Biallelic mutations in the 3′ exonuclease *TOE1* cause pontocerebellar hypoplasia and uncover a role in snRNA processing

Rea M Lardelli[1,36], Ashleigh E Schaffer[1–3,36], Veerle R C Eggens[4,36], Maha S Zaki[5], Stephanie Grainger[6], Shashank Sathe[3], Eric L Van Nostrand[3], Zinayida Schlachetzki[2], Basak Rosti[2], Naiara Akizu[2], Eric Scott[2], Jennifer L Silhavy[2], Laura Dean Heckman[2], Rasim Ozgur Rosti[2], Esra Dikoglu[2], Anne Gregor[2], Alicia Guemez-Gamboa[2], Damir Musaev[2], Rohit Mande[2], Ari Widjaja[2], Tim L Shaw[1], Sebastian Markmiller[3], Isaac Marin-Valencia[2], Justin H Davies[7], Linda de Meirleir[8], Hulya Kayserili[9], Umut Altunoglu[10], Mary Louise Freckmann[11], Linda Warwick[12], David Chitayat[13,14], Susan Blaser[15], Ahmet Okay Çağlayan[16,17], Kaya Bilguvar[18], Huseyin Per[19], Christina Fagerberg[20], Henrik T Christesen[21], Maria Kibaek[21], Kimberly A Aldinger[22], David Manchester[23], Naomichi Matsumoto[24], Kazuhiro Muramatsu[25], Hirotomo Saitsu[24,26], Masaaki Shiina[27], Kazuhiro Ogata[27], Nicola Foulds[28], William B Dobyns[22], Neil C Chi[29], David Traver[6], Luigina Spaccini[30], Stefania Maria Bova[31], Stacey B Gabriel[32], Murat Gunel[17], Enza Maria Valente[33], Marie-Cecile Nassogne[34], Eric J Bennett[1], Gene W Yeo[3,35], Frank Baas[4], Jens Lykke-Andersen[1] & Joseph G Gleeson[1,2]

Deadenylases are best known for degrading the poly(A) tail during mRNA decay. The deadenylase family has expanded throughout evolution and, in mammals, consists of 12 Mg^{2+}-dependent 3′-end RNases with substrate specificity that is mostly unknown[1]. Pontocerebellar hypoplasia type 7 (PCH7) is a unique recessive syndrome characterized by neurodegeneration and ambiguous genitalia[2]. We studied 12 human families with PCH7, uncovering biallelic, loss-of-function mutations in *TOE1*, which encodes an unconventional deadenylase[3,4]. *toe1*-morphant zebrafish displayed midbrain and hindbrain degeneration, modeling PCH-like structural defects *in vivo*. Surprisingly, we found that TOE1 associated with small nuclear RNAs (snRNAs) incompletely processed spliceosomal. These pre-snRNAs contained 3′ genome-encoded tails often followed by post-transcriptionally added adenosines. Human cells with reduced levels of TOE1 accumulated 3′-end-extended pre-snRNAs, and the immunoisolated TOE1 complex was sufficient for 3′-end maturation of snRNAs. Our findings identify the cause of a neurodegenerative syndrome linked to snRNA maturation and uncover a key factor involved in the processing of snRNA 3′ ends.

The onset of pontocerebellar neurodegeneration occurs so early that it overlaps with neurodevelopment, and it is thus alternatively referred to as pontocerebellar hypoplasia (PCH)[5]. PCH7 (MIM 614969) is characterized by neurological deterioration, atrophy or hypoplasia of the pons and cerebellum, muscular hypotonia and breathing abnormalities, in combination with hypogonadism[2]. This combination of rare conditions suggests a unique syndromic association due to mutation of a single gene, but no locus or causative gene has been identified to date. We recruited 12 families meeting criteria for PCH7, including the index family on the basis of which the condition was defined (**Fig. 1a**), and we confirmed that the clinical features of these families matched those published for PCH7 (**Supplementary Table 1**). These features included reduced pons and cerebellum parenchyma (**Fig. 1b**), ventriculomegaly, thin corpus callosum and variable hypogonadism—ranging from absent gonads to ovarian and uterine remnants or atrophic and undescended testes (**Supplementary Fig. 1a,b**). All patients and families were enrolled in institutional review board (IRB)-approved protocols at referral institutions and provided consent for study. We performed whole-exome sequencing in the proband and parents from Egyptian families 1275 and 1603. Aligning the genomic variants uncovered a homozygous mutation in *TOE1* encoding p.Glu220Lys (NC_000001.11) within a shared haplotype of 500 kb (**Fig. 2a** and **Supplementary Fig. 2a,b**), which is indicative of shared ancestry for these families. All ten additional families enrolled subsequently proved positive for biallelic mutations in *TOE1* (**Fig. 2a** and **Supplementary Table 2**). The variants were all predicted to impair expression of full-length TOE1 or affect protein function[6] by altering well-conserved amino acids (**Fig. 2b**) and were not observed

