De Novo Variants in *CNOT1*, a Central Component of the CCR4-NOT Complex Involved in Gene Expression and RNA and Protein Stability, Cause Neurodevelopmental Delay

Lisenka E.L.M. Vissers,[1,43,*] Sreehari Kalvakuri,[2,43] Elke de Boer,[1] Sinje Geuer,[1] Machteld Oud,[1] Inge van Outersterp,[1] Michael Kwint,[1] Melde Witmond,[1] Simone Kersten,[1,3] Daniel L. Polla,[1,4] Dilys Weijers,[1] Amber Begtrup,[5] Kirsty McWalter,[5] Anna Ruiz,[6] Elisabeth Gabau,[6] Jenny E.V. Morton,[7] Christopher Griffith,[8] Karin Weiss,[9] Candace Gamble,[11] James Bartley,[12] Hilary J. Vernon,[13] Kendra Brunet,[14] Claudia Ruivenkamp,[15] Sarina G. Kant,[15] Paul Kruszka,[10] Austin Larson,[16] Alexandra Afenjar,[17] Thierry Billette de Villemeur,[18] Kimberly Nugent,[18] the DDD Study,[42] F. Lucy Raymond,[19] Hanka Venselaar,[20] Florence Demurger,[21] Claudia Soler-Alfonso,[22] Dong Li,[23] Elizabeth Bhoj,[23] Ian Hayes,[24] Nina Powell Hamilton,[25] Ayesha Ahmad,[25] Rachel Fisher,[25] Myrthe van den Born,[26] Marjolaine Willems,[27] Arthur Sorlin,[28,29] Julian Delanne,[28,29] Sebastien Moutton,[28,30] Philippe Christophe,[28,31] Frederic Tran Mau-Them,[28,31] Antonio Vitobello,[28,31]

(Author list continued on next page)

CNOT1 is a member of the CCR4-NOT complex, which is a master regulator, orchestrating gene expression, RNA deadenylation, and protein ubiquitination. We report on 39 individuals with heterozygous *de novo CNOT1* variants, including missense, splice site, and nonsense variants, who present with a clinical spectrum of intellectual disability, motor delay, speech delay, seizures, hypotonia, and behavioral problems. To link CNOT1 dysfunction to the neurodevelopmental phenotype observed, we generated variant-specific *Drosophila* models, which showed learning and memory defects upon *CNOT1* knockdown. Introduction of human wild-type *CNOT1* was able to rescue this phenotype, whereas mutants could not or only partially, supporting our hypothesis that *CNOT1* impairment results in neurodevelopmental delay. Furthermore, the genetic interaction with autism-spectrum genes, such as *ASH1L*, *DYRK1A*, *MED13*, and *SHANK3*, was impaired in our *Drosophila* models. Molecular characterization of *CNOT1* variants revealed normal *CNOT1* expression levels, with both mutant and wild-type alleles expressed at similar levels. Analysis of protein-protein interactions with other members indicated that the CCR4-NOT complex remained intact. An integrated omics approach of patient-derived genomics and transcriptomics data suggested only minimal effects on endonucleolytic nonsense-mediated mRNA decay components, suggesting that *de novo CNOT1* variants are likely haploinsufficient hypomorph or neomorph, rather than dominant negative. In summary, we provide strong evidence that *de novo CNOT1* variants cause neurodevelopmental delay with a wide range of additional co-morbidities. Whereas the underlying pathophysiological mechanism warrants further analysis, our data demonstrate an essential and central role of the CCR4-NOT complex in human brain development.

Master regulators controlling development include, but are not limited to, paired box (PAX) proteins,[1] SRY-related HMG-box (SOX) proteins,[2] and the relatively unknown CCR4-NOT protein complex.[3] Although the full spectrum of functional diversity for the CCR4-NOT complex has not yet been established, it has already become apparent that it is active on all levels of gene expression, from accessibility of the DNA to translation and degradation of mRNA.[4]

[1]Department of Human Genetics, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, the Netherlands; [2]Development, Aging and Regeneration Program, Sanford Burnham Prebys Medical Discovery Institute, 10901 North Torrey Pines Rd, La Jolla, CA 92037, USA; [3]Department of Internal Medicine, Radboud Institute for Molecular Life Sciences, Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen the Netherlands; [4]CAPES Foundation, Ministry of Education of Brazil, 70040-031 Brasília, Brazil; [5]GeneDx, 207 Perry Pkwy, Gaithersburg, MD 20877, USA; [6]Paediatric Unit, Parc Taulí Hospital Universitari, Institut d'Investigació i Innovació Parc Taulí I3PT, Universitat Autònoma de Barcelona, 08208 Sabadell, Barcelona, Spain; [7]West Midlands Regional Genetics Laboratory, Birmingham Women's Hospital, Birmingham Women's and Children's NHS Foundation Trust, Edgbaston B15 2TG, UK; [8]Department of Pediatrics, University of South Florida, Tampa, FL 33606, USA; [9]The Genetics Institute, Rambam Health Care Center, Haifa, Israel; The Ruth and Bruce Rappaport Faculty of Medicine, Technion – Israel Institute of Technology, Haifa, Israel; [10]National Human Genome Research Institute, National Institutes of Health, 10 Center Dr, Bethesda, MD 20814, USA; [11]Cook Children's, 801 7th Ave, Fort Worth, TX 76104, USA; [12]Pediatric Specialty Clinics, Loma Linda University, 11234 Anderson St., Loma Linda, CA 92354, USA; [13]Department of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; [14]Porcupine Health Unit, 169 Pine St S, Timmins, ON P4N 2K3, Canada; [15]Department of Clinical Genetics, Leiden University Medical Centre, PO Box 9600, 2300 RC Leiden, the Netherlands; [16]Department of Pediatrics, Section of Genetics, University of Colorado School of Medicine, Aurora, CO 80045, USA; [17]CRMR malformations et maladies congénitales du cervelet et déficiences intellectuelles de causes rares, Département de génétique, Sorbonne Université, AP-HP, Hôpital Trousseau, 75012 Paris, France; [18]Children's Hospital of San Antonio, 333 N Santa Rosa St, San Antonio, TX 78207, USA; [19]Department of Medical Genetics, University of Cambridge, CB2 0XY Cambridge, UK; [20]Center for Molecular and Biomolecular Informatics, Radboud Institute for Molecular Life Sciences (RIMLS), PO Box 9101, 6500 HB Nijmegen, the Netherlands; [21]Centre Hospitalier Bretagne Atlantique, 20 Boulevard Général Maurice Guillaudot, BP 70555, 56017 Vannes Cedex, France; [22]Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA; [23]Center for Applied Genomics, The Children's Hospital of Philadelphia, 3615 Civic Center Blvd, Philadelphia, PA 19104, USA; [24]Genetic Health Service New Zealand, 2 Park Road, Grafton, Auckland

(Affiliations continued on next page)

Himanshu Goel,[32] Lauren Massingham,[33] Chanika Phornphutkul,[33] Jennifer Schwab,[33] Boris Keren,[34] Perrine Charles,[34] Maaike Vreeburg,[35] Lenika De Simone,[36] George Hoganson,[36] Maria Iascone,[37] Donatella Milani,[38] Lucie Evenepoel,[39] Nicole Revencu,[39] D. Isum Ward,[40] Kaitlyn Burns,[40] Ian Krantz,[41] Sarah E. Raible,[41] Jill R. Murrell,[41] Kathleen Wood,[41] Megan T. Cho,[5] Hans van Bokhoven,[1] Maximilian Muenke,[9] Tjitske Kleefstra,[1] Rolf Bodmer,[2,44,*] and Arjan P.M. de Brouwer[1,44]

The human CCR4-NOT complex contains up to 11 different subunits,[3] with each of the subunits having specified functions. For instance, the catalytic activity of CNOT6, CNOT6L, CNOT7, and CNOT8 plays an important role in the deadenylation step leading to mRNA degradation, and the E3 ligase activity[5] of CNOT4 is involved in protein substrate recognition and ubiquitination. The residual subunits seem to have a scaffolding function,[3,6,7] with CNOT1 being the central scaffolding protein (in)directly binding to all CCR4-NOT partners.[8] In addition to its scaffolding function, CNOT1 has been considered as a translational regulator through the binding of nuclear receptors and as a regulator of deadenylase activity.[9,10] For the latter, CNOT1 also exhibits the capacity to bind proteins that are not part of the CCR4-NOT complex, but are known to either be involved in general or tissue-specific mRNA degradation pathways.[11–15] Interrogation of gnomAD, as well as visualization of the CNOT1 tolerance landscape using MetaDome[16] (Figure S1), has shown that CNOT1 is significantly depleted of loss-of-function (LoF; pLI = 1.00) and missense variation (Z-score = 7.25), suggesting that such variants may lead to genetic disease. Indeed, two recent studies together reported five individuals with a recurrent *de novo* missense variant in *CNOT1* (GenBank: NM_016284.4; c.1603C>T [p.Arg535Cys]) causing a novel syndrome characterized by pancreatic agenesis and holoprosencephaly (MIM: 618500).[17,18] *De novo* variants elsewhere in the gene, however, have not been systematically been reported.

We collected 39 individuals (19 females, 20 males), from 37 nuclear families, with a (likely) pathogenic variant *CNOT1* through international collaborations, facilitated by MatchMaker Exchange,[19] for further molecular and clinical studies to establish the role of CNOT1 in neurodevelopmental disorders (Figure 1, Table S1). For 31 individuals, it was established that the variant had occurred *de novo*. For three individuals, the variant was inherited from a mildly affected (n = 2) or mosaic (n = 1) mother. Additional segregation to establish the inheritance of the *CNOT1* variant in the two transmitting mothers by study of the maternal grandparents was not performed. For three individuals, inheritance could not be established because of absence of (one of the) parental samples (Table S1). The ages of the index subjects ranged from newborn to 22 years (Table S2). In essence, individuals with a (*de novo*) *CNOT1* variant show a broad phenotypic spectrum (Figures 1 and S2 and Table S2). The most consistent features observed included intellectual disability (ID) (72%) of varying degree, development delay (DD) (92%), speech delay (83%), motor delay (83%), and hypotonia (74%) (Tables 1 and S2). Although facial abnormalities were very common (92%, Figure 1), these did not yield a typical gestalt. Similarly, abnormal growth (74%), behavioral problems (65%), abnormalities of the brain (65%) and skeletal, and muscle and soft tissue abnormalities (67%) were frequently observed, albeit with different features in each individual. From the clinical characteristics of the cohort, we concluded that there is no recognizable "CNOT1 phenotype." This fits previous observations that genes involved in newly discovered neurodevelopmental disorders are often clinically not recognizable.[20]

Given the relative large number of individuals (n = 39) collected, we next reasoned that there may be a more subtle genotype-phenotype correlation, which could be based on the type of variant (truncating versus missense) or the location of the missense variants (e.g., different CNOT1 protein domains; Figure 1). Whereas previously a striking resemblance was observed for individuals with the same *de novo* p.Arg535Cys variant within the HEAT domain, consisting of holoprosencephaly and absence or insufficient function of the pancreas,[17,18] no such other features were observed

1023, New Zealand; [25]Department of Pediatric Genetics, Metabolism and Genomic Medicine, University of Michigan, 1522 Simpson Road East, Ann Arbor, MI 48109-5718, USA; [26]Department for Clinical Genetics, Erasmus MC, Postbus 2040, 3000 CA Rotterdam, the Netherlands; [27]Centre Hospitalier Universitaire de Montpellier, 191 av. du Doyen Giraud, 34295 Montpellier Cedex 5, France; [28]INSERM U1231, LNC UMR1231 GAD, Burgundy University, 21000 Dijon, France; [29]Reference Center for Developmental Anomalies, Department of Medical Genetics, Dijon University Hospital, 2 Boulevard du Maréchal de Lattre de Tassigny, 21000 Dijon, France; [30]CPDPN, Pôle mère enfant, Maison de Santé Protestante Bordeaux Bagatelle, 33401 Talence, France; [31]Laboratoire de génétique, Innovation en diagnostic génomique des maladies rares UF6254, Plateau Technique de Biologie, CHU Dijon, 14 rue Paul Gaffarel, BP 77908, 21079 Dijon, France; [32]Hunter Genetics, Waratah, NSW 2298, Australia; University of Newcastle, Callaghan, NSW 2308, Australia; [33]Division of Human Genetics, Department of Pediatrics, Warren Alpert Medical School of Brown University, Hasbro Children's Hospital/Rhode Island Hospital, Providence, RI 02905, USA; [34]Genetic Department, APHP, Sorbonne Université, Pitié-Salpêtrière Hospital, 47-83 Boulevard de l'Hôpital, 75013 Paris, France; [35]Department of Clinical Genetics, Maastricht UMC+, Postbus 5800, 6202 AZ Maastricht, the Netherlands; [36]UIC Pediatric Genetics, 840 South Wood Street, Chicago, IL 60612, USA; [37]Laboratorio di genetica Medica, ASST Papa Giovanni XXIII, Bergamo 24127, Italy; [38]Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Pediatric Highly Intensive Care Unit, Milan, 20122, Italy; [39]Centre de Génétique Humaine, Cliniques Universitaires Saint-Luc, Université Catholique de Louvain, Avenue Hippocrate 10-1200, Brussels, Belgium; [40]Sanford Health 1600 W. 22th St, Sioux Falls, SD 57105, USA; [41]The Children's Hospital of Philadelphia, Philadelphia, 3401 Civic Center Blvd, Philadelphia, PA 19104, USA; [42]Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK
[43]These authors contributed equally to this work
[44]These authors contributed equally to this work
*Correspondence: lisenka.vissers@radboudumc.nl (L.E.L.M.V.), rolf@sbpdiscovery.org (R.B.)
https://doi.org/10.1016/j.ajhg.2020.05.017.

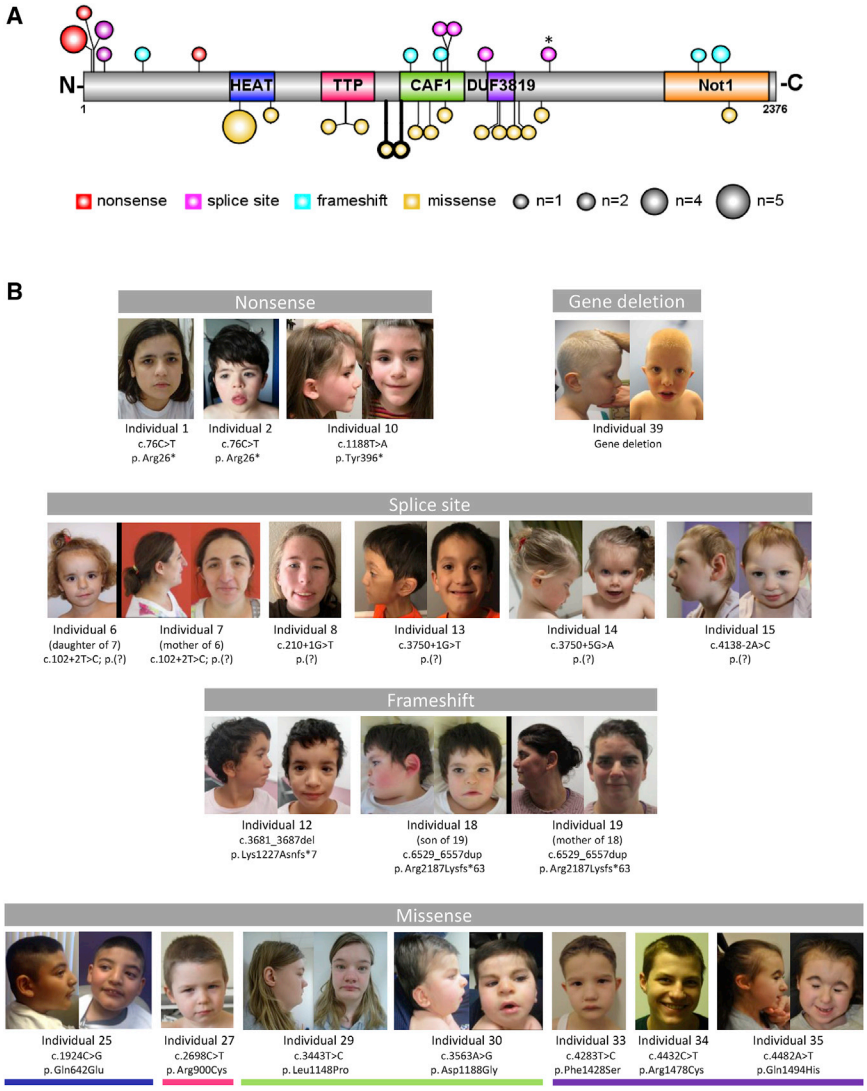


Figure 1. Schematic Representation of CNOT1 and Facial Photos of Individuals Carrying a (Likely) Pathogenic *CNOT1* Variant

(A) Graphical representation of the linear protein structure of CNOT1 with the functional domains indicated by colored boxes (blue, HEAT; pink, TTP; green, CAF; purple, DUF3819; orange, Not1). Variants observed in this study are indicated by circles, in which the size of the circle corresponds to the number of recurrences and the color of the circle to the variant type. Note that the variant in individual 16, affecting the last base of exon 34, creates a splice site (indicated by an asterisk). The variants in individual 28, who was shown to have two *de novo* missense variants, are outlined in bold.

