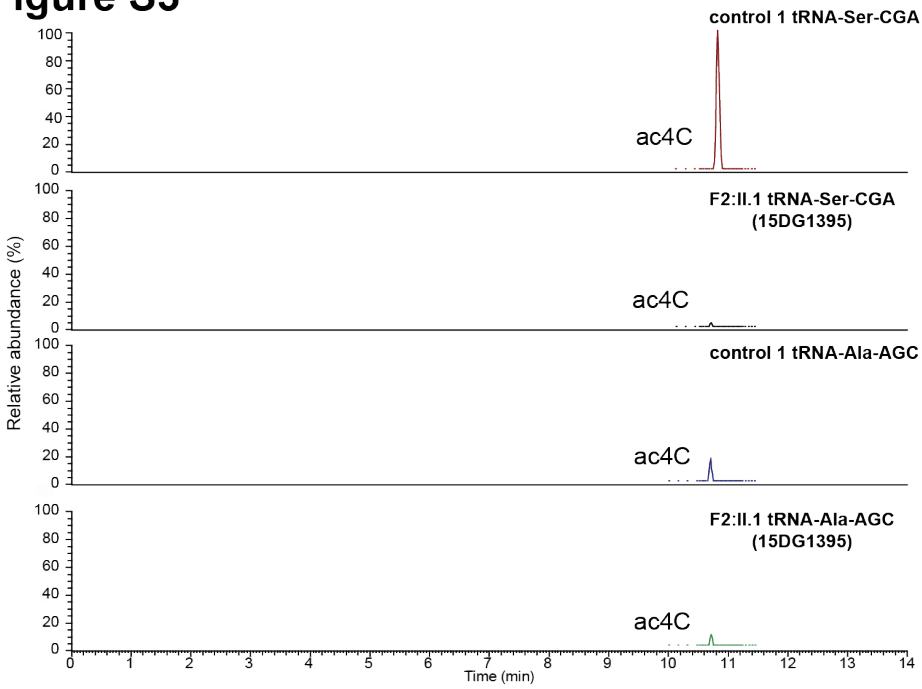

Figure S3. LC-MS nucleoside modification analysis of the purified tRNA samples prepared from control and individual F2:II.1 lymphoblasts. Representative LC-separation chromatograms for purified tRNA-Ser-CGA and tRNA-Ala-AGC samples from control and individual F2:II.1 lymphoblasts, as indicated. The ac4C peaks are shown in green and unmodified cytidine peaks are in blue (as controls).