A full list of affiliations appears at the end of the paper.

in our in-house database of 4,000 ancestry-matched exomes. We confirmed that each variant segregated according to a recessive mode of inheritance in all genetically informative members of each family, suggesting that *TOE1* biallelic mutations underlie PCH7.

To test whether *TOE1* missense mutations were likely to interfere with protein function, we modeled them on the structure of the protein encoded by paralogous *CNOT7* (UniProt, Q9UIV1) (**Supplementary Fig. 3a**). Most variant residues were on the surface rather than within the RNA-binding cleft, suggesting that they are not likely to directly affect deadenylase activity (**Fig. 2c**). Next, we established primary fibroblast cultures from individuals carrying homozygous mutations encoding p.Glu220Lys, p.Phe148Tyr and
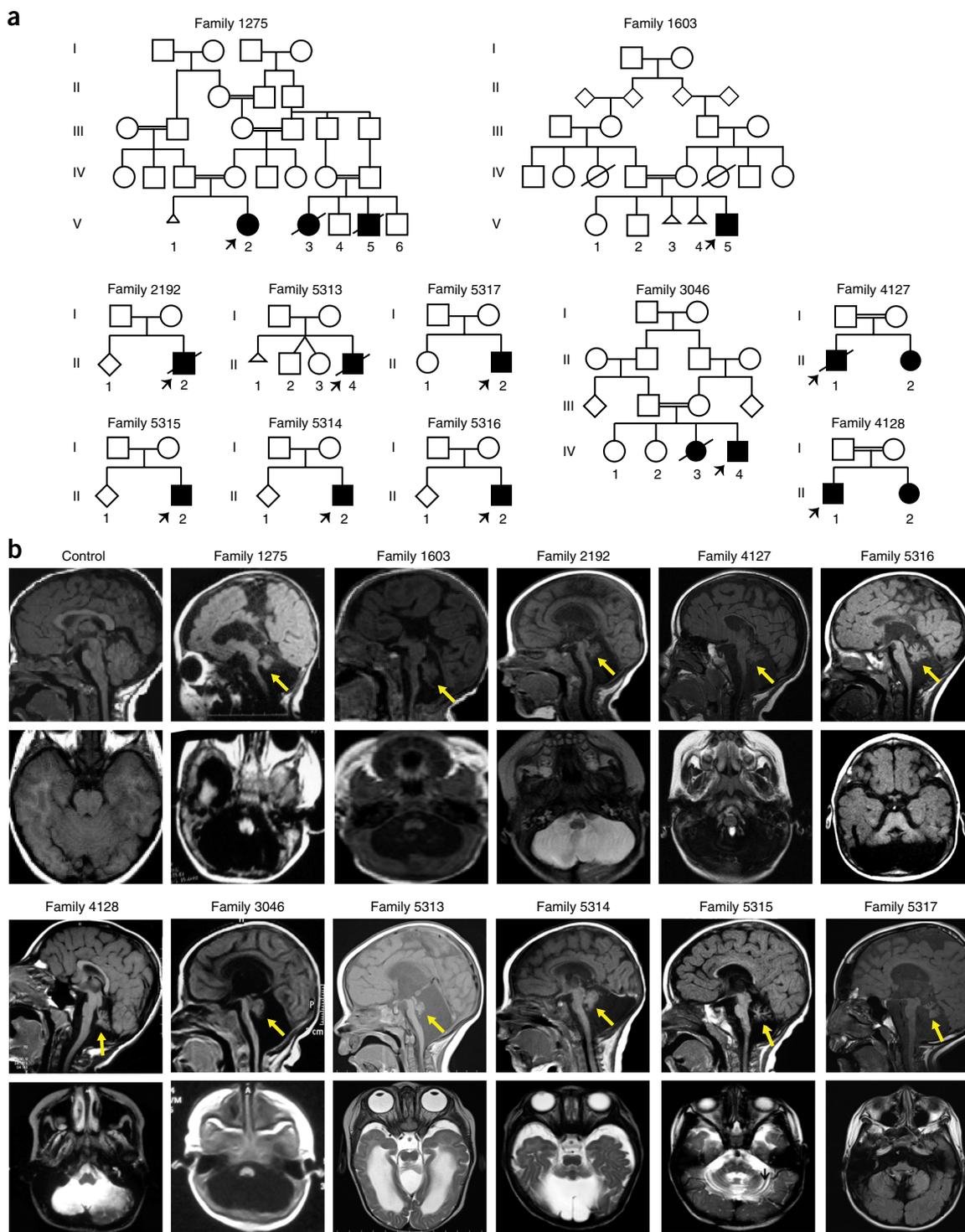
Figure 1 *TOE1* mutations lead to pontocerebellar hypoplasia with abnormal genitalia (PCH7). (**a**) Pedigrees of affected families showing recessive inheritance. Double bar, consanguineous marriage; open circle, unaffected female; open square, unaffected male; filled circle, affected female; filled square, affected male; triangle, spontaneous abortion; open diamond, unaffected individual of unknown sex; diagonal line, deceased. Arrows indicate probands. (**b**) Magnetic resonance midline sagittal (top) and axial (bottom) images showing reduced cerebellar volume in affected individuals (yellow arrows).