(B) Facial photos of individuals with a (likely) pathogenic *CNOT1* variant, categorized by the type of variant observed. The colored lines underneath the individuals with missense variants correspond to the functional domains indicated in (A). Although shared facial features can be observed between individuals in the cohort, e.g., straight eyebrows (individuals 2, 5, 6, 12, 18, 19, 27, 33), hypertelorism (individuals 18, 25, 29, 33, 39), epicanthus inversus (individuals 8, 18, 30, 33), low nasal bridge (individuals 2, 14, 18, 29, 39), large ears (individuals 15, 18, 25, 39), thickened helix (individuals 10, 14, 18, 25, 29, 35), and protruding helix (individuals 5, 7, 10, 14, 33, 39), a syndromic facial phenotype is lacking. Also, no evident facial resemblance is observed when comparing between individuals based on variant type. Clinical photos of hands and feet are provided in Figure S2 and additional clinical information for all individuals in this study is provided in Table S2 with a summary in Table 1.

(Fishers' exact tests), neither when discriminated by the type of variants nor for missense variants clustering to specific domains (Tables 1 and S1). Of note, individual 25, who had a *de novo* p.Gln642Glu, which is also located in the HEAT domain, did not have pancreatic agenesis and/or holoprosencephaly (Table S2), suggesting that variants leading to p.Arg535Cys represent a separate clinically recognizable entity within the spectrum of *CNOT1* (neuro)developmental disorders.

All types of genetic variation were observed in *CNOT1*: three different nonsense variants of which one was recurrent in four individuals; six unique splice site variants, of which one was observed twice in an index and mother; five unique frameshifts, one of which was observed twice in an index and mother; 15 different missense variants, one recurrent in five individuals and one individual with two *de novo* missense variants; and lastly two (partial) gene deletions (Figure 1; Table S1). Interestingly, the missense variants of 16 of 18 individuals suggested clustering to CNOT1 functional domains: six affected the HEAT domain, which modulates substrate specificity; two in the TTP-binding domain involved in deadenylation; four in the CAF1 domain, binding proteins with catalytic properties; four in a domain of unknown function (DUF3819); and one in the Not1 domain, which is associated with interaction of multiple protein partners (Figure S1). Effect prediction of these missense variants shows that they in general occur in evolutionary well-conserved regions (Figure S3) and that they may disturb hydrogen bonds, salt bridges, and structural stability, thereby often affecting interactions with protein binding partners (Table S3). In addition, the position of the missense variants, buried in the core of the protein, may affect protein folding and stability (Figure S3).

To functionally elucidate the pathophysiological effects of the observed *de novo* variants on CNOT1 scaffolding capacity (Figure S4), we generated eight different *CNOT1* variant constructs, including seven missense variants and one loss-of-function variant (Table S4) and transfected these in COS-1 cells. Using a PalmMyr-CFP-tagged construct, targeting CNOT1 to the cell membrane, we

Table 1. Summary of Clinical Characteristics Associated with *De Novo CNOT1* Individuals

	All Individuals (n = 39)		Individuals with Presumed Loss-of-Function Variants (n = 20)		Individuals with Missense Variants (n = 19)	
	% (present/assessed)		% (present/assessed)		% (present/assessed)	
Neurological Abnormalities	97	(38/39)	100	(20/20)	95	(18/19)
Intellectual disability	72	(23/32)	67	(12/18)	79	(11/14)
Normal	13	(4/32)	11	(2/18)	14	(2/14)
Borderline	16	(5/32)	22	(4/18)	7	(1/14)
Mild	34	(11/32)	28	(5/18)	43	(6/14)
Moderate	9	(3/32)	17	(3/18)	0	(0/14)
Severe	13	(4/32)	6	(1/18)	21	(3/14)
Unspecified	16	(5/32)	17	(3/18)	14	(2/14)
Developmental delay	92	(34/37)	95	(18/19)	89	(16/18)
Motor delay	83	(29/35)	89	(17/19)	75	(12/16)
Speech delay	83	(29/35)	79	(15/19)	88	(14/16)
Dysarthria	34	(10/29)	28	(5/18)	45	(5/11)
Epilepsy	25	(9/36)	22	(4/18)	28	(5/18)
Hypotonia	74	(26/35)	79	(15/19)	69	(11/16)
Behavioral disturbances	63	(20/32)	61	(11/18)	64	(9/14)
Sleep disturbances	32	(10/31)	28	(5/18)	38	(5/13)
Abnormal Brain Imaging	65	(20/31)	64	(9/14)	65	(11/17)
Holoprosencephaly	13	(4/31)	0	(0/14)	24	(4/17)
Other MRI findings	62	(18/29)	64	(9/14)	60	(9/15)
Abnormal[a] Growth	74	(28/38)	79	(15/19)	68	(13/19)
Abnormal term of delivery	16	(6/38)	21	(4/19)	11	(2/19)
Preterm (<37 weeks of gestation)	16	(6/38)	21	(4/19)	11	(2/19)
Postterm (>42 weeks of gestation)	0	(0/38)	0	(0/19)	0	(0/19)
Abnormal birth weight	23	(8/35)	17	(3/18)	29	(5/17)
Small for gestational age	23	(8/35)	17	(3/18)	29	(5/17)
Large for gestational age	0	(0/35)	0	(0/18)	0	(0/17)
Abnormal head circumference at birth	24	(4/17)	20	(2/10)	29	(2/7)
Decreased head circumference	18	(3/17)	10	(1/10)	29	(2/7)
Increased head circumference	6	(1/17)	10	(1/10)	0	(0/7)
Abnormal height	61	(20/33)	72	(13/18)	47	(7/15)
Short stature	55	(18/33)	72	(13/18)	33	(5/15)
Tall stature	6	(2/33)	0	(0/18)	13	(2/15)
Abnormal head circumference	26	(9/34)	28	(5/18)	25	(4/16)
Decreased head circumference	21	(7/34)	17	(3/18)	25	(4/16)
Increased head circumference	6	(2/34)	11	(2/18)	0	(0/16)
Abnormal weight	23	(7/31)	29	(5/17)	14	(2/14)
Underweight	19	(6/31)	24	(4/17)	14	(2/14)
Overweight	3	(1/31)	6	(1/17)	0	(0/14)

(*Continued on next page*)

Table 1. *Continued*

	All Individuals (n = 39)	Individuals with Presumed Loss-of-Function Variants (n = 20)	Individuals with Missense Variants (n = 19)
	% (present/assessed)	% (present/assessed)	% (present/assessed)
Other Abnormalities			
Facial abnormalities	92 (36/39)	100 (20/20)	84 (16/19)
Cardiac abnormalities	33 (8/24)	42 (5/12)	25 (3/12)
Urogenital abnormalities	39 (7/18)	63 (5/8)	20 (2/10)
Gastrointestinal abnormalities	50 (9/18)	57 (4/7)	45 (5/11)
Dysphagia/feeding difficulties	50 (14/28)	33 (5/15)	69 (9/13)
Pulmonal abnormalities	24 (4/17)	13 (1/8)	33 (3/9)
Immunological abnormalities	14 (2/14)	33 (2/6)	0 (0/8)
Endocrine abnormalities	56 (10/18)	50 (3/6)	58 (7/12)
Skeletal, muscle and soft tissue abnormalities	67 (16/24)	77 (10/13)	55 (6/11)
Hearing abnormalities	12 (3/25)	8 (1/13)	17 (2/12)
Vision abnormalities	28 (7/25)	38 (5/13)	17 (2/12)
Ectodermal abnormalities	52 (12/23)	64 (9/14)	33 (3/9)
Hands and feet abnormalities	64 (14/22)	64 (7/11)	64 (7/11)

Only features present in at least 10% of all individuals are listed.
[a]Abnormal growth parameters defined as $> +2$ SD and < -2 SD. Clinical details per individual are specified in Table S2.

confirmed co-localization, as proxy for interaction, of wild-type CNOT1 with its partners CNOT2, CNOT4, and CNOT8 (Figures S5). We next assessed whether the *de novo* missense variants in CNOT1 impacted these interactions. Hereto, each *CNOT1* variant was co-transfected with the interaction partner most likely to be disrupted (Figures S4 and S6; Table S5). Apart from one missense variant (p.Lys1241Arg), none of the variants seemed to affect binding of the respective CCR4-NOT1 interaction partner (Figure S6). Although the effects might be (1) more subtle than quantified here, (2) only identifiable in a different cell type system, or (3) dependent on the subunit composition which differs among tissues and during neural development,[21] we concluded that in our cell-based assays the functional consequence of *de novo* CNOT1 variants is not mediated by major changes in the composition of the CCR4-NOT1 complex.

The CCR4-NOT complex has an important function in cell viability, and it has been shown that functional depletion of CNOT1 results in endoplasmic reticulum stress-induced apoptosis.[9] We therefore set out to measure the consequence of *de novo CNOT1* variants on apoptosis. We used fluorescence-activated cell sorting (FACS) to discriminate viable cells from those in apoptosis, by using 7-AAD and Annexin V as markers. This failed to observe increased apoptosis in three patient-derived Epstein-Barr virus immortalized B-lymphoid cell lines (Figure S7).

With the CCR4-NOT1 complex being involved in mRNA deadenylation, we then reasoned that its role in endonucleolytic nonsense-mediated mRNA decay (NMD) might be impaired. Therefore, we performed QuantSeq 3′mRNA sequencing (Lexogen) for five individuals with a *de novo CNOT1* variant and 15 control subjects. As a proof-of-principle, we first evaluated the expression levels of *CNOT1* in individuals with a *de novo CNOT1* variant. Unexpectedly, the mRNA expression levels of individual 1 were normal, despite her *de novo* p.Arg26* variant (Figures S8A and S8B). This indicates that transcripts harboring the *de novo* p.Arg26* nonsense variant did not undergo NMD. By using allele-specific PCR primers for qPCR (Table S7), we determined that the wild-type and the mutant allele were indeed equally expressed (Figure S8C). While this observation would support the recently suggested non-canonical rule of NMD escape if the premature termination codon is located <150 nt from the start codon (start-proximal rule),[22] this rule does not explain the observed normal levels of full-length CNOT1 protein by using CNOT1 antibodies against the C terminus (Figure S8D). Since the second methionine is located 234 amino acid residues downstream of the canonical one, this suggests translational upregulation of the single normal *CNOT1* allele. Of note, any CNOT1 protein translated from the premature termination codon-containing allele escaping NMD would be only a maximum of 33 amino acids in size, and in the absence of an antibody directed against this, part of the protein remains undetectable.

There are two NMD pathways by which the cell degrades RNA.[23] One is mainly performed by SMG6-dependent endocleavage, which degrades aberrant RNA with a premature stop codon. The other goes by SMG7-dependent exonucleolytic activity, which is usually directed toward transcripts with an upstream open reading frame or a long 3′ untranslated region. The latter is part of normal regulation of RNA

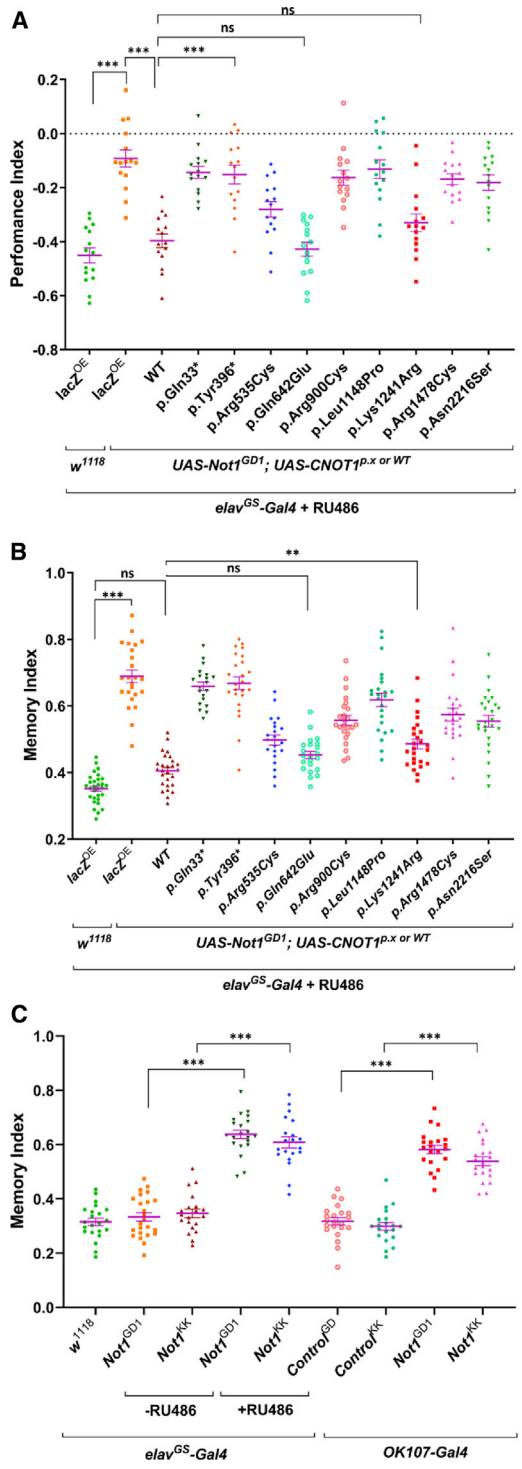


Figure 2. Disease-Associated Variants of Human CNOT1 Induce Memory Deficits in *Drosophila* Model

(A) Depletion of *Drosophila Not1* function in all neurons by inducible RNAi knockdown in adults only by using two different UAS-RNAi lines – $Not1^{GD1}$ and $Not1^{KK}$ with $elav^{GS}$-Gal4 and exposure to RU (GeneSwitch[24]), or specifically in the mushroom body with OK107-Gal4, produces memory deficits.

(B and C) RU-induced expression of disease-associated human CNOT1 variant cDNAs (UAS-CNOT1$^{p.X\ or\ WT}$) in a pan-neuronal $Not1^{GD1}$–RNAi knockdown background (using $elav^{GS}$-Gal4) induces memory deficits both during adult (B) and larval (C) stages. One-way ANOVA, **$p < 0.01$, ***$p < 0.0001$.

expression, also referred to as "regulatory NMD." We integrated the exome-sequencing data of five individuals with *de novo CNOT1* variants with 3′mRNA Quantseq data, where analysis was aimed at the identification of SMG6-dependent NMD failure as a consequence of CCR4-NOT dysfunction. Hence, we focused on nonsense and frameshift variants (Figure S9), which are under normal physiological circumstances degraded by the endonucleolytic NMD pathway, but in individuals with *de novo CNOT1* mutations would not. For all variants identified, we cross-referenced with data in GTEX to confirm that indeed the expression in controls with nonsense and frameshift variants in these genes show a reduced expression, fitting our hypothesis of NMD. This led to the identification of only three genes with a nonsense or frameshift variant in control subjects and individuals with *de novo CNOT1* variants for which we were able to assess whether these genes showed reduced (proxy for NMD functional expression) or normal expression (proxy for NMD dysfunctional gene expression; Table S7 and Table S8). For these, no differences in endonucleolytic NMD were observed between control subjects and individuals with *de novo CNOT1* variants (Figure S10). This suggests no influence of the five *CNOT1* variants on SMG6-dependent endonucleolytic NMD, although effects at spatio-temporal level and/or subtle, more global (transcript-dependent) NMD effects cannot be excluded.