Figure S4

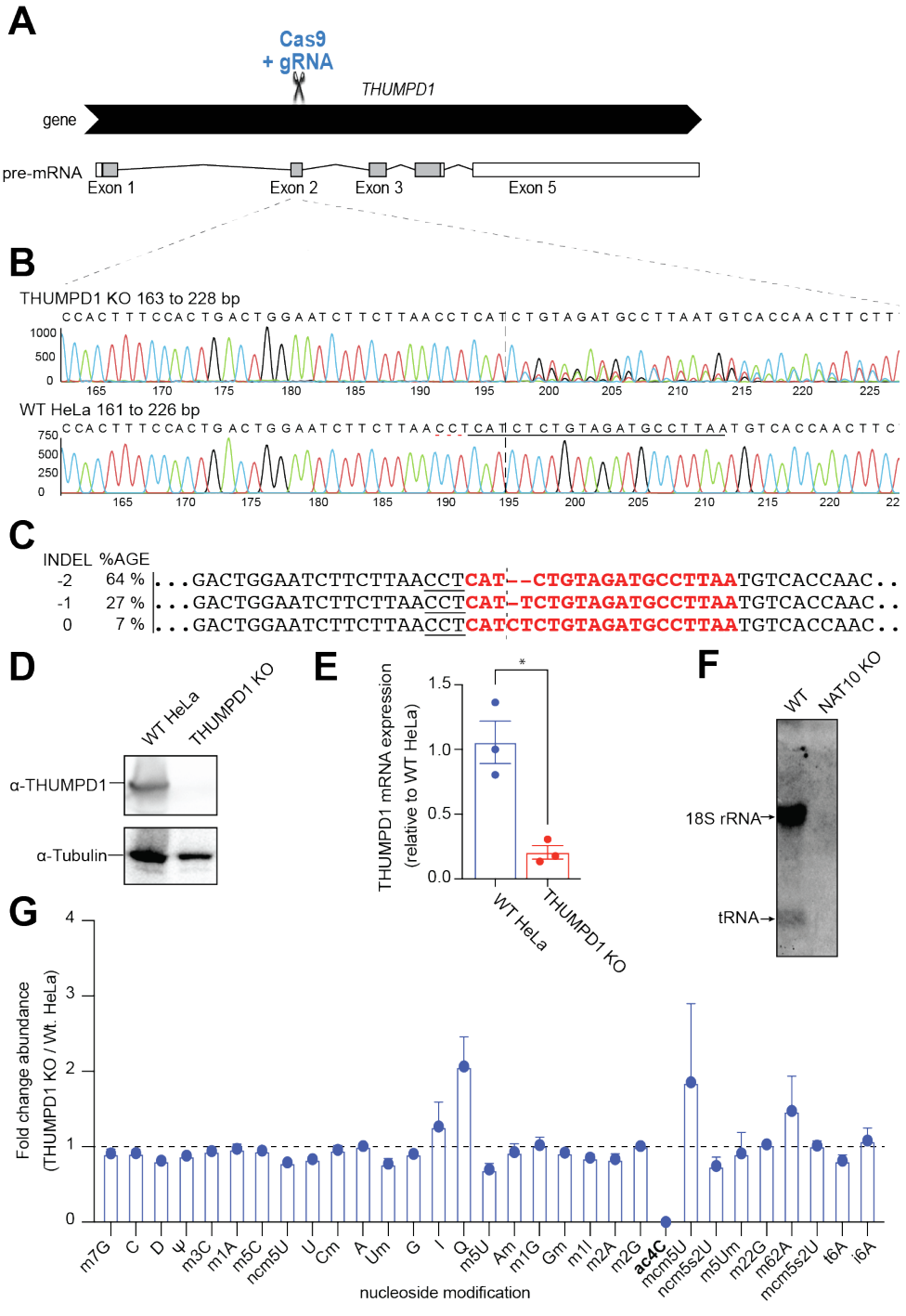

Figure S4. THUMPD1 KO in HeLa cells results in the loss of ac4C modification in small RNA samples similar to small RNA in individual F2:II.1 (15DG1395 lymphoblasts)

(A) Schematic representation of the *THUMPD1* KO in a HeLa cell line. exon 2 to produce an aberrant PTC-containing transcript (see Table S1 for gRNA sequence). (B) Sanger DNA sequencing data of the *THUMPD1* exon 2 genomic region from selected *THUMPD1* KO clone (top) and WT HEK293T cells (bottom). Underlined DNA sequence in WT HEK293T sample is complementary to the designed sgRNA. Dotted vertical line in both sequences indicates a Cas9 target cleavage site. (C) ICE analysis (Synthego) results of the *THUMPD1* exon 2 genomic DNA region from a selected *THUMPD1* KO clone. Target nucleotide sequence complementary to sgRNA is in red font and dotted vertical line indicates a Cas9 target cleavage site. The indel type and percentage of detected sequence are shown on the left. (D) Western blot analysis of the protein cell extracts prepared from HeLa control and *THUMPD1* KO cell lines probed with anti-THUMPD1 antibody and anti-tubulin antibody (as loading control). (E) *THUMPD1* mRNA expression in HeLa control and *THUMPD1* KO cells determined by RT-qPCR of the corresponding RNA samples. Samples were analyzed in triplicate, error bars represent -/+ SD. *: $p < 0.05$, as determined by a Welch's t test. (F) Northern immunoblot analysis of the total RNA samples prepared from HeLa control (WT) and *NAT10* KO cells. Positions of 18S rRNA and tRNA on the gel (membrane) are shown by the arrows on the left. (G) LC-MS nucleoside modification analysis of the small RNA samples prepared from HeLa control and *THUMPD1* KO cells. Alterations (fold change) for each nucleoside modification was calculated as a ratio of the corresponding value in *THUMPD1* KO RNA to HeLa control RNA normalized based on the abundance of all standard non-modified nucleosides (A,C,G and U). Samples were ran in duplicate, error bars represent -/+ SD.