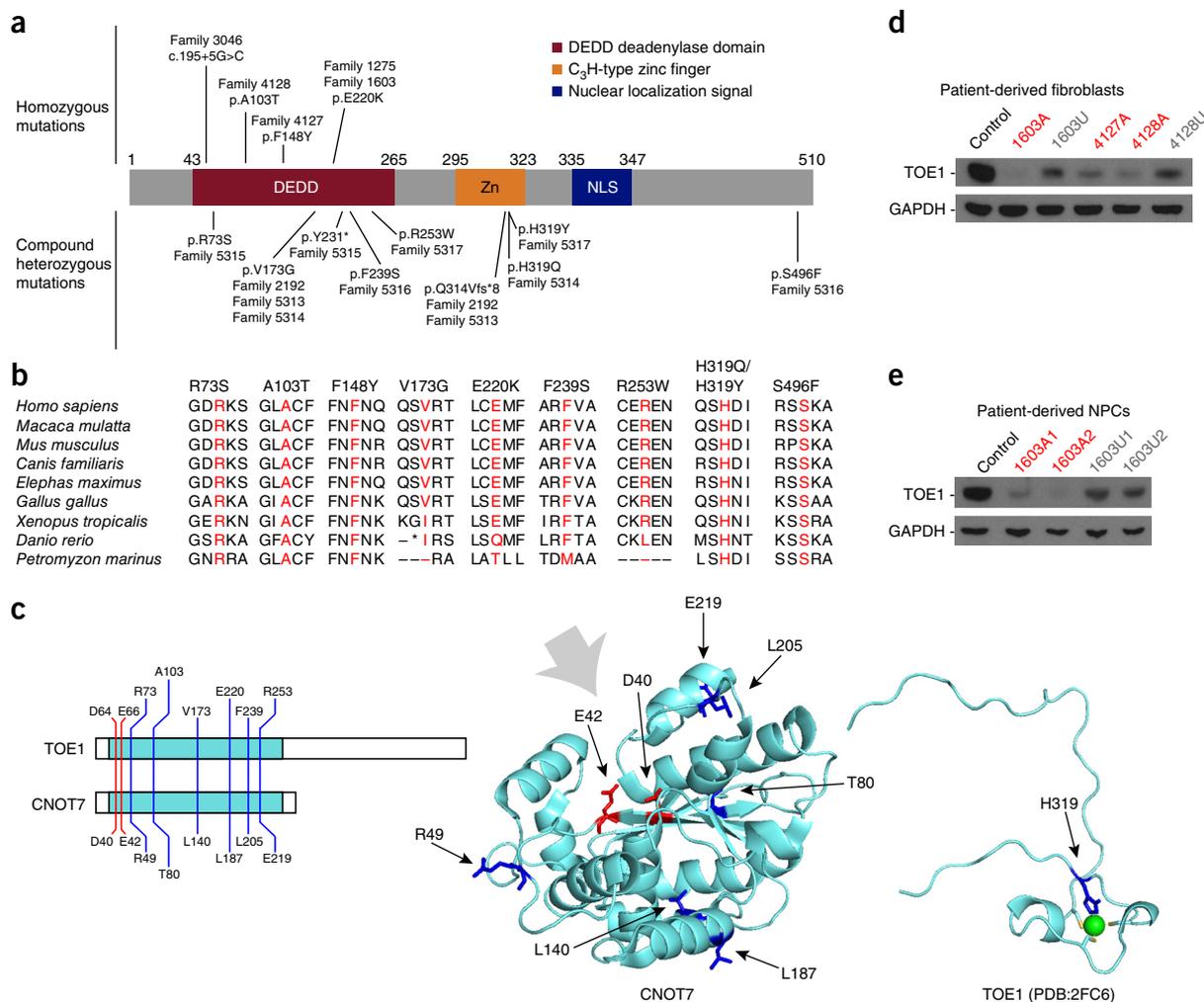
Figure 2 *TOE1* mutations map to conserved domains and amino acid residues and reduce protein levels. (**a**) TOE1 is a 510-amino-acid protein containing a DEDD deadenylase domain, a C_3H-type zinc finger and a nuclear localization signal (NLS). Identified homozygous mutations (bold) are shown above and compound heterozygous mutations are shown below, with corresponding family number. (**b**) The TOE1 positions affected by missense alterations (red) show amino acid conservation between species among vertebrates. (**c**) TOE1 alterations (blue) modeled onto the solved structures for CNOT7 (PDB 2D5R) and the C_3H zinc finger of TOE1 (both teal). Residues affecting deadenylase activity (Asp64, Glu66; red) are located in the RNA cleft (gray arrow), while the residues altered by mutations in patients are located on the protein surface. (**d**) Protein blots (cropped) showing reduced expression of TOE1 in fibroblasts from affected subjects (A) as compared with fibroblasts from related unaffected individuals (U). Control, ATCC fibroblast cell line. GAPDH, loading control. See **Supplementary Figure 3b** for full-length blots. (**e**) Impaired accumulation of TOE1 protein in NPCs derived from affected subjects (A) as compared with NPCs from related unaffected individuals (U). GAPDH, loading control.