We next set out to study the effects of individual variants in a model organism, as our assays did not yield a convincing pathophysiological mechanism to disease. Hereto, we chose *Drosophila melanogaster*, in which Not1 is the ortholog of human CNOT1, showing 62% overall identity (Figure S11A). To test the effect of *de novo CNOT1* variants on learning and memory, we used the UAS-GAL4 system in two different RNAi lines targeting *Not1* (GD12571 for UAS-$Not1^{GD1}$ and KK196587 for UAS-$Not1^{KK}$), as well as for two different Gal4-driver lines, *elav* and OK107, for pan-neuronal or (memory-pivotal) mushroom body-specific knockdown, respectively, and assessed memory in a courtship suppression behavior (Figure S11B). As constitutive *Not1* knockdown by *elav*-Gal4 led to early lethality after eclosion (Figure S12A), we adopted the *elav*-Gal4 GeneSwitch system ($elav^{GS}$) inducible by RU486.[24] Both RU induced pan-neuronal knockdown in adult flies and mushroom body-specific knockdown (OK107) of *Not1* with both RNAi lines, showed a statistically increased memory index indicating impaired learning and memory in those flies (Figures 2A, S12A, and S12C). We next aimed to rescue the learning and memory phenotypes by introducing human wild-type *CNOT1* cDNA, as well as by nine of the *de novo* variants identified in this study, which includes two loss-of-function variants and seven missense variants, all targeting different functional domains (Figure S11C). As both RNAi lines showed similar results upon *Not1* knockdown, we only assessed the rescue of memory and learning loss by wild-type and variants of CNOT1 with the RU486-induced ($elav^{GS}$>$Not1^{GD1}$) knockdown line. When human wild-type *CNOT1* is introduced,

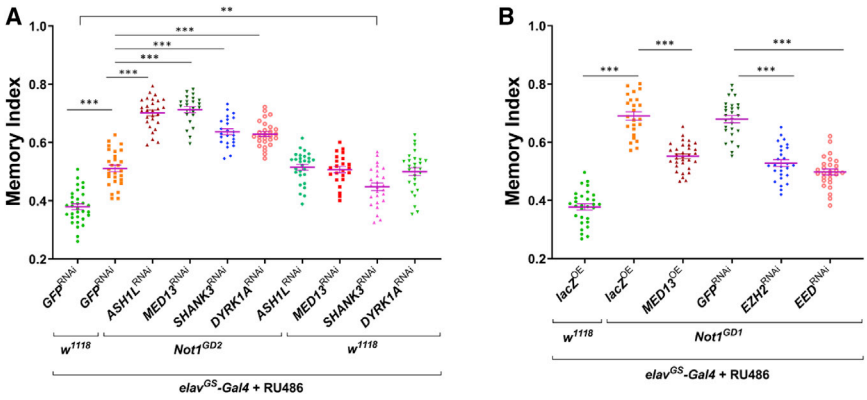


Figure 3. Genetic Interactions between *Not1* and Known ID/ASD Genes and Rescue of Neuronal *Not1* Knockdown-Induced Neurodevelopmental Defects with Transcriptional Modifiers

(A) Neuronal RNAi knockdown of known ID/ASD genes (*ASH1*, *DYRK1A*, *MED13*, *SHANK3*; in flies *ash1*, *mnb*, *skd*, *Prosap*, respectively) exacerbates *Not1* knockdown-induced memory defects in adults. (B) Overexpression of *MED13* or knockdown of PRC2 components *E(z)* (*EZH2* in humans) and *esc* (*EED* in humans) ameliorates *Not1* knockdown-induced memory deficits in adults. One-way ANOVA, **$p < 0.01$, ***$p < 0.0001$.

a statistically significant rescue of the memory phenotype was observed, whereas for eight of the nine mutants the phenotype could not or only partially be rescued (Figure 2B). The impact of these variants on learning and memory was further substantiated by two additional memory tests performed at larval stages, i.e., aversive and appetitive associative learning (Figures 2C and S12D). These findings suggest that the *de novo* CNOT1 mutations as observed in this study are likely sufficient to cause a neurodevelopmental disorder.

In order to test for a potential link between *CNOT1* to genes previously implicated to impact ID/ASD cognitive functions, we used a weaker *Not1* RNAi line, which produces only moderate memory defects, in conjunction with knockdown of *ash1*, *MED13*, *shank*, and *Dyrk1a* for the human genes *ASH1L* (MIM: 617796), *MED13* (MIM: 618009), *SHANK3* (MIM: 606232), and *DYRK1A* (MIM: 614104).[25–27] Whereas we find that pan-neuronal knockdown of *ash1*, *MED13*, *shank*, and *Dyrk1a* by themselves also caused moderate memory defects, the defects are greatly enhanced by weaker *Not1* knockdown ($elav^{GS} > Not1^{GD2}$; Figure 3A). Conversely, pan-neuronal overexpression of *MED13* or knockdown of potential mediators of neurodevelopmental disorders (i.e., Polycomb Repressor Complex 2 [PRC2] components, *E(z)* and *esc*, for human *EZH2* and *EED*)[28,29] significantly reduces the severe memory defect of strong *Not1* knockdown (Figure 3B). These results were further supported by an independent larval olfactory-mediated aversive-associative learning test (Figure S13): knockdown of *ash1*, *MED13*, *shank*, and *Dyrk1a* exacerbates the moderate larval memory deficits induced by the weaker *Not1*-RNAi line (Figure S13A). Conversely, overexpression of *MED13* or reduction of PRC2 components rescued the memory phenotypes upon *Not1* knockdown in larval brains (Figure S13B). Thus, in addition to its role in causing ID/ASD-associated developmental cognitive dysfunctions, our genetic data also implicate CNOT1 as a candidate in contributing to other neurodevelopmental phenotypes.

In conclusion, we have collected 39 individuals with a neurodevelopmental disorder who have a broad clinical phenotypic spectrum, ranging from individuals with severe ID with co-morbidities such as seizures, hypotonia, and behavioral problems, to individuals with borderline normal IQ and normal everyday functioning. Functional characterization of the (*de novo*) variants at both the level of RNA and protein of *CNOT1*, as well as its functional roles in important biological processes such as mRNA decay and cell viability, have not revealed the pathophysiological mechanisms underlying this novel neurodevelopmental disorder. Nonetheless, knockdown and rescue experiments in *Drosophila* have indicated the plausibility of the pathogenicity of the variants on neurodevelopment, which is further strengthened by the interaction studies with autism-spectrum genes. Given the importance of CNOT1 in many essential biological processes, it is anticipated that *de novo* variants in *CNOT1* impact its normal function in a more complex manner, involving either transcript- and/or tissue-dependent hypomorphic or neomorphic alleles.

Supplemental Data

Supplemental Data can be found online at https://doi.org/10.1016/j.ajhg.2020.05.017.

Acknowledgments

We wish to thank all families participating in this study. In addition, we wish to thank the members of the Genome Technology Center and Cell culture facility, Department of Human Genetics, Radboud University Medical Center, Nijmegen, for data processing and cell culture of patient-derived cell lines, as well as the Radboud Technology Center for Flow Cytometry for help in FACS sorting. This work was financially supported by an Aspasia grant of the Dutch Research Council (015.014.066 to L.E.L.M.V.). D.L.P. is recipient of a CAPES Fellowship (99999.013311/2013-01). F.L.R. is funded from the Cambridge Biomedical Centre. In addition, the collaborations in this study were facilitated by the ERN ITHACA, one of the 24 European Reference Networks (ERNs) approved by the ERN Board of Member States, co-funded by European Commission. For more information about the ERNs and the EU health strategy please visit https://ec.europa.eu/health/ern. The aims of this study contribute to the Solve-RD project (to L.E.L.M.V.) which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 779257.

Declaration of Interests

A.B. and K.M. are, and M.T.C. was, an employee of GeneDx, Inc.

Received: April 9, 2020
Accepted: May 26, 2020
Published: June 17, 2020

Web Resources

GenBank, https://www.ncbi.nlm.nih.gov/genbank/
gnomAD, https://gnomad.broadinstitute.org/
GTEX, https://gtexportal.org/home/index.html
HOPE, https://www3.cmbi.umcn.nl/hope/
Metadome, https://stuart.radboudumc.nl/metadome/about
OMIM, https://www.omim.org/
YASARA, http://www.yasara.org/

References

1. Relaix, F. (2015). Pax genes: Master regulators of development and tissue homeostasis. Semin. Cell Dev. Biol. *44*, 62–63.
2. Kamachi, Y., and Kondoh, H. (2013). Sox proteins: regulators of cell fate specification and differentiation. Development *140*, 4129–4144.
3. Shirai, Y.T., Suzuki, T., Morita, M., Takahashi, A., and Yamamoto, T. (2014). Multifunctional roles of the mammalian CCR4-NOT complex in physiological phenomena. Front. Genet. *5*, 286.
4. Ukleja, M., Valpuesta, J.M., Dziembowski, A., and Cuellar, J. (2016). Beyond the known functions of the CCR4-NOT complex in gene expression regulatory mechanisms: New structural insights to unravel CCR4-NOT mRNA processing machinery. BioEssays *38*, 1048–1058.
5. Albert, T.K., Hanzawa, H., Legtenberg, Y.I., de Ruwe, M.J., van den Heuvel, F.A., Collart, M.A., Boelens, R., and Timmers, H.T. (2002). Identification of a ubiquitin-protein ligase subunit within the CCR4-NOT transcription repressor complex. EMBO J. *21*, 355–364.
6. Collart, M.A. (2016). The Ccr4-Not complex is a key regulator of eukaryotic gene expression. Wiley Interdiscip. Rev. RNA *7*, 438–454.
7. Azzouz, N., Panasenko, O.O., Deluen, C., Hsieh, J., Theiler, G., and Collart, M.A. (2009). Specific roles for the Ccr4-Not complex subunits in expression of the genome. RNA *15*, 377–383.
8. Chen, Y., Boland, A., Kuzuoğlu-Öztürk, D., Bawankar, P., Loh, B., Chang, C.T., Weichenrieder, O., and Izaurralde, E. (2014). A DDX6-CNOT1 complex and W-binding pockets in CNOT9 reveal direct links between miRNA target recognition and silencing. Mol. Cell *54*, 737–750.
9. Ito, K., Takahashi, A., Morita, M., Suzuki, T., and Yamamoto, T. (2011). The role of the CNOT1 subunit of the CCR4-NOT complex in mRNA deadenylation and cell viability. Protein Cell *2*, 755–763.
10. Winkler, G.S., Mulder, K.W., Bardwell, V.J., Kalkhoven, E., and Timmers, H.T. (2006). Human Ccr4-Not complex is a ligand-dependent repressor of nuclear receptor-mediated transcription. EMBO J. *25*, 3089–3099.
11. Fabian, M.R., Cieplak, M.K., Frank, F., Morita, M., Green, J., Srikumar, T., Nagar, B., Yamamoto, T., Raught, B., Duchaine, T.F., and Sonenberg, N. (2011). miRNA-mediated deadenylation is orchestrated by GW182 through two conserved motifs that interact with CCR4-NOT. Nat. Struct. Mol. Biol. *18*, 1211–1217.
12. Suzuki, A., Saba, R., Miyoshi, K., Morita, Y., and Saga, Y. (2012). Interaction between NANOS2 and the CCR4-NOT deadenylation complex is essential for male germ cell development in mouse. PLoS ONE *7*, e33558.
13. Murakawa, Y., Hinz, M., Mothes, J., Schuetz, A., Uhl, M., Wyler, E., Yasuda, T., Mastrobuoni, G., Friedel, C.C., Dölken, L., et al. (2015). RC3H1 post-transcriptionally regulates A20 mRNA and modulates the activity of the IKK/NF-κB pathway. Nat. Commun. *6*, 7367.
14. Fabian, M.R., Frank, F., Rouya, C., Siddiqui, N., Lai, W.S., Karetnikov, A., Blackshear, P.J., Nagar, B., and Sonenberg, N. (2013). Structural basis for the recruitment of the human CCR4-NOT deadenylase complex by tristetraprolin. Nat. Struct. Mol. Biol. *20*, 735–739.
15. Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat. Commun. *7*, 12626.
16. Wiel, L., Baakman, C., Gilissen, D., Veltman, J.A., Vriend, G., and Gilissen, C. (2019). MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. Hum. Mutat. *40*, 1030–1038.
17. De Franco, E., Watson, R.A., Weninger, W.J., Wong, C.C., Flanagan, S.E., Caswell, R., Green, A., Tudor, C., Lelliott, C.J., Geyer, S.H., et al. (2019). A Specific CNOT1 Mutation Results in a Novel Syndrome of Pancreatic Agenesis and Holoprosencephaly through Impaired Pancreatic and Neurological Development. Am. J. Hum. Genet. *104*, 985–989.
18. Kruszka, P., Berger, S.I., Weiss, K., Everson, J.L., Martinez, A.F., Hong, S., Anyane-Yeboa, K., Lipinski, R.J., and Muenke, M. (2019). A CCR4-NOT Transcription Complex, Subunit 1, CNOT1, Variant Associated with Holoprosencephaly. Am. J. Hum. Genet. *104*, 990–993.
19. 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.
20. Kaplanis, J., Samocha, K.E., Wiel, L., Zhang, Z., Arvai, K.J., Eberhardt, R.Y., Gallone, G., Lelieveld, S.H., Martin, H.C., McRae, J.F., et al. (2019). Integrating healthcare and research genetic data empowers the discovery of 49 novel developmental disorders. bioRxiv. https://doi.org/10.1101/797787.
21. Chen, C., Ito, K., Takahashi, A., Wang, G., Suzuki, T., Nakazawa, T., Yamamoto, T., and Yokoyama, K. (2011). Distinct expression patterns of the subunits of the CCR4-NOT deadenylase complex during neural development. Biochem. Biophys. Res. Commun. *411*, 360–364.
22. Hoek, T.A., Khuperkar, D., Lindeboom, R.G.H., Sonneveld, S., Verhagen, B.M.P., Boersma, S., Vermeulen, M., and Tanenbaum, M.E. (2019). Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. Mol. Cell *75*, 324–339.e11.
23. Ottens, F., Boehm, V., Sibley, C.R., Ule, J., and Gehring, N.H. (2017). Transcript-specific characteristics determine the contribution of endo- and exonucleolytic decay pathways during the degradation of nonsense-mediated decay substrates. RNA *23*, 1224–1236.

24. McGuire, S.E., Roman, G., and Davis, R.L. (2004). Gene expression systems in Drosophila: a synthesis of time and space. Trends Genet. *20*, 384–391.
25. van Bon, B.W., Coe, B.P., Bernier, R., Green, C., Gerdts, J., Witherspoon, K., Kleefstra, T., Willemsen, M.H., Kumar, R., Bosco, P., et al. (2016). Disruptive de novo mutations of DYRK1A lead to a syndromic form of autism and ID. Mol. Psychiatry *21*, 126–132.
26. De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al.; DDD Study; Homozygosity Mapping Collaborative for Autism; and UK10K Consortium (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature *515*, 209–215.
27. Snijders Blok, L., Hiatt, S.M., Bowling, K.M., Prokop, J.W., Engel, K.L., Cochran, J.N., Bebin, E.M., Bijlsma, E.K., Ruivenkamp, C.A.L., Terhal, P., et al.; DDD study (2018). De novo mutations in MED13, a component of the Mediator complex, are associated with a novel neurodevelopmental disorder. Hum. Genet. *137*, 375–388.
28. Moccia, A., and Martin, D.M. (2018). Nervous system development and disease: A focus on trithorax related proteins and chromatin remodelers. Mol. Cell. Neurosci. *87*, 46–54.
29. Kuehner, J.N., Bruggeman, E.C., Wen, Z., and Yao, B. (2019). Epigenetic Regulations in Neuropsychiatric Disorders. Front. Genet. *10*, 268.