Figure S5

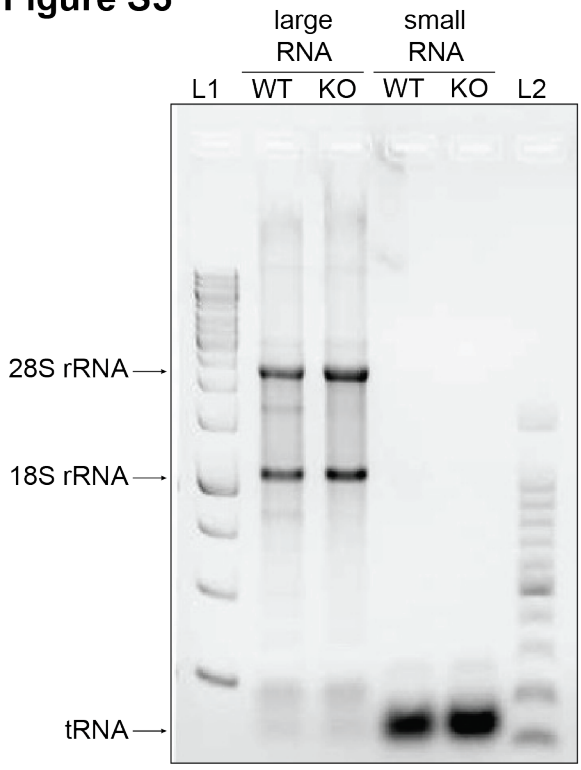

Figure S5. Typical large and small RNA samples

A representative 1.5% agarose gel image with large and small RNA samples prepared from HeLa WT and *THUMPD1* KO cells. L1, 1 kb DNA ladder (Promega); L2, 100 nt DNA ladder (Promega). Position 28S rRNA, 18S rRNA and tRNA on the gel are indicated by arrows on the left.

Figure S6

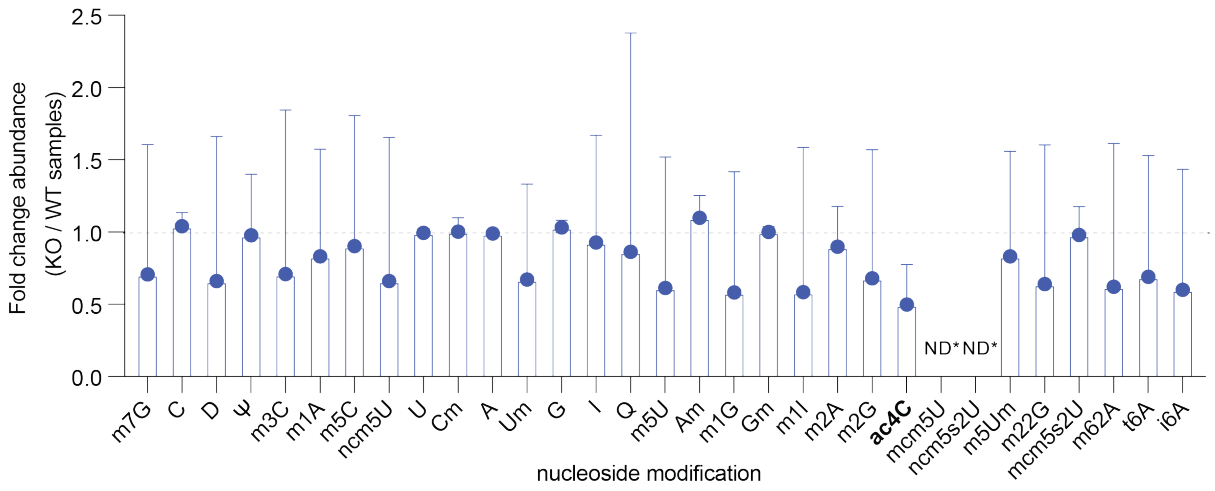

Figure S6. LC-MS nucleoside modification analysis of the large RNA samples prepared from HEK293T control and THUMPD1 KO cells.