The American Journal of Human Genetics, Volume *107*

Supplemental Data

De Novo Variants in *CNOT1*, a Central Component of the CCR4-NOT Complex Involved in Gene Expression and RNA and Protein Stability, Cause Neurodevelopmental Delay

Lisenka E.L.M. Vissers, Sreehari Kalvakuri, Elke de Boer, Sinje Geuer, Machteld Oud, Inge van Outersterp, Michael Kwint, Melde Witmond, Simone Kersten, Daniel L. Polla, Dilys Weijers, Amber Begtrup, Kirsty McWalter, Anna Ruiz, Elisabeth Gabau, Jenny E.V. Morton, Christopher Griffith, Karin Weiss, Candace Gamble, James Bartley, Hilary J. Vernon, Kendra Brunet, Claudia Ruivenkamp, Sarina G. Kant, Paul Kruszka, Austin Larson, Alexandra Afenjar, Thierry Billette de Villemeur, Kimberly Nugent, the DDD Study, F. Lucy Raymond, Hanka Venselaar, Florence Demurger, Claudia Soler-Alfonso, Dong Li, Elizabeth Bhoj, Ian Hayes, Nina Powell Hamilton, Ayesha Ahmad, Rachel Fisher, Myrthe van den Born, Marjolaine Willems, Arthur Sorlin, Julian Delanne, Sebastien Moutton, Philippe Christophe, Frederic Tran Mau-Them, Antonio Vitobello, Himanshu Goel, Lauren Massingham, Chanika Phornphutkul, Jennifer Schwab, Boris Keren, Perrine Charles, Maaike Vreeburg, Lenika De Simone, George Hoganson, Maria Iascone, Donatella Milani, Lucie Evenepoel, Nicole Revencu, D. Isum Ward, Kaitlyn Burns, Ian Krantz, Sarah E. Raible, Jill R. Murrell, Kathleen Wood, Megan T. Cho, Hans van Bokhoven, Maximilian Muenke, Tjitske Kleefstra, Rolf Bodmer, and Arjan P.M. de Brouwer

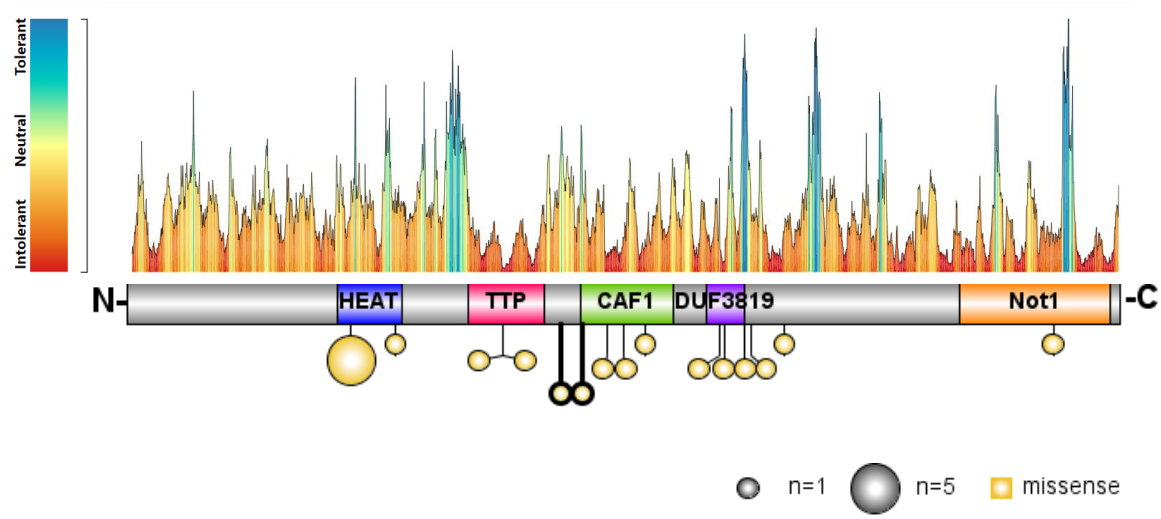
Supplemental Information

De novo variants in *CNOT1*, a central component of the CCR4-NOT complex involved in gene expression and RNA and protein stability, cause neurodevelopmental delay

Supplemental Figures

Supplemental Figure 1: Graphical representation of the CNOT1 missense variants in relation to the CNOT1 tolerance landscape

A

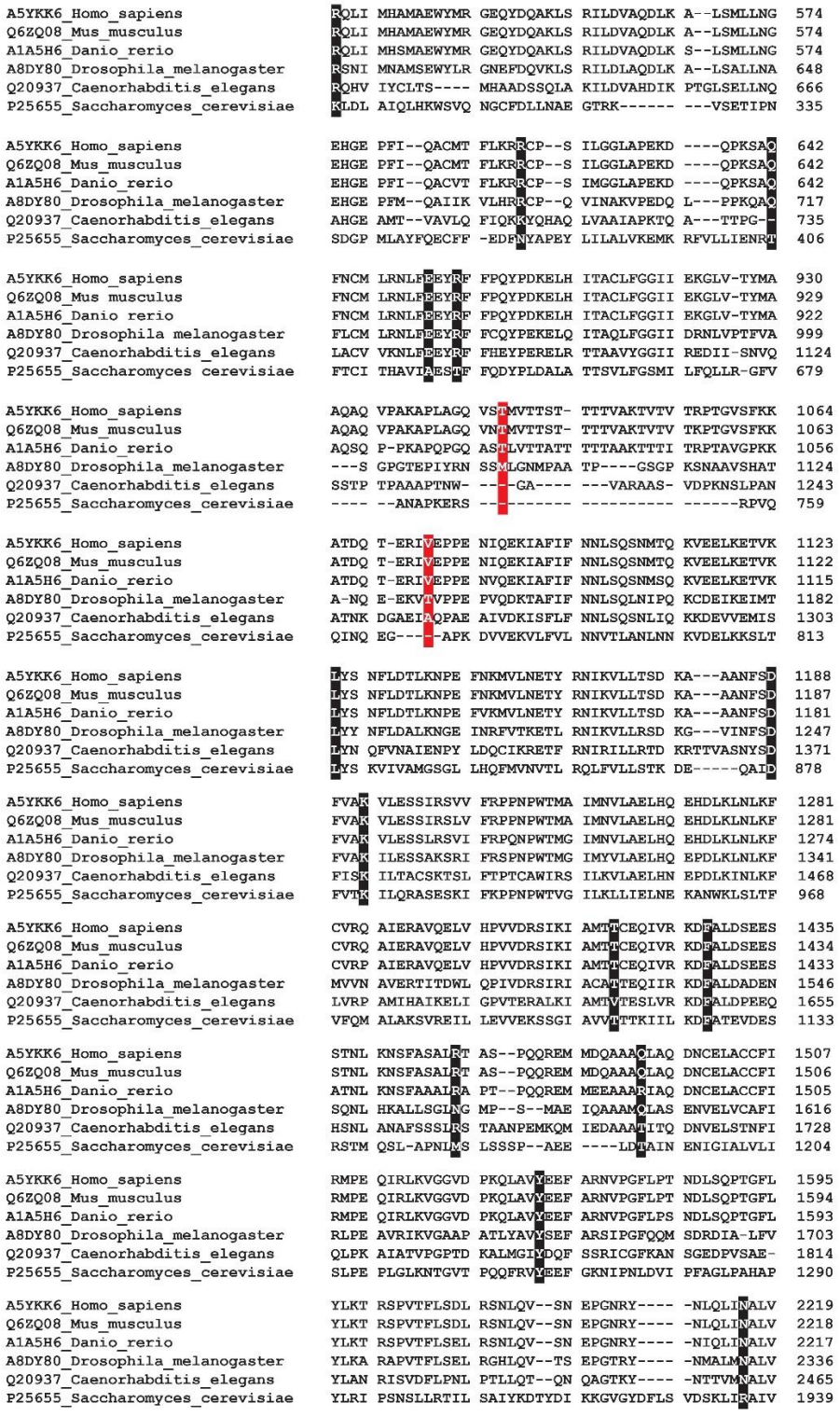


(**A**) Graphical representation of the linear protein structure of CNOT1 with the functional domains indicated by coloured boxes (blue: HEAT; pink: TTP; green: CAF; purple: DUF3819 and orange: Not1). Missense variants observed in this study are indicated by circles, in which the size of the circle corresponding to the number of recurrences of a given variant, and the colour of the circle the type of mutation. Note that Individual 28 was shown to have two *de novo* missense mutations, which are here shown by a bold outline. The tolerance landscape of CNOT1 was generated using Metadome (https://stuart.radboudumc.nl/metadome)[1], which is based on the non-synonymous over synonymous ratio in gnomAD. The colour in the plot is an indication for the tolerance (red: intolerant, blue: tolerant). The linear protein structure of CNOT1 with the missense variants is depicted underneath the tolerance profile. Together these indicate that the missense variants identified in CNOT1 are located in regions that are mostly intolerable for variation.

Supplemental Figure 2: Photos of hands and feet of Individuals with (*de novo*) *CNOT1* variants

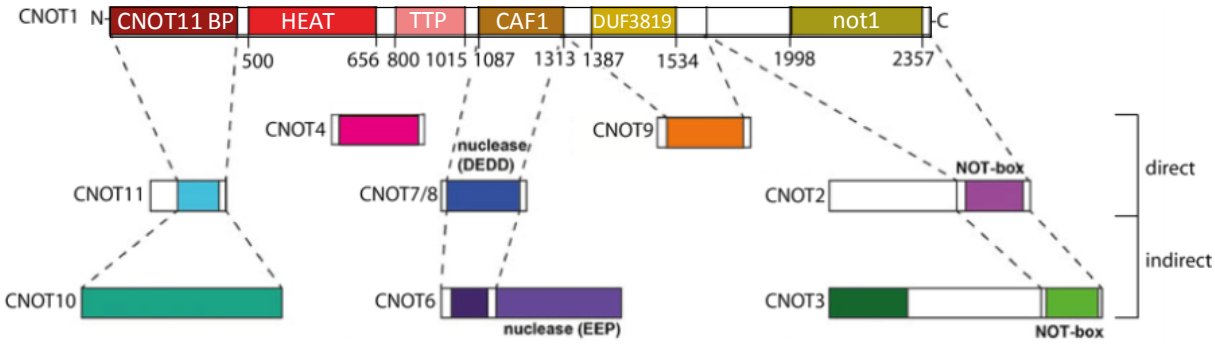
Hands and feet of Individuals with (likely) pathogenic *CNOT1* variants are shown, with individuals grouped based on the type of mutation identified. A variety of subtle abnormalities of hands and feet can be observed in the cohort, e.g. clinodactyly (Individual 7, 12, 18, 29, 39), tapered fingers (Individual 1, 19, 30), small hands (Individual 6, 7, 12 18, 19), broad forefeet with short broad toes (Individual 7, 12, 30). The subtle abnormalities were detected in both individuals with a presumed loss-of-function variant as well as in individuals with a missense variant. The green line below photos of individuals missense variants indicates that variants were located in the CAF domain.

Supplemental Figure 3: Evolutionary conservation of missense variants observed in CNOT1.



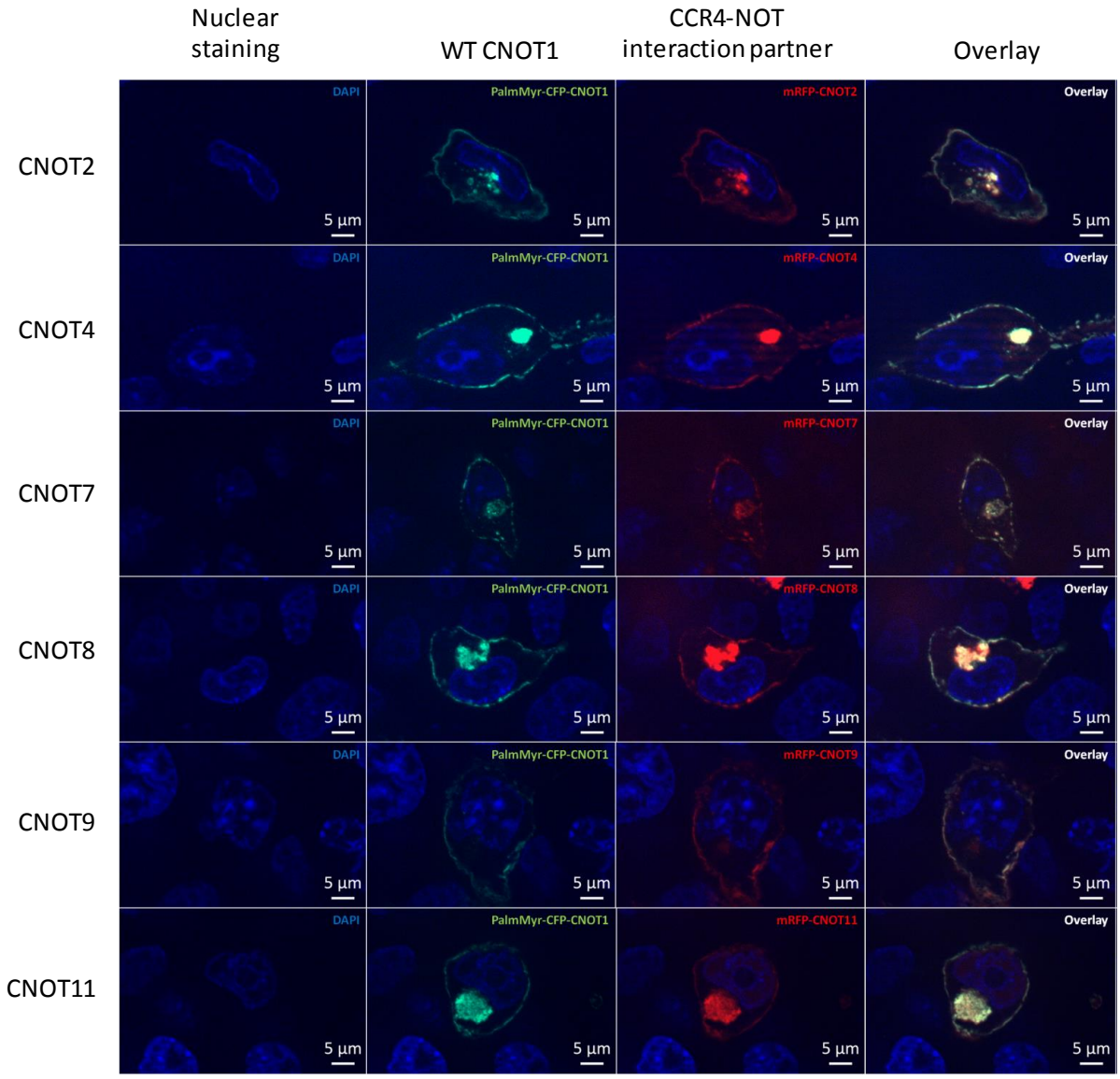
Positions of the *CNOT1* variants in CCR4-NOT transcription complex subunit 1. Cross-species alignment by Clustal Omega[2] of the protein sequences directly surrounding the missense variants. The unique changes are highlighted by the black boxes. The two *de novo* changes identified in the same individual are indicated by the red boxes. Protein accession numbers used for alignment are given before the sequences and include the specific species. The position of the last amino acid residue in each row is given right after the respective sequences.

Supplemental Figure 4: Direct and indirect CNOT1 CCR4-NOT interaction partners



Schematic representation of interaction partners of CNOT1 in its scaffolding function in the CCR4-NOT complex. Direct interaction partners include CNOT11, binding to the CNOT11-binding domain (CNOT11 BP), CNOT7/8 to the CAF1 domain, CNOT9 to the DUF3819 domain and CNOT2 to the Not1 domain. In addition, through indirect interactions, CNOT1 fulfils a scaffolding function for CNOT10, CNOT6 and CNOT3. Prior to this study, it was unknown whether CNOT4 directly or indirectly binds CNOT1.

Supplemental Figure 5: Co-localization of wildtype CNOT1 and wildtype CCR4-NOT1 interaction partners in COS-1 cells



Prior to assessing the (co-)localization of wildtype (WT) CNOT1 and its CCR4-NOT1 interaction partners, transfection rates were determined. Approximately 10% of all cells was transfected (CNOT1) and between 28-53% of the transfected cells was double transfected (CFP-CNOT1+ CCR4-NOT interaction partner). For these double transfected cells, colocalization was studied. The palmMyr-CFP-CNOT1 (green) was targeted to the cell membrane and co-localized with the mRFP-labeled CNOT2, 4, 7, 8, 9 and 11 (red), respectively. Of note, in addition to the membrane localization, partners CNOT7, 8, 9 and 11 also located to the nucleus and cytosol.

Supplemental Figure 6: CNOT1 missense variants do not seemingly disrupt binding of CCR4-NOT1 interaction partners

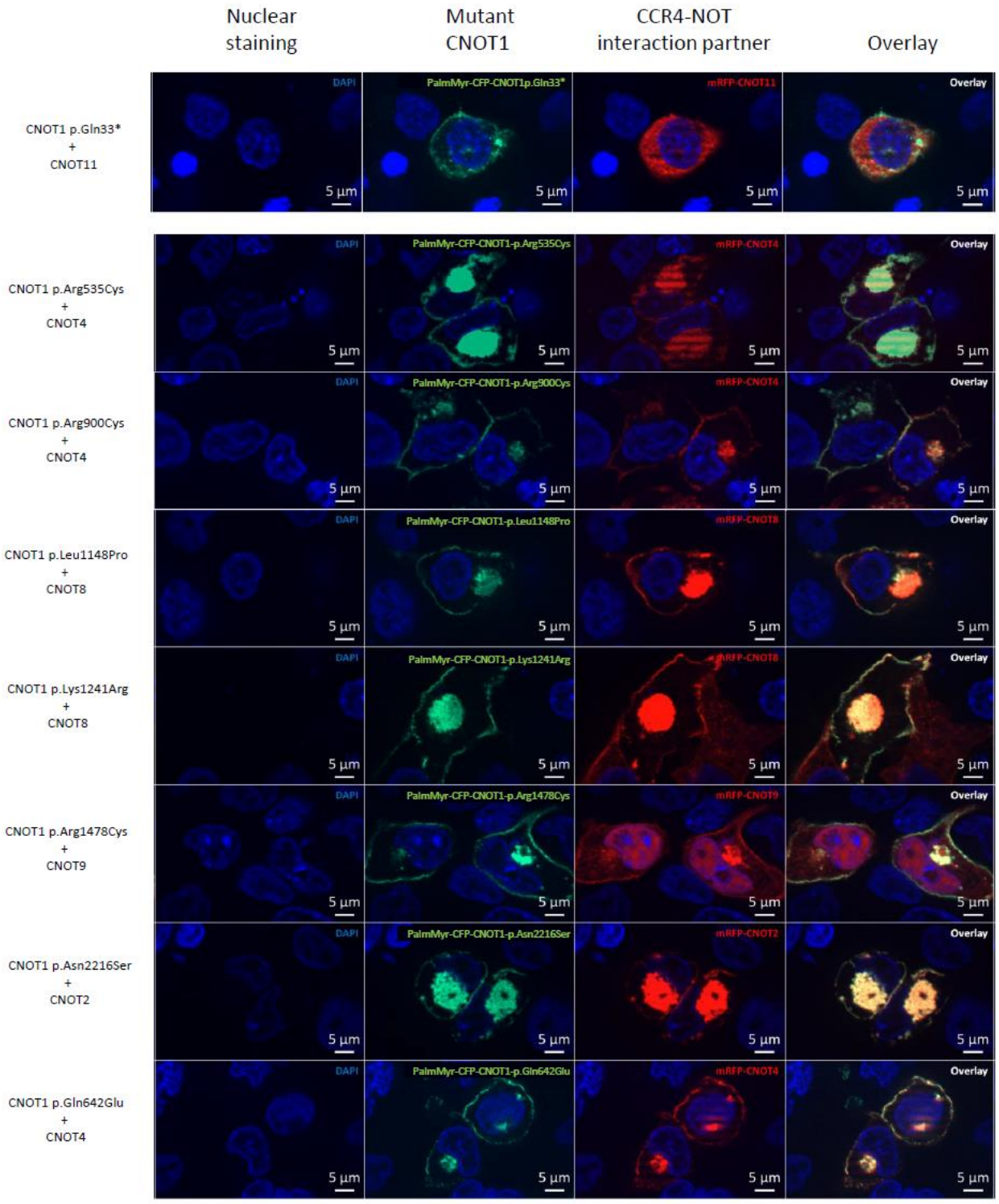
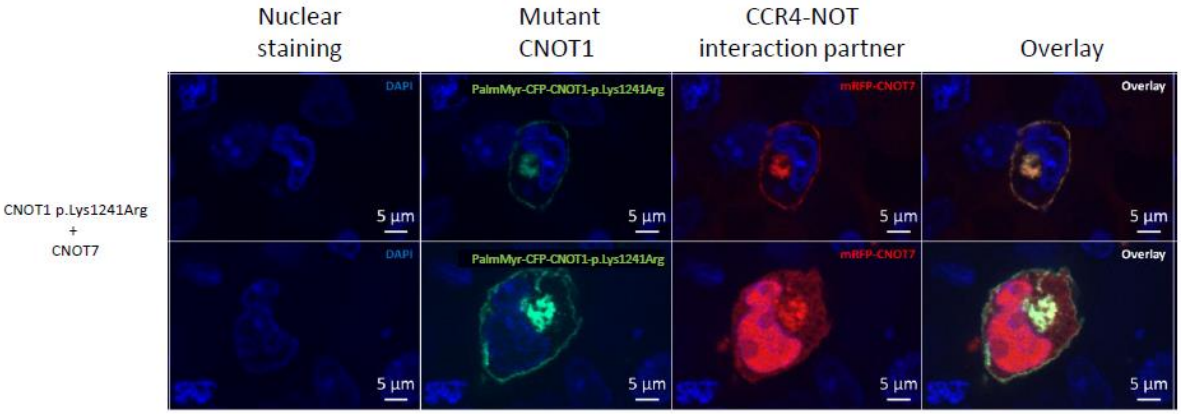


Figure S6 continued



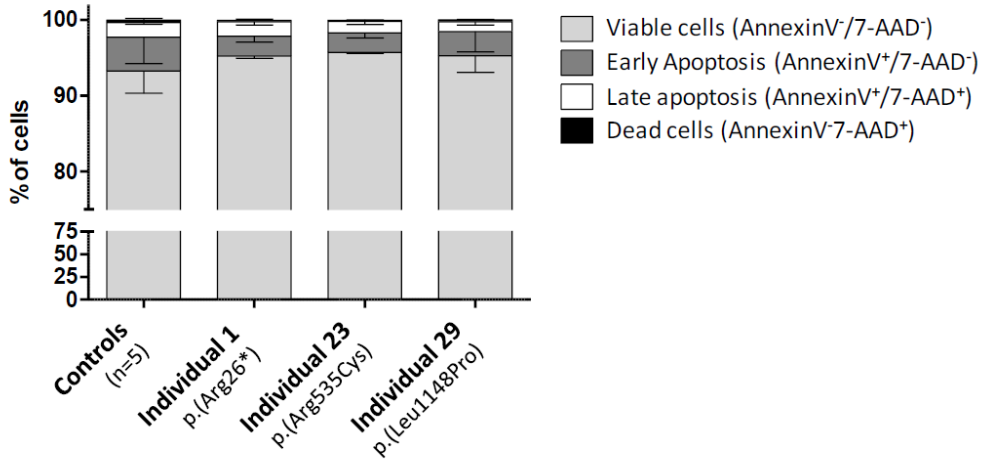
Co-transfection of mutant CNOT1 and CNOT2, 4, 8, 9 and 11 in COS-1 cells. PalmMyr-CFP-CNOT1-mut (green) were targeted to the cell membrane and transfected with RFP-tagged (red) subunits. Top panel shows the CNOT1-p.Gln33* with its interaction partner CNOT11. The variant clearly disrupts co-localization. Whereas this nonsense variant is expected to be degraded by nonsense mediated RNA decay in patient-derived cells, the loss of interaction shows the validity of the assay. For all (but one) missense variants, co-localization is not affected. For the co-transfection of CNOT1-Lys1241Arg with CNOT7, we noted that the interaction between was disturbed in approximately half of the cells, while in the other part the interaction remained intact.

Supplemental Figure 7: Apoptosis assay in CNOT1-mutant patient-derived EBV cell lines

A

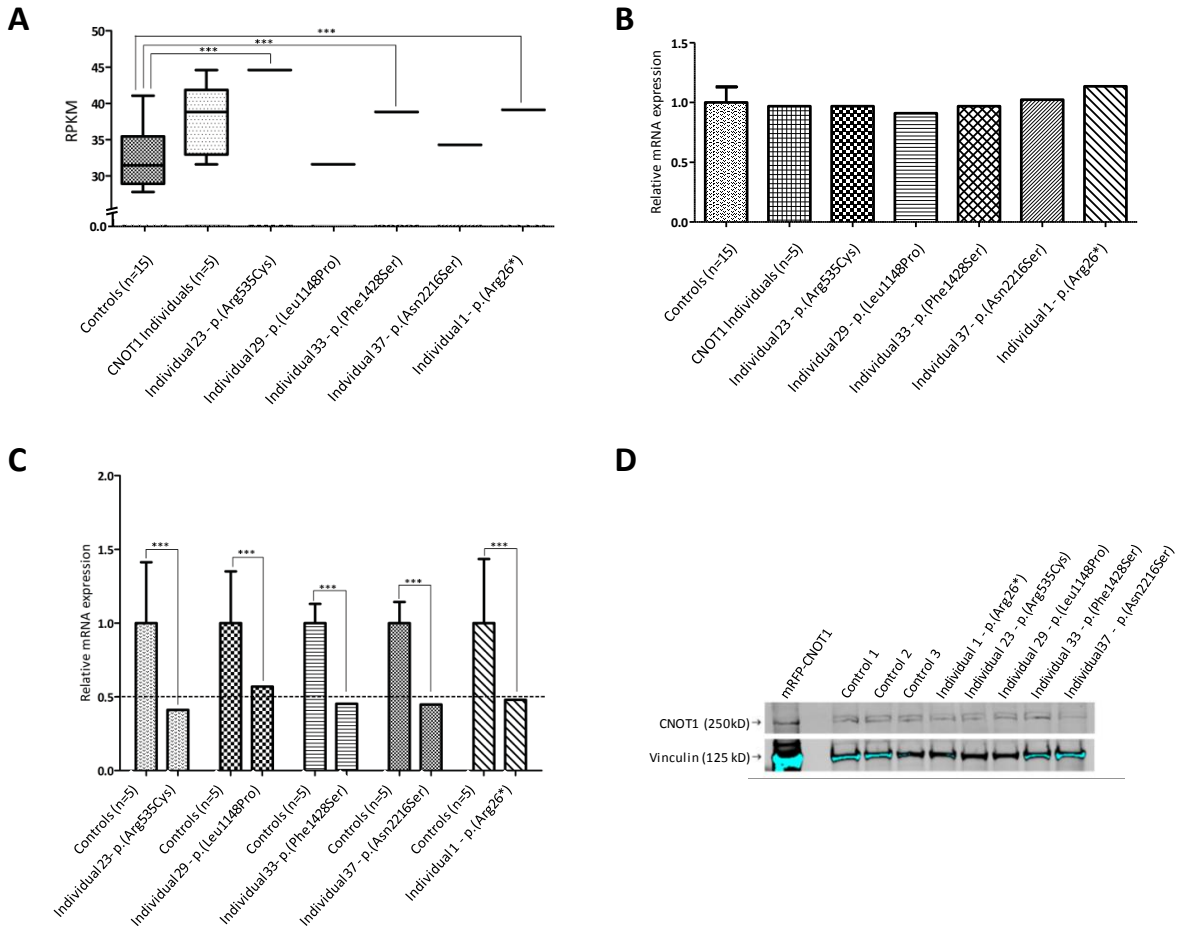
	AnnexinV $^-$ / 7-AAD $^-$ Alive cells (%)		AnnexinV $^+$ / 7-AAD $^-$ Early apoptosis (%)		AnnexinV $^+$ / 7-AAD $^+$ Late apoptosis (%)		AnnexinV $^-$ / 7-AAD $^+$ Dead cells (%)	
	exp 1	exp 2	exp 1	exp 2	exp 1	exp 2	exp 1	exp 2
Controls (n=5)	95.4	91.2	2.0	6.9	2.1	1.8	0.5	0.1
Individual 1 - p.(Arg26*)	95.4	95.0	2.1	3.2	2.2	1.6	0.3	0.1
Individual 23 - p.(Arg535Cys)	95.8	95.6	2.1	3.1	1.8	1.2	0.2	0.1
Individual 29 - p.(Leu1148Pro)	96.8	93.7	1.3	5.1	1.6	1.0	0.3	0.2

B



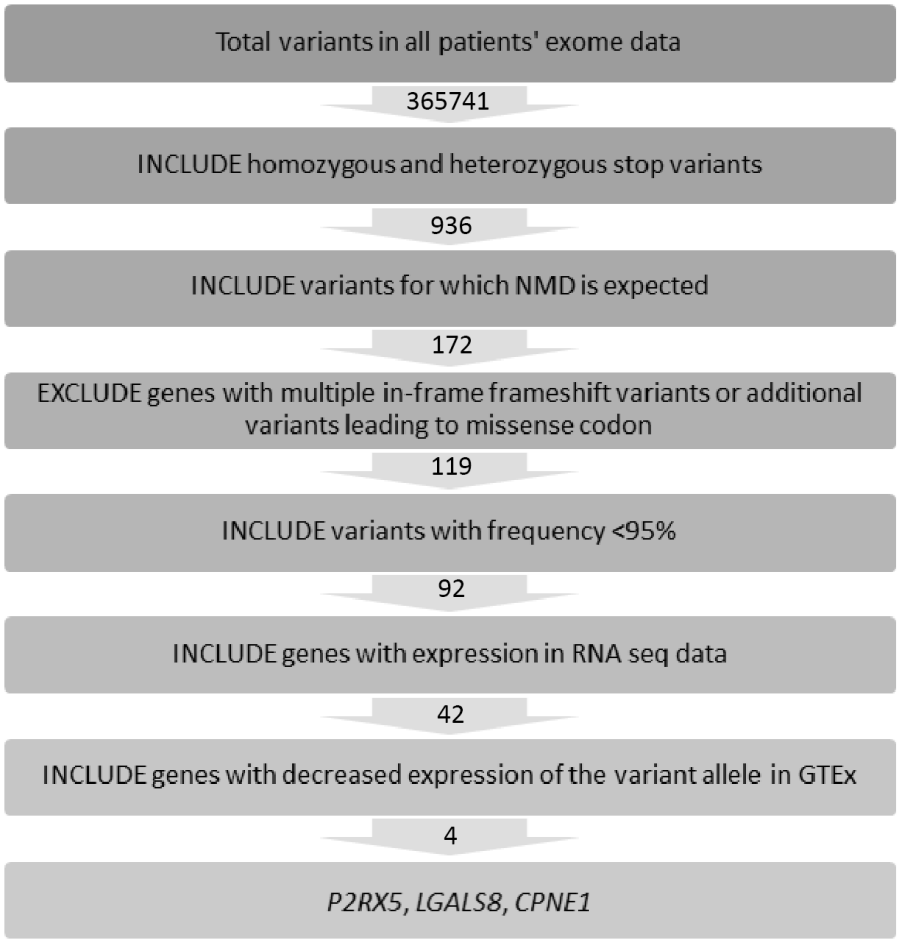
(**A**) FACS results denoting the percentage of cells in various apoptotic stages using Annexin V staining for apoptotic cells, and 7-AAD for viability. The experiment was repeated twice (exp 1 and exp 2). (**B**) Graphical representation of FACS results presented in A. From these data, it was concluded that there is no significant increase in apoptosis in patient-derived EBV-LCLs with *de novo CNOT1* variants when compared to controls (n=5).