Alterations (fold change) for each nucleoside modification was calculated as a ratio of the corresponding value in *THUMPD1* KO RNA to HEK293T control RNA normalized based on the abundance of all four normal nucleosides (A,C,G and U). Samples were ran in duplicate, error bars represent -/+ SD.

Figure S7

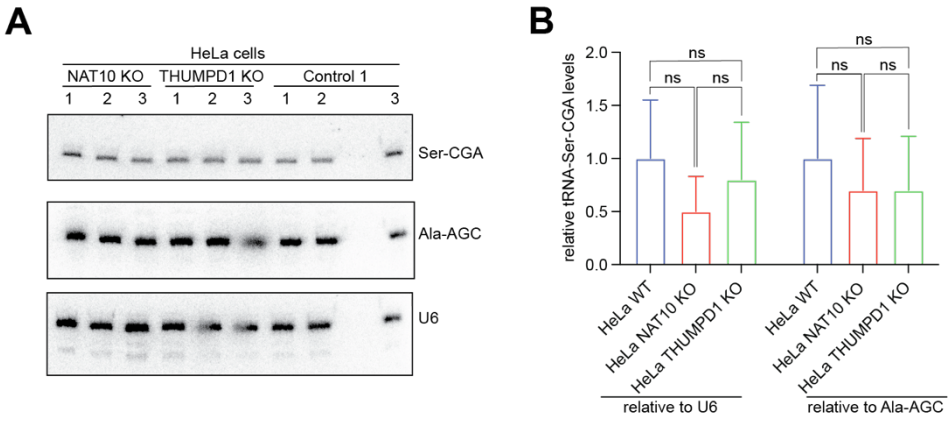

Figure S7. Northern blot analysis of the small RNA samples prepared from HeLa control (WT), THUMPD1 KO and NAT10 KO cells.

(A) A representative Northern blot for small RNA samples from HeLa WT (control 1), *THUMPD1* KO and *NAT10* KO cells, all loaded in triplicate and probed with tRNA-Ser-CGA1-1 (top panel), tRNA-Ala-AGC9-1 (middle panel) or U6 RNA (bottom panel) specific probes. (B) Quantitation of the Northern blot data showing relative tRNA-Ser-CGA levels in HeLa control, *THUMPD1* KO and *NAT10* KO cells as normalized to U6 RNA or tRNA-Ala-AGC levels in the corresponding samples. Error bars represent -/+ SD. nd: not significant (adjusted p valve > 0.05) as determined by a Brown-Forsythe and Welch one-way ANOVA with Benjamini Hochberg correction (FDR < 5%).

Table S1. Oligonucleotide sequences used in this study

RNA sequences

THUMPD1 KO sgRNA guide sequence

5'- UUAAGGCAUCUACAGAGAUG-3'

DNA oligonucleotides

THUMPD1 CRISPR KO screening oligonucleotides:

THUMPD1 KO F	5'-CCCAAAGAGCAACAACTGGC-3'
THUMPD1 KO R	5'-GCATGGTAGCTCACACCTGT-3'

THUMPD1 LIC cloning into pET-6×His-TEV:

THUMPD1 pET F	5'-TACTCCCAATCCAATGCA GCGGCCCCTGCCCAGCAG-3'
THUMPD1 pET R	5'-TTATCCACTTCCAATGTATTA CTATGAGAAGTCATTTT CA-3'

Real-time PCR oligos:

THUMPD1_primer_set_1 (pre-mRNA) F	5'-GTCATGCCCAAGAAGCCATTTT-3'
THUMPD1_primer_set_1 (pre-mRNA) R	5'-AGTCAGTGGAAAGTGGAGCAA-3'
THUMPD1_primer_set_2 (pre-mRNA) F	5'-CCATTGAATGCCTATGGAAAATCAC-3'
THUMPD1_primer_set_2 (pre-mRNA) R	5'-AAGCTCCAAACAAAGGGACAT-3'
THUMPD1_primer_set_3 (mature mRNA) F	5'-CCTCAACGAATACGGCGAC-3'
THUMPD1_ primer_set_3 (mature mRNA) R	5'-TCAGGCTCTATCCCAAGTGTC-3'
THUMPD1_primer_set_4 (mature mRNA) F	5'-CCCTGCCCAGCAGACTAC-3'
THUMPD1_primer_set_4 (mature mRNA) R	5'- CTGATCCTTGTCTGTAAACTTTTCT-3'
GAPDH F	5'-GAGTCAACGGATTTGGTCG-3'
GAPDH R	5'-GACAAGCTTCCCGTTCTCAG-3'

Mutagenesis:

THUMPD1 P164S F	5'-TTTGCGAATG TTATCCATCT CAGGC-3'
THUMPD1 P164S R	5'-GCCTGAGATG GATAACATTC GCAAA-3'

Northern probes:

tRNASer-CGA	5'-CGCTGTGAGCAGGATTCGA-3'
tRNAAla-AGC	5'-TCCCGCTACCTCTCGCATG-3'
U6 RNA	5'-CGTTCCAATTTTAGTATATGTGCTGCCGAAGCGA-3'

Biotinylated probes for individual tRNA isolation:

tRNASer-CGA-BTN	5'-Biotin-TTCGAACCTG CGCGGGGAGA CCCCATT-3'
tRNAAla-ACG-BTN	5'-Biotin-AATGCGGGCG TCGATCCCGC TACCTCT-3'

tRNA DNA templates:

tRNA-Ser-CGA-1-1	5'-GCTGTGATGG CCGAGTGGTT AAGGCGTTGG ACTCGAAATC CAATGGGGT TCCCCGCGCA GGTTCGAATC CTGCTCACAG CG-3'
tRNA-Ala-AGC-9-1	5'-GGGGAATTAG CTCAGGCGGT AGAGCGCTCG CTTAGCATGC GAGAGGTAGC GGGATCGACG CCCGCATTCT CCA-3'

Table S2: List of unmodified and modified nucleosides detected in this study

Name	Abbreviated Name	Precursor m/z	Fragment m/z	Retention Time (min)
Cytidine	C	244.0928	112.0503	1.42
Pseudouridine	Ψ	245.0768	125.0344	1.58
Uridine	U	245.0768	113.0344	2.48
Dihydrouridine	D	247.7859	115.0500	1.66
2′-O-methylcytidine	Cm	258.1085	112.0504	3.08
3-methylcytidine	m3C	258.1085	126.0660	1.68
5-methylcytidine	m5C	258.1085	126.0660	1.91
2′-O-methyluridine	Um	259.0925	169.0605	3.49
5-methyluridine	m5U	259.0925	127.0500	5.5
Adenosine	A	268.104	136.0616	3.61
Inosine	I	269.0881	137.0460	4.72
5,2′-O-dimethyluridine	m5Um	273.1081	127.0500	11.53
1-methyladenosine	m1A	282.1197	150.0772	1.9
2-methyladenosine	m2A	282.1197	150.0772	10.5
2′-O-methyladenosine	Am	282.1197	136.0616	8.45
1-methylinosine	m1I	283.1037	151.0614	10.42
Guanosine	G	284.099	152.0565	4.57
N4-acetylcytidine	ac4C	286.1034	154.0609	10.76
N6,N6-dimethyladenosine	m62A	296.1354	164.0929	12.27
1-methylguanosine	m1G	298.1146	166.0722	8.64
N2-methylguanosine	m2G	298.1146	166.0722	10.48
2′-O-methylguanosine	Gm	298.1146	152.0564	9.68
7-methylguanosine	m7G	300.1303	151.0611	1.26
5-carbamoylmethyluridine	ncm5U	302.0983	153.0292	2.21
5-carboxymethyluridine	cm5U	303.0823	152.0564	11.05
N2,N2-dimethylguanosine	m22G	312.1303	180.0876	11.61
5-methoxycarbonylmethyluridine	mcm5U	317.098	153.0294	10.84
5-methoxycarbonylmethyl-2′-O-methyluridine	mcm5Um	331.1136	153.0294	12.38
5-methoxycarbonylmethyl-2-thiouridine	mcm5s2U	333.0751	169.0064	12.33
N6-isopentenyladenosine	i6A	336.1667	204.1241	13.18
Queuosine	Q	410.167	163.0613	5.2
N6-threonylcarbamoyladenosine	t6A	413.1416	136.0616	12.55