Supplemental Figure 8: CNOT1 RNA and protein levels in Individuals with a CNOT1 de novo variant do not differ from controls



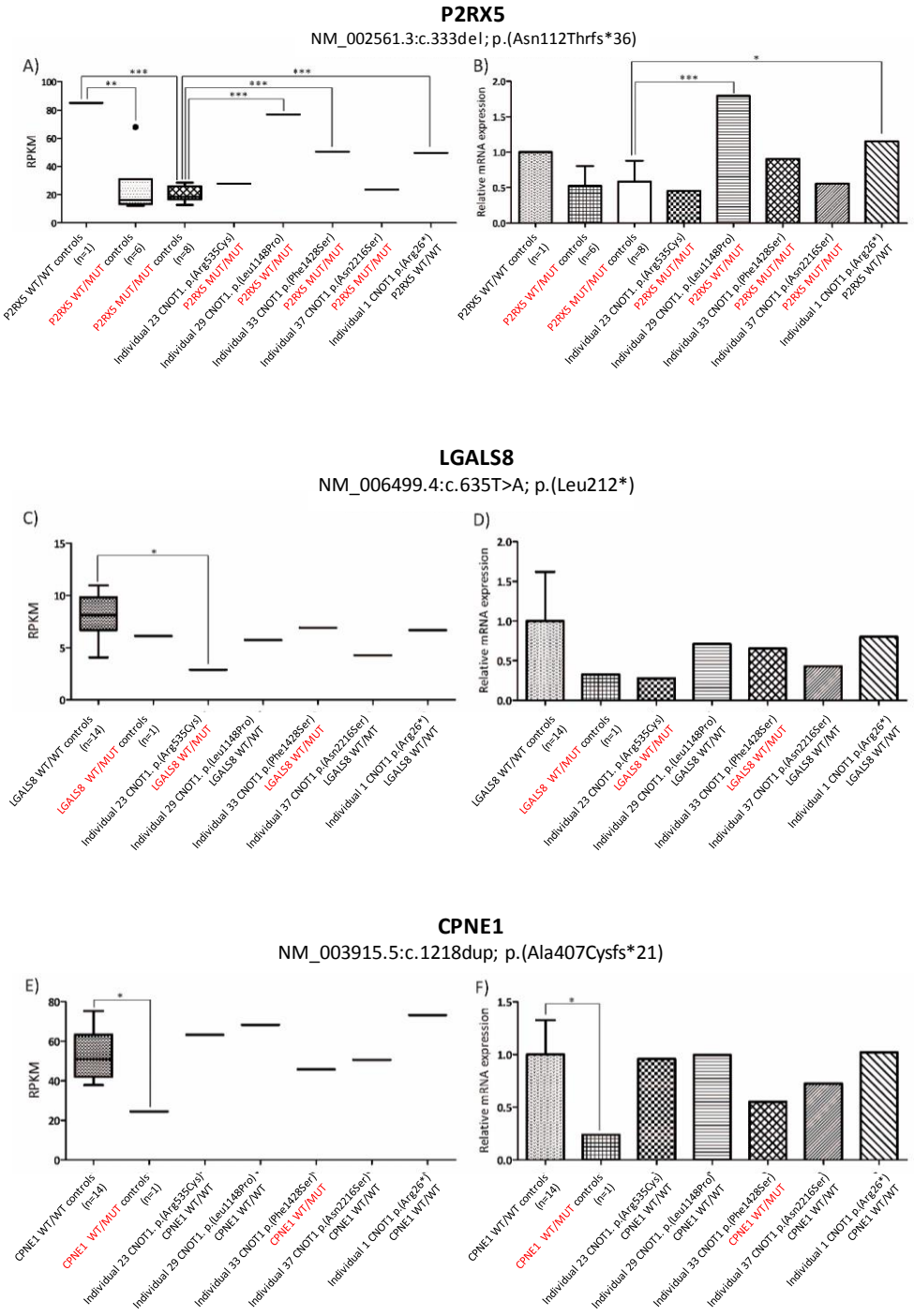
Evaluation of CNOT1 RNA (**A,B**, **C**) and protein (**D**) levels in patient-derived EBV cell lines of individuals with CNOT1 de novo variants and controls. In (**A**) normalized QuantSeq 3'mRNA counts are visualized for controls (n=15) and Individuals with *de novo* CNOT1 variants (n=5). Comparison between Individuals with *CNOT1* variants and controls does not show a statistical difference in CNOT1 expression. When individual controls are compared such difference can be observed (***= p<0.001). (**B**) Validation of QuantSeq results by conventional qPCR, indicating that there is no difference in expression between individuals with a *de novo* variant in CNOT1 and controls. (**C**) More detailed allele-specific qPCR shows that when only the WT allele is targeted by qPCR (tailor-made per individual; Table S6), only half the expression is observed in the patient-derived cell lines. These results indicating that the total amount of expression as observed in (**B**) is obtained by equal expression of both the WT and mutant allele. (**D**) Analysis of CNOT1 protein levels in these individuals again do not show significance difference when compared to controls. Vinculin is used as loading control.

Supplemental Figure 9: Filter strategy to identify genes harbouring PTCs expected to be targeted by NMD in controls and individuals with *de novo CNOT1* variants



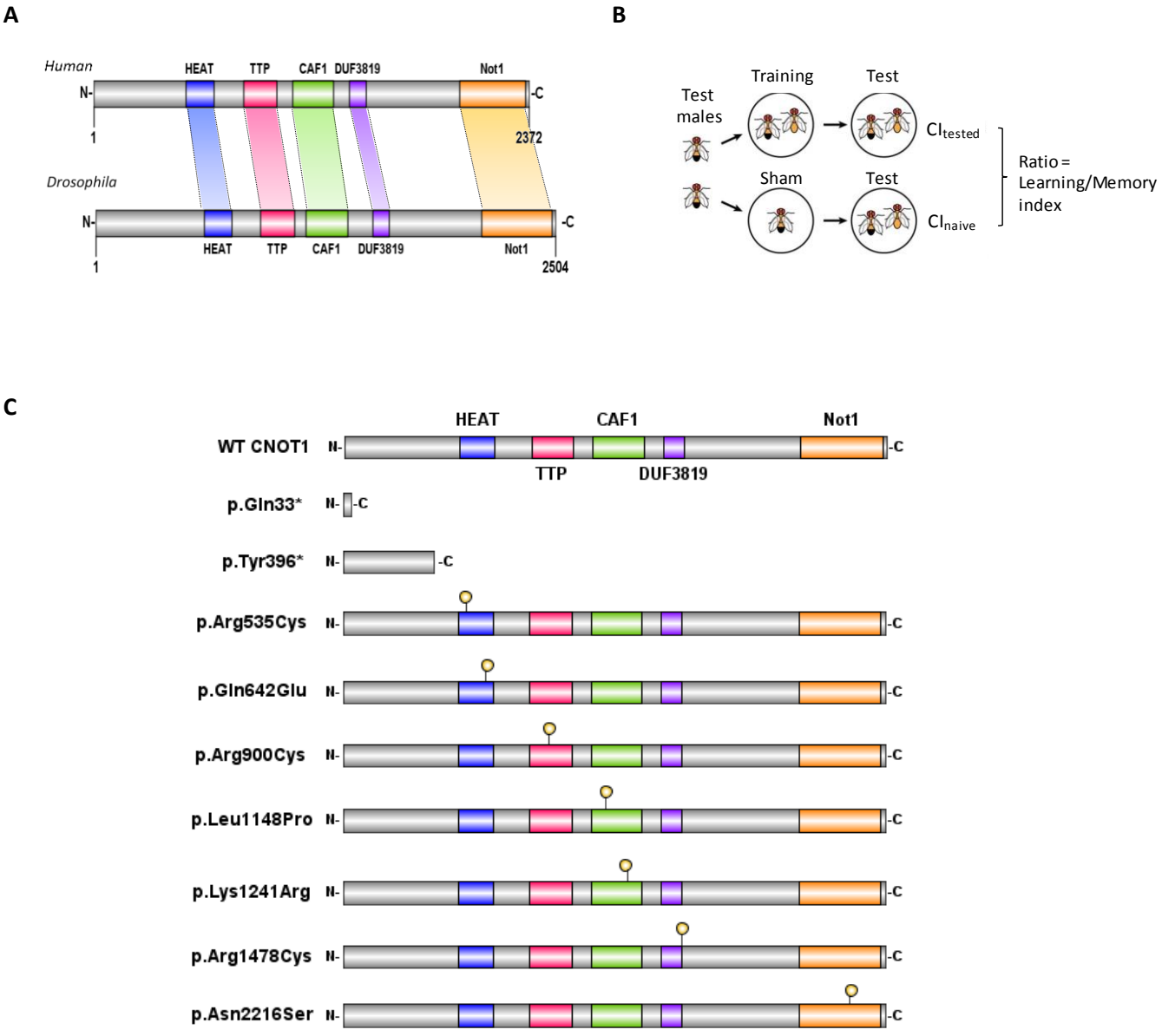
Selection criteria for filtering of exome sequencing data of individuals with *de novo CNOT1* variants (Individuals 1, 22, 28, 32 and 36) to evaluate NMD in QuantSeq 3'mRNA seq data. Overall QC thresholds used for filtering WES data: read counts for heterozygous and homozygous were set at 6x and 4x coverage respectively, with homozygous variants showing >90% of variant reads and heterozygous variants between 20-80%. Only "stop gain", "canonical splice sites" and "frameshift" were selected if they had a frequency higher than 10% of the alleles in a population (gnomAD frequency).

Supplemental Figure 10: Expression of genes of interest for which NMD is expected



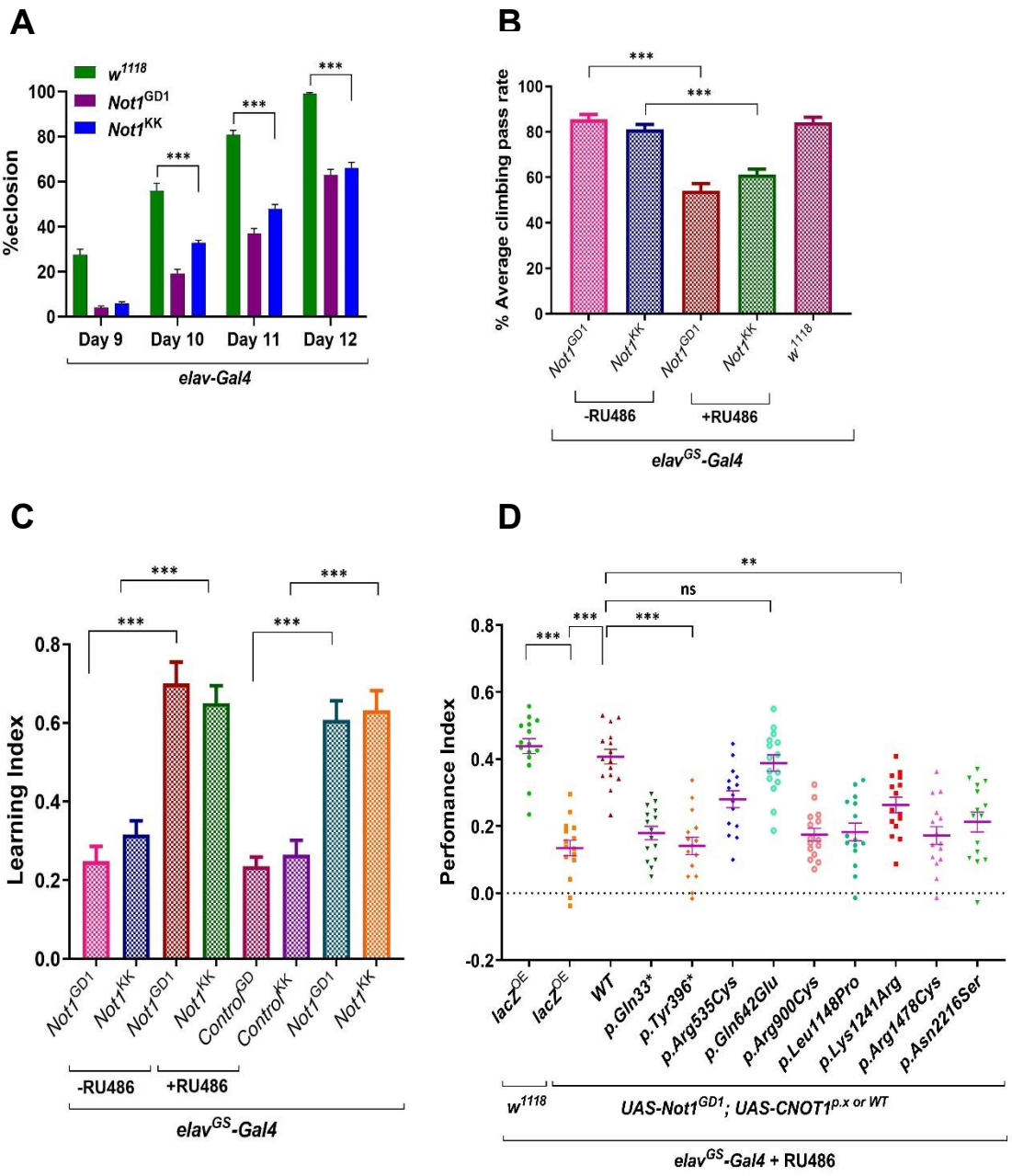
The genotype for the nonsense of frameshift variant in the gene of interest is indicated below each plot (in red, individuals with heterozygous or homozygous premature termination codons). For each gene, left panel shows the RPKM values by QuantSeq, with on the right panel, the validations thereof as relative mRNA expression obtained by qPCR. Significant results of two-tailed unpaired t-tests comparing control groups and two-tailed z-test comparing single samples with the control groups are indicated with stars ($p \leq 0.05$=*, $p \leq 0.01$=**, $p \leq 0.001$=***). Whereas subtle differences in NMD between controls and individuals with CNOT1 variants are observed from RNA seq data, none of them withstand validation by qPCR and statistical analysis. Hence, we concluded that individuals with de novo CNOT1 variants degrade transcripts with protein truncating variants in similar ways as controls, and that the endonucleolytic NMD pathway is not affected.

Supplemental Figure 11: Conservation of CNOT1 in Drosophila and schematic representation of courtship assay



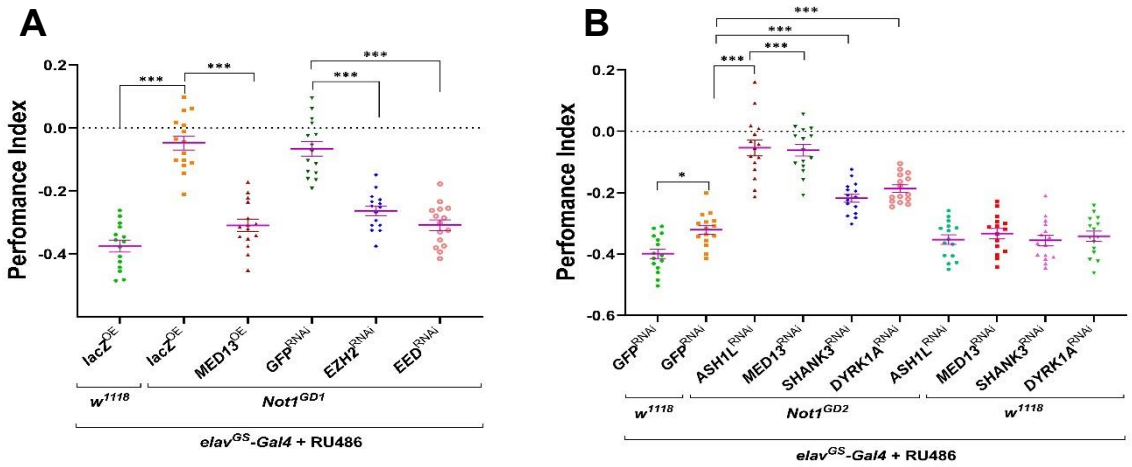
Schematic representation (**A**) of homology between human CNOT1 and its drosophila homolog Not1, and the principle of the courtship assay (**B**). In essence, if male flies have normal learning and memory, the learning and memory index, calculated by CI_{tested}/CI_{naive}, is low. Impaired learning and memory is shown by an index closer to 1.0. (**C**) Overview of CNOT1 constructs generated for the Drosophila functional follow-up, each harbouring a variant observed in an Individual with neurodevelopment disorders as observed in this study.

Supplemental Figure 12: Loss-of-Not1 function in Drosophila neurons results in multiple deficits.



(**A**) Persistent knockdown of *Not1* in all neurons under the control of *elav*-Gal4 induces a significant developmental delay. (**B**) Adult-only knockdown of CNOT1 for 3 weeks lead to motor dysfunction as assessed with the climbing assay described in Methods. (**C,D**) elav^{GS}-Gal4-induced knockdown of *Not1* in adults (**C**) or larval (**D**) stages results in learning and memory deficits. One–way ANOVA, **p<0.01, ***p<0.0001

Supplemental Figure 13: Genetic interactions between Not1 and known ID/ASD genes and rescue of Not1-induced neurodevelopmental defects at larval stages



(**A**) Overexpression of *MED13* or knockdown of PRC2 components *E(z)* and *esc* ameliorates *Not1* knockdown-induced memory deficits in larvae. (**B**) Neuronal RNAi knockdown of known ID/ASD genes (*ash1, MED13, shank3* and *Dyrk1a*) exacerbates *Not1* knockdown-induced memory defects in larvae. One–way ANOVA, *p<0.05, **p<0.01, ***p<0.0001

Supplemental Tables

Supplemental Table 1

Individual	Gender	Genomic annotation on Chr16(GRCh37):	c.DNA (NM_016284.4)	Protein	Type¥	Inheritance
Nonsense, Frameshift and Splice site variants						
Individual 1	F	g.58633166G>A	c.76C>T	p.(Arg26*)	Nonsense	*De novo*
Individual 2	M	g.58633166G>A	c.76C>T	p.(Arg26*)	Nonsense	Maternal mosaicism
Individual 3	F	g.58633166G>A	c.76C>T	p.(Arg26*)	Nonsense	*De novo*
Individual 4	F	g.58633166G>A	c.76C>T	p.(Arg26*)	Nonsense	*De novo*
Individual 5	M	g.58633145G>A	c.97C>T	p.(Gln33*)	Nonsense	*De novo*
Individual 6 *(daughter of 7)*	F	g.58633138A>G	c.102+2T>C	p.(?)	Splice site	Maternal
Individual 7 *(mother of 6)*	F	g.58633138A>G	c.102+2T>C	p.(?)	Splice site	Unknown
Individual 8	F	g.58622702C>A	c.210+1G>T	p.(?)	Splice site	*De novo*
Individual 9	M	g.58620475_58620478delTCTA	c.608_611delTAGA	p.(Ile203Thrfs*32)	Frameshift	*De novo*
Individual 10	F	g.58615276A>T	c.1188T>A	p.(Tyr396*)	Nonsense	Unknown: adopted
Individual 11	F	g.58583781_58583782insTTTAGCTTACCTTA	c.3363_3364insTAAGGTAAGCTAAA	p.(Val1122*)	Frameshift	*De novo*
Individual 12	F	g.58581153_58581159delTTGTCCT	c.3681_3687del	p.(Lys1227Asnfs*7)	Frameshift	Unknown: mother NA, not paternal
Individual 13	M	g.58581089C>A	c.3750+1G>T	p.(?)	Splice site	*De novo*
Individual 14	F	g.58581085C>T	c.3750+5G>A	p.(?)	Splice site	*De novo*
Individual 15	F	g.58577805T>G	c.4138-2A>C	p.(?)	Splice site	*De novo*
Individual 16	M	g.58575405C>G	c.4800G>C	(p.Lys1600Asn) / p.(?)	Missense / Splice site	*De novo*
Individual 17	M	g.58562529dup	c.6303dup	p.(Leu2102Serfs*4)	Frameshift	*De novo*

Individual 18 *(son of 19)*	M	g.58559939_58559967dup	c.6529_6557dup	p.(Arg2187Lysfs*63)	Frameshift	Maternal
Individual 19 *(mother of 18)*	F	g.58559939_58559967dup	c.6529_6557dup	p.(Arg2187Lysfs*63)	Frameshift	Unknown
Missense variants						
Individual 20*	M	g.58610468G>A	c.1603C>T	p.(Arg535Cys)	Missense	*De novo*
Individual 21*	F	g.58610468G>A	c.1603C>T	p.(Arg535Cys)	Missense	*De novo*
Individual 22#	M	g.58610468G>A	c.1603C>T	p.(Arg535Cys)	Missense	*De novo*
Individual 23#	F	g.58610468G>A	c.1603C>T	p.(Arg535Cys)	Missense	Unknown: mother NA, not paternal
Individual 24#	F	g.58610468G>A	c.1603C>T	p.(Arg535Cys)	Missense	*De novo*
Individual 25	M	g.58608568G>C	c.1924C>G	p.(Gln642Glu)	Missense	*De novo*
Individual 26	M	g.58589357C>T	c.2689G>A	p.(Glu897Lys)	Missense	*De novo*
Individual 27	M	g.58589348G>A	c.2698C>T	p.(Arg900Cys)	Missense	*De novo*
Individual 28	M	g.58585581G>A g.58585113C>G	c.3113C>T c.3265G>C	p.(Thr1038Ile) p.(Val1089Leu)	Missense Missense	*De novo* *De novo*
Individual 29	F	g.58583702A>G	c.3443T>C	p.(Leu1148Pro)	Missense	*De novo*
Individual 30	M	g.58581546T>C	c.3563A>G	p.(Asp1188Gly)	Missense	*De novo*
Individual 31	F	g.58581118T>C	c.3722A>G	p.(Lys1241Arg)	Missense	*De novo*
Individual 32	F	g.58577690T>C	c.4255A>G	p.(Thr1419Ala)	Missense	*De novo*
Individual 33	M	g.58577662A>G	c.4283T>C	p.(Phe1428Ser)	Missense	*De novo*
Individual 34	M	g.58577513G>A	c.4432C>T	p.(Arg1478Cys)	Missense	*De novo*
Individual 35	F	g.58576425T>A	c.4482A>T	p.(Gln1494His)	Missense	*De novo*
Individual 36	M	g.58575491A>C	c.4714T>G	p.(Tyr1572Asp)	Missense	*De novo*
Individual 37	M	g.58559220T>C	c.6647A>G	p.(Asn2216Ser)	Missense	*De novo*
(Partial) gene deletions						
Individual 38	M	46,XY,del(16)(q21) arr 16q21(58589134_58594329)x1	deletions of exon 17-21	p.(?)	Intragenic deletion	*De novo*
Individual 39	M	46,XY,del(16)(q12.2q21) arr 16q12.2q21(55,741,501-62,145,478)x1	-	-	Whole gene deletion	*De novo*

**Individual previously described by Kruszka et al. 2019; #: Individual previously described by De Franko et al. 2019.*

Supplemental Table 2: Clinical details for each individual with a (*de novo*) *CNOT1* variant

Excel file describing the clinical details for each individual included in this study.

Supplemental Table 3: Effect predictions of de novo missense variants in *CNOT1*

Unique missense variants	Domain	PDB structure	Location of missense	Predicted effect
p.(Arg535Cys)	HEAT	-		Loss of positive charge, and more hydrophobic than wildtype residue.
p.(Gln642Glu)	HEAT	-		Introduction of a negative charge.
p.(Glu897Lys)	TTP	4J8S	Buried in the core of the protein.	Introduction of a positive charge instead of a negative one. Loss of hydrogen bond, and sidechain might clash with Arg in vicinity. Possible destabilization of ZFP36 binding groove which affects the interaction with this protein.
p.(Arg900Cys)	TTP	4J8S	Buried in the core of the protein.	Loss of positive charge, and more hydrophobic than wildtype residue. Loss of salt-bridge formation with Glutamic Acid at position 940, which might cause loss of stability close to the ZFP63 binding groove. Interaction with ZFP36 might be disturbed.
p.(Thr1038Ile)	-	-	-	*No 3D structure available for effect prediction*
p.(Val1089Leu)	CAF1	4CT4	Surface of the protein	Change to hydrophobic amino acid. May disturb interactions with DDX6 and CNOT6/CNOT6L/CNOT7/CNOT8.
p.(Leu1148Pro)	CAF1	4CT4	Buried in the core of the protein.	The proline is expected to disturb the alpha-helix which may have severe effects on the structure of the protein. Variant will affect the interaction with DDX6 and CNOT6/CNOT6L/CNOT7/CNOT8.
p.(Asp1188Gly)	CAF1	4CT4	Semi-buried in core	Change of a relatively large to smaller residue ca affect the general stability of the conformation. No specific interactions of function found for this residue.
p.(Lys1241Arg)	CAF1	4CT4	Surface of the protein.	Loss of hydrogen bonds with Glutamic Acid at position 1244 as well as the salt bridges with Glutamic Acids 1244 and 1282. Interaction with CNOT6/CNOT6L/CNOT7/CNOT8 might be disturbed.
p.(Thr1419Ala)	DUF3819	4CRU	Surface of the protein.	Alanine is a bit smaller and hydrophobic. Maybe interaction with CNOT9 might be disturbed.
p.(Phe1428Ser)	DUF3819	4CRU	Buried in the core of the protein.	Less hydrophobic than wild type residue. Phe1428 reported by PISA-assembly be involved in protein interaction. Interaction with CNOT9 might be disturbed.
p.(Arg1478Cys)	DUF3819	4CRU	Surface of the protein.	Loss of positive charge, and more hydrophobic than wildtype residue. Structural information is unclear about function of this residue.
p.(Gln1494His)	DUF3819	4CRU	Surface of the protein	Size difference, distant from CNOT9 interaction site, no severe effect on structure expected
p.(Tyr1572Asp)	-	4CRU	Buried in the core of the protein.	Important for hydrophobic interactions in core. Loss of stability, close to CNOT9 interaction surface.
p.(Asn2216Ser)	Not1	4C0D	Buried in the core of the protein.	The variant will cause loss of hydrogen bonds (with Aspartic Acid at position 2219) in the core of the protein and as a result disturb correct folding.

Supplemental Table 4: *CNOT1* variants created by site directed mutagenesis

Protein annotation based on NM_001265612.1	Primer Orientation	Primer sequence (5'->3')
p.(Gln33*)	FW	GCAGGAAATATAGCATATTGTGAATC
	RV	TGGCTGGCTCGGTAATTT
p.(Tyr396*)	FW	ACCTCATATAAAGACCTTGGAAAC
	RV	CTACTGGGAACACTTCCATAC
p.(Arg535Cys)	FW	TCCCTCAATTTGCCAACTTATCATGC
	RV	GACTGTCCCTGCCCATGC
p.(Gln642Glu)	FW	CAAAAGTGCTGAACTTCCTCC
	RV	GGCTGGTCTTTTTCTGGG
p.(Arg900Cys)	FW	TGAAGAATATTGTTTTTTTCCCC
	RV	AACAAGTTCCTTAGCATAC
p.(Leu1148Pro)	FW	TTTCATAGCCCGTATTCAAACTTC
	RV	GTTTGGCTCAATACTGAC
p.(Lys1241Arg)	FW	TTTGTTGCCAGAGTCTTAGAATC
	RV	GGGCACTACATAGAGCAA
p.(Phe1428Ser)	FW	AGGAAGGATTCTGCCCTGGAT
	RV	GACTATTTGCTCACAAGTAGTCATG
p.(Arg1478Cys)	FW	CTCAGCCCTTTGTACTGCTTC
	RV	GCAAAACTGTTTTTTAAGTTGG
p.(Asn2216Ser)	FW	CAGCTCATCAGTGCACTGGTG
	RV	GAGGTTGTAGCGATTCCC
p.(Leu2102Serfs*4)	FW	TTTCTTTTGCATGATTTCCCAG
	RV	ACCAGCAGCACTCTTAAAG

Supplemental Table 5: Experimental set-up co-localization studies

CNOT1 mut	CNOT2	CNOT4	CNOT7	CNOT8	CNOT9	CNOT11
p.(Gln33*)						X
p.(Arg535Cys)		X				
p.(Gln642Glu)		X				
p.(Arg900Cys)		X				
p.(Leu1148Pro)			X	X		
p.(Lys1241Arg)			X	X		
p.(Arg1478Cys)					X	
p.(Asn2216Ser)	X					

Supplemental Table 6: Allele-specific primers to assess contribution of WT allele to total *CNOT1* expression

WT allele of	Sequence (5'-3')	(5'-3')
CNOT1 p.Arg26*	CAATTTAACCAAGAAAAATGACC	GAAATCCACATGCGAAAATAGGC
CNOT1 p.Arg535Cys	GGACAGTCTCCCTCACTTC	AAGTCCTGGGCCACATCAAG
CNOT1 p.Leu1148Pro	TGAGCCAAACTTTCATATCCT	CAGAGGTCAGGAGCACTTTAATG
CNOT1 p.Phe1428Ser	GCAAATAGTCAGGAAGGCTTT	AGCTGTCAAGTTACGCATCATG
CNOT1 p.Asn2216Ser	TACAACCTCCAGCTCAGCAA	TCCATGTGTGCTGAGTGAGT

For CNOT1, a general primer pair was used (Table S7), as well as primer pairs that are designed to be specific for the WT allele of each individual. For WT-allele-specific primers, the mutated nucleotide is marked red and the introduced mismatch is marked blue. The genes GUSB, PPIB and CLK2 were used for normalization (primer sequences in Table S7).

Supplemental Table 7: qPCR primers used for validation of RNAseq data and allele-specific expression

Gene	Sequence (5'-3')	Sequence (5'-3')
P2RX5	GTCTGTGCTGAGAATGAAGGC	GTCTTCACTCCGTTTCCAGC
LGALS8	GCATGTTCCTAGTGACGCAG	ACGAGGATTGAAATGAAAGGCC
CPNE1	ACTGACTCTCCCCTTGATGC	GGTTTCTGGCCTCTACCTCC
CNOT1 total expression	GCCCAAAAGTGCTCAACTTC	TGGCTACCATGGTGAGGATAG
RUVBL2	ATTGATCGACCAGCAACAGG	GACTCAATCATCTTGGTGCCC
PABPC1	CGCGTATGTGAACTTCCAGC	ACTGGCTTGCCCTTTATAACATC
PPP2CA	TGGTGGATGGGCAGATCTTC	TGGGGAACTTCTTGTAGGCG
GUSB*	AGAGTGGTGCTGAGGATTGG	CCCTCATGCTCTAGCGTGTC
PPIB*	CGGAAAGACTGTTCCAAAAAC	GATTACACGATGGAATTTGCTG
CLK2*	CGGCGAGAGGACAGCTAC	AGTATCGCCGGTCATACACC

** used as reference/control*

Supplemental Table 8: Targets identified for NMD assessment by integrating exome sequencing and QuantSeq data of Individuals with a *de novo CNOT1* variants

Gene	gDNA (GRCh37)	cDNA	Protein	Freq. (%)	Individual (allelic composition for variant)
P2RX5	chr17:g.3594277delG	NM_002561.4: c.333delC	p.(Asn112Thrfs*36)	58.51	1(**MUT**/**MUT**) 23 (**MUT**/**MUT**) 29 (**WT**/**MUT**) 33 (**MUT**/**MUT**)
LGALS8	chr1:g.236706300T>A	NM_006499.4: c.635T>A	p.(Leu212*)	8.041	23 (**WT**/**MUT**) 37 (**WT**/**MUT**)
CPNE1	chr20:g.34215234insA	NM_003915.5: c.1218dupT	p.(Ala407Cysfs*21)	9.653	33 (**WT**/**MUT**)

The positions of the variants in the genomic DNA (GRCh37), effect at transcript level (cDNA), and protein level are provided. The last column indicates the individuals who carry the nonsense or frameshift variant. In parenthesis is indicated whether the variant is identified as heterozygous (WT/MUT) or homozygous (MUT/MUT) variant. Freq. = frequency that refers to the overall frequency in GnoMAD.