Table S3: Complete phenotype/clinical information for NDD-affected individuals with THUMPD1 bi-allelic pathogenic variants. Please see attached excel spreadsheet for this table (Table_S3.xlsx).

Supplementary text

Section 1: Exome sequencing methods, analysis platforms and pipelines used

Nantes: We performed exome sequencing on the proband with the Twist Human core + refseq exome (Twist Bioscience) according to manufacturer's instructions and we generated 75-bp paired-end reads on an Illumina NextSeq500. Fastq files were aligned to human genome hg19 with bwa mem (v0.7.3). We then called SNVs and INDELs following GATKs best practices (v3.4). We achieved an average mean target coverage of 104X. Variants were annotated using ANNOVAR and filtered with in-house scripts to keep variants with at least 9 reads and with a variant read frequency over 20 percent impacting exonic sequences or splice sites (+/- 10bp from the junction) and with an allele frequency <0.5% in 1000 genomes, genome aggregation database (gnomAD, 123,136 exomes and 15,496 whole genome sequences; accessed on 11/10/2018)[58] and in a local database. The possible functional impact of amino-acid changes was predicted by SIFT (Sorting Intolerant from Tolerant)[59], PolyPhen-2 hvar[60] and CADD score (Combined Annotation Dependent Depletion).[61] The Alamut software (Interactive biosoftware) was used to study retained variant sites. Variants were confirmed by Sanger sequencing in the proband and its parents (details available upon request).

Oman: To further explore the genetic cause of the phenotype, the family was recruited after obtaining informed consent under an IRB-approved research protocol (SQU-MREC#1362). We performed whole-exome sequencing (WES) analysis on DNA isolated from blood from the three affected siblings. In brief, genomic DNA of the individual was isolated from peripheral blood using a DNeasy Blood and Tissue Kit (Qiagen, Courtaboeuf, France). DNA was barcoded and enriched for the coding exons of targeted genes using hybrid capture technology (Agilent SureSelect-V6-60 MB). Prepared DNA libraries were then sequenced using a Next Generation Sequencing (NGS) technology [Hiseq4000, 150 bp paired-end, at 200X coverage]. Variant filtration was conducted to only keep novel or rare variants (≤ 1%). Publically available variant databases (1000 Genomes, Exome Variant Server, and gnomAD) were used to determine the frequency. Also, an in-house database of 1564 exome was used to filter out common variants specific to our population. Only coding/splicing variants considered. The phenotype and mode of inheritance (autosomal recessive) were considered. Sanger sequencing was used to confirm the segregation.

58. Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alföldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., et al.; Genome Aggregation Database Consortium (2020). The mutational constraint spectrum quantified from variation in 141,456 humans. Nature *581*, 434–443.

59. Ng, P.C., and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. *31*, 3812–3814.

60. Adzhubei, I., Jordan, D.M., and Sunyaev, S.R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. Curr. Protoc. Hum. Genet. Chapter 7, 20.

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