Supplemental Material and Methods

Generation of expression vectors of the CCR4-NOT complex. To study the functional consequences of *de novo* variants in *CNOT1*, the coding region of wild-type (WT) cDNA (NM_001265612) was cloned into template vector pT7-EGFP-C1-HsNot1 (a kind gift from Elisa Izaurralde, Addgene plasmid # 37370)[3], after which the cDNA was subsequently cloned into destination vector pDONR201 using the gateway system according to the manufacturer's instructions (ThermoFisher Scientific, Landsmeer, the Netherlands), and transformed into DH5α cells using standard protocols. Plasmid DNA was isolated using a NucleoSpin® Plasmid QuickPure kit (Macherey-Nagel, Dueren, Germany), and, after digestion using HINDIII-HF (20 U/µl, New England Biolabs, Bioké, Leiden, the Netherlands), sequenced to exclude the presence of variants. Then, pDONR201-CNOT1 was cloned into two expression vectors: pPalmMyr/DEST(CFP), encoding PalmMyr-CFP-CNOT1, which contains a green fluorescently-labeled (CFP) PalmMyr cell membrane linker; and pDEST-733 (mRFP, ThermoFisher Scientific, Landsmeer, the Netherlands), encoding red fluorescent-labeled (RFP)-CNOT1, in which CNOT1 contains a RFP tag allowing visualization in the cytosol and/or nucleus. Using similar protocols, also the CCR4-NOT complex subunits CNOT2, CNOT4, CNOT7, CNOT8, CNOT9 and CNOT11 were cloned into expression vector pDEST-733 (mRFP; ThermoFisher Scientific, Landsmeer, the Netherlands).

Detailed information of CCR4-NOT complex members used in this study:

Subunit	Catalog number	Accession number (EST/RefSeq)
CNOT2	GC-Z1136	ENST00000418359/NM_001199302
CNOT4	GC-A1384	ENST00000423368/NM_001190847
CNOT7	GC-T8806	ENST00000361272/NM_001322090
CNOT8	GC-A6058	ENST00000519404/NM_001301074
CNOT9	GC-U0151	ENST00000273064/NM_005444
CNOT11	GC-U0969	ENST00000289382/NM_017546

Site directed mutagenesis. In total, 11 CNOT1 variants were generated using site directed mutagenesis in pDONR201-CNOT1 (Table S3). In short, primers were designed and ordered (IDT) for each variant using NEBaseChanger, after which variants were introduced using the Q5 mutagenesis protocol of the

manufacturer (New England Biolabs, via Bioké Leiden the Netherlands). PCR products were sequenced to ensure that the variants were introduced correctly, and subsequently transformed into DH5α cells. Plasmids were isolated routine Miniprep kit (NucleoSpin® Plasmid QuickPure (Macherey-Nagel, Dueren, Germany). After digestion using PstI-HF (20 U/µl, New England Biolabs, via Bioké Leiden the Netherlands), the mutated CNOT1 constructs were cloned into expression vector pPalmMyr-CFP, thus encoding a variant-specific PalmMyr-CFP-CNOT1 expression.

Co-localization studies of CNOT1 and CCR4-NOT complex partners in COS-1 cells. COS-1 cells were grown to ~90% confluency on a glass coverslip in a 12-well plate, after which they were transfected with plasmid isolates according to the manufacturer's instructions (Lipofectamine® 3000; Cat# L3000001; Life Technologies, Landsmeer, the Netherlands) with minor modifications. In brief, 3µl Lipofectamine® 3000 was used for co-transfection of 500ng of the CFP-PalmMyr-CNOT1/mut plasmid DNA and 300ng of the RFP-CNOT subunits CNOT2, CNOT4, CNOT7, CNOT8, CNOT9 or CNOT11 (**Table S4**), respectively, while for single transfection 500ng of PalmMyr-CFP-CNOT1/mut or 300ng of the RFP-CNOT subunits CNOT2, CNOT4, CNOT7, CNOT8, CNOT9 or CNOT11, and 1 µl lipofectamine® 3000 was used. After ca. 6 hours the media was replaced by the normal DMEM media. The cells were washed two times in 1X PBS 36-48 hours after transfection and fixed in 2% paraformaldehyde (PFA, stock 4% PFA diluted in PBS) at room temperature for 15 min. Cells were stained with 4',6-diamidino-2-phenylindole (DAPI; Vector laboratories, via Brunschwig Chemie, Amsterdam, the Netherlands) while placed on a microscope slide, and localization of CNOT1 WT/mut and its subunits was analyzed using a Zeiss Axio Imager Z1/Z2 fluorescence microscope equipped with a 63x objective lens. Image processing was performed using Zen software (blue edition, Zeiss).

Apoptosis assay using patient-derived cell lines. For three Individuals (p.(Arg26*), p.(Arg535Cys), p.(Leu1148Pro)) and five control lines, an apoptosis assay was performed using cultured Epstein-Barr-

virus transformed lymphoblastoid cell lines (EBV-LCLs). Hereto, cells were incubated for 20-22 hours in medium (RPMI + 15%FSC + HEPES (1:50) and 1%Pen/Strep) containing 1µM Bortezomib to induce apoptosis. Next, cells from 2 ml culture were harvested and counted, to be re-suspended at a concentration of $2.5x10^6$ cells/ml using ice-cold PBS. Of these, 500,000 cells (200 µl) were spun down and re-suspended in 1xFACS buffer (PBS with 0.5% BSA) containing Annexin V-FITC (apoptosis marker) and 7-AAD (viable cell marker). FACS analysis, discriminating viable cells from those in apoptosis, was performed in duplicate using routine procedures.

qPCR Experiments for validation and allele-specific expression of individuals with a *de novo CNOT1* variant. For the quantitative real-time PCR (qPCR) experiments, RNA samples were isolated from EBV cell lines and converted to cDNA using routine procedures. qPCR primers at cDNA level were designed with Primer3web v4.1.0 (Table S7)[4]. Allele-specific primers were designed for the wildtype (WT) alleles of the patients' CNOT1 variants, according to Gaudet et al. (2009)[5]. *GUSB*, *PPIB*, and *CLK2* were taken along in the qPCR experiments as reference genes[6]. Standard curves were made for all primers with a cDNA control sample in a dilution series (20x, 80x, 320x, 1280x, 5120x dilutions in MilliQ). For all qPCR experiments, samples were tested in duplicate and a blank was taken along for each primer pair.

QuantSeq 3'mRNA sequencing. A selection of five individuals carrying *de novo CNOT1* mutations (Table S6) and 15 healthy controls were subjected to quantitative RNA sequencing (QuantSeq 3'mRNA-Seq Library Prep kit-FWD, Lexogen) in accordance with protocol. Prior to library generation, RNA from individuals with *CNOT1* variants and controls was isolated from EBV lymphoblastoid cell lines with either the NucleoSpin RNA kit (Macherey-Nagel) or the RNeasy mini kit (Qiagen) including on-column DNAse digestion for all samples (RNAse-Free DNAse kit, Qiagen), according to the manufacturer's protocol. Respective quantification of purified RNA was determined using Qubit RNA HS assay (Thermo Fisher Scientific). RNA input was normalized to 1000 ng for all samples. A mock qPCR on a 1:10 aliquot

of ds cDNA libraries indicated 17 cycles to be optimal for endpoint PCR. Quantification and quality control, including identification of the average fragment size, of the generated libraries was performed using the dsDNA HS assay (Qubit, Thermo Fisher Scientific) and HS-D1000 ScreenTape (2200 Tapestation, Agilent), respectively. Libraries were pooled equimolar to 100 fmol and diluted to a final concentration of 4 nM before sequencing on a NextSeq 500 instrument (75-cycle high output kit, Illumina). For RNAseq analysis, first quality was assessed using FastQC (v0.11.5, Babraham Bioinformatics), after which the raw data was subjected to TrimGalore! (v.0.4.4_dev, Babraham Bioinformatics) and Cutadapt[7] (v1.18) for removal of adapter sequences and poly(A) tails. Subsequent mapping of filtered and trimmed reads to the human reference genome (hg38) was performed using STAR aligner[8] (v2.6.0a). Lastly, gene level count data was generated using the HTSeq-count tool[9] (v0.11.0).

Evaluation of NMD efficiency using transcriptomics in patient-derived cell lines. Genes of interest for studying NMD of aberrant RNA transcripts contain a loss-of-function (LoF) in at least one individual with a de *novo CNOT1* variant. Selection was based on the following criteria (**Figure S8**). Homozygous and heterozygous loss-of-function variants were selected by filtering the exome sequencing (ES) data of five individuals with *de novo* CNOT1 variants. Next, variants were evaluated to fulfil the canonical rules for NMD: variants located >50-55 nucleotides upstream of the last exon-exon boundary or that were located in the last exon were excluded. Genes with multiple frameshift variants that were in-frame with each other and genes with additional variants that in combination with the selected stop variant led to a missense codon were excluded. In addition, only variants with a global SNP frequency <95% in the Genome Aggregation Database (gnomAD) were included to overcome problems with MAF wrongly represented as the common allele.[10] All genes without expression in the RNA sequencing data set were excluded from the analysis. The threshold for expression (QuantSeq) was set to a total read count of ≥10 in all twenty samples together. Lastly, the Genotype-Tissue Expression (GTEx) database

was consulted regarding the expression of the genes of interest with and without the LoF variants in EBV-LCLs[11]. Those genes for which the expression quantitative trait locus (eQTL) of the LoF variant showed a decreased expression of the variant allele in EBV-LCLs in the GTEx database were included.

***Drosophila* Genetics.** Fruit fly husbandry was done at 25℃ unless otherwise mentioned. Fly RNAi lines that include *Not1* (GD1#v12571, GD2#v41680, KK#v106587), *ash1* (GD#v28982, KK#v108832), *MED13* (SH#v330726), *shank3* (GD#44830, KK#v103592), *Dyrk1a* (GD#v28628, KK#v107066) were obtained from Vienna stock center (Vienna Biocenter Core Facilities GmbH, Vienna, Austria). The pan-neuronal Gal4 driver lines, *elav-Gal4* (BL-458), *elav^{GS}-Gal4* (BL-43642), *OK107-Gal4* (BL-854) were obtained from Bloomington stock center (Dept Biology, Indiana University, Indianapolis, United States). Additional RNAi lines for *MED13* (BL-34630), *Dyrk1a* (BL-35222) and UAS-*MED13* (BL-63800) were obtained from Bloomington stock center. Human *CNOT1* cDNA variants were cloned into a *pUAST*-attB vector and transgenes were produced (BestGene, Chino Hills, California, Unites States) using a site-specific integration method to avoid position effects on gene expression. Flies were raised on a standard molasses-corn meal food. For the courtship memory assays, adult flies of desired genotypes were treated with 100µM RU486 (Mifepristone) mixed in fly food for 1-week after eclosion in order to conditionally induce *elav^{GS}-Gal4*-mediated transgene expression before testing. For larval olfactory assays, larvae were allowed to hatch on fly food containing 100µM RU486 and stay there for 5-days before testing with the larval memory assays during late third instar larval stage.

Courtship suppression assay. Flies were raised at 25°C before collecting the F1 males and females separately. For courtship assays, naïve male (3-4 day old) and a mated female (3-4-day old virgin female mated one day before training) were transferred into the courtship chamber and the courtship was observed for 60mins. Courtship Index (CI) is defined as fraction of time a male spent in courtship activity during 10 min observation period. Initial 10mins ($CI_{initial}$) and final 10mins (CI_{final}) were carefully

monitored for the time the test male spends in courting the mated female. After 2-hours, the trained male was subsequently transferred to another courtship chamber with an immobilized virgin tester female and the courtship was monitored immediately for 10mins (CI_{test}). As a sham-trained control, the tester male was isolated for 60mins in the courtship chamber before testing towards the tester female (CI_{sham}). Courtship index, learning index and memory indexes were calculated as previously reported[12,13]:

1. Learning index = $CI_{final}/CI_{initial}$
2. Memory Index = CI_{test}/mean CI_{sham}

Larval olfactory-mediated memory tests. Larval memory tests were performed as per established protocols[14,15]. Briefly, assay plates were filled with a thin layer of 2.5% agarose (Sigma Aldrich Cat. No.#A5093, St. Louis, MO, Unites States) with or without a gustatory reinforcer. We used 2M sodium chloride (Sigma Aldrich Cat. No.#S7653, St. Louis, MO, Unites States), 2M fructose (Sigma Aldrich Cat. No.#47740, St. Louis, MO, Unites States) as reinforcers. Olfactory stimuli provided using 10µl amyl acetate (AM; Acros Organics, via Thermo Fisher, New Jersey, United States, Cat No#AC149182500 diluted 1:250 in paraffin oil, Sigma Aldrich Cat. No.#18512, St. Louis, United States) and benzaldehyde (BA; Sigma Aldrich Cat. No.#418099, St. Louis, United States). Odorants were placed in specific containers with perforated lids to minimize evaporation. First group of 30 larvae were placed on petri dish containing sodium chloride as a negative reinforcer or fructose as a positive reinforcer and AM as an olfactory stimulator. After 5 min, these larvae were transferred to pure agarose petri dish containing BA. A second group of larvae received the reciprocal training. After three cycles of training, larvae were transferred to another agarose plate before further testing. Memory was tested by placing larvae onto test plates that contain positive or negative gustatory reinforcers with AM and BA placed on opposite sides. After 5 min, individuals on both sides were counted and preference indexes (PREF) and

performance indexes (PI) were calculated as mentioned below. Negative PI numbers indicate aversive associative learning, whereas positive PI numbers reflect appetitive associative learning:

1. (1a) $PREF_{AM+/BA}$ = (# AM—# BA) / # TOTAL
2. (1b) $PREF_{AM/BA+}$ = (# AM—# BA) / # TOTAL
3. (2) PI = ($PREF_{AM+/BA}$–$PREF_{AM/BA+}$) / 2test

Climbing Assay. Motor function in *Drosophila* was quantified as shown previously[16,17]. Briefly, flies were collected and aged at 25℃ for three weeks before testing. Using a negative geotaxis assay, where flies (10 per vial) were challenged to climb a vertical height of 10cm in an empty fly food vial. This trial is repeated 10 times per vial with a gap of one minute between trials and approximately 150 flies were analyzed for each genotype. The percentage of flies that crossed the 10cm mark was determined and plotted as % climbing pass rate.

References to the Supplemental Information

1. Wiel, L. *et al.* MetaDome: Pathogenicity analysis of genetic variants through aggregation of homologous human protein domains. *Hum Mutat* **40**, 1030-1038 (2019).
2. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
3. Braun, J.E., Huntzinger, E., Fauser, M. & Izaurralde, E. GW182 proteins directly recruit cytoplasmic deadenylase complexes to miRNA targets. *Mol Cell* **44**, 120-33 (2011).
4. Primer3web. *http://primer3.ut.ee/ (Accessed: 13th June 2019)* (2019).
5. Gaudet, M., Fara, A.-G., Beritognolo, I. & Sabatti, M. . Allele-specific PCR in SNP genotyping. in Single Nucleotide Polymorphisms. . *Methods in Molecular Biology (ed. Komar, A.)* (2009).
6. de Brouwer, A.P., van Bokhoven, H. & Kremer, H. Comparison of 12 reference genes for normalization of gene expression levels in Epstein-Barr virus-transformed lymphoblastoid cell lines and fibroblasts. *Mol Diagn Ther* **10**, 197-204 (2006).
7. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *2011* **17**, 3 (2011).
8. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
9. Anders, S., Pyl, P.T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166-9 (2015).
10. Karczewski. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. *BioRxiv* doi:10.1101/531210, 39 (2019).
11. GTEX Portal. *GTEX project* Available at: *https://gtexportal.org/home/.* (Accessed: 23rd April 2019).
12. Ejima, A. & Griffith, L.C. Measurement of Courtship Behavior in Drosophila melanogaster. *CSH Protoc* **2007**, pdb prot4847 (2007).
13. Ejima, A. & Griffith, L.C. Assay for courtship suppression in Drosophila. *Cold Spring Harb Protoc* **2011**, pdb prot5575 (2011).
14. Apostolopoulou, A.A., Widmann, A., Rohwedder, A., Pfitzenmaier, J.E. & Thum, A.S. Appetitive associative olfactory learning in Drosophila larvae. *J Vis Exp* (2013).
15. Widmann, A. *et al.* Genetic Dissection of Aversive Associative Olfactory Learning and Memory in Drosophila Larvae. *PLoS Genet* **12**, e1006378 (2016).
16. Collart, M.A. The Ccr4-Not complex is a key regulator of eukaryotic gene expression. *Wiley Interdiscip Rev RNA* **7**, 438-54 (2016).
17. Kamachi, Y. & Kondoh, H. Sox proteins: regulators of cell fate specification and differentiation. *Development* **140**, 4129-4144 (2013).