p.Ala103Thr, as well as their unaffected relatives (families 1603, 4127 and 4128, respectively). Patient-derived fibroblasts had reduced levels of TOE1 protein as determined by protein blot (**Fig. 2d** and **Supplementary Fig. 3b**). Additionally, neural progenitor cell (NPC) lines that were derived from an affected individual of family 1603 showed less TOE1 protein than NPC lines from an unaffected relative and control NPCs (**Fig. 2e** and **Supplementary Fig. 3b,c**). Together, these results indicate that TOE1 amino acid substitutions negatively affect protein accumulation.

To assess whether *TOE1* mutations affect protein levels, we generated single-site-integration T-REx-293 cell lines for tetracycline-regulated expression of small interfering RNA (siRNA)-resistant transcripts for wild-type human TOE1 and the Glu220Lys, Phe148Tyr, Val173Gly and Ala103Thr variants. When cells were depleted of endogenous *TOE1* and induced with a concentration of tetracycline that promoted accumulation of wild-type TOE1 to near-endogenous levels, we observed reduced levels of mutant TOE1 as compared with wild-type

TOE1, despite similar mRNA levels (**Supplementary Fig. 3d,e**). In contrast, a previously characterized, catalytically inactive mutant of TOE1 (DE) accumulated to similar levels as wild-type protein (**Supplementary Fig. 3e**)[4]. These results are consistent with our findings from patient-derived cell lines and indicate that the mutations in affected individuals are deleterious by reducing TOE1 levels.

Because human *TOE1* and mouse *Toe1* were expressed in all tested tissues (**Supplementary Fig. 4a,b**), we generated *Toe1*-mutant mice to test for defects *in vivo*. Embryos with homozygous *Toe1* frameshift mutations showed uniform lethality before embryonic day (E) 11.5 (**Supplementary Fig. 4c**), demonstrating that *Toe1* is required for mouse development. Because the mutations in human patients allow for partial expression of TOE1 protein, we next turned to morpholino (MO)-based knockdown in zebrafish, where protein dosage could be regulated, to create a PCH7 disease model (**Supplementary Fig. 5a**). Knockdown of the single *toe1* ortholog (NM_001256682.1) led to a reproducible phenotype comprising microcephaly, small eyes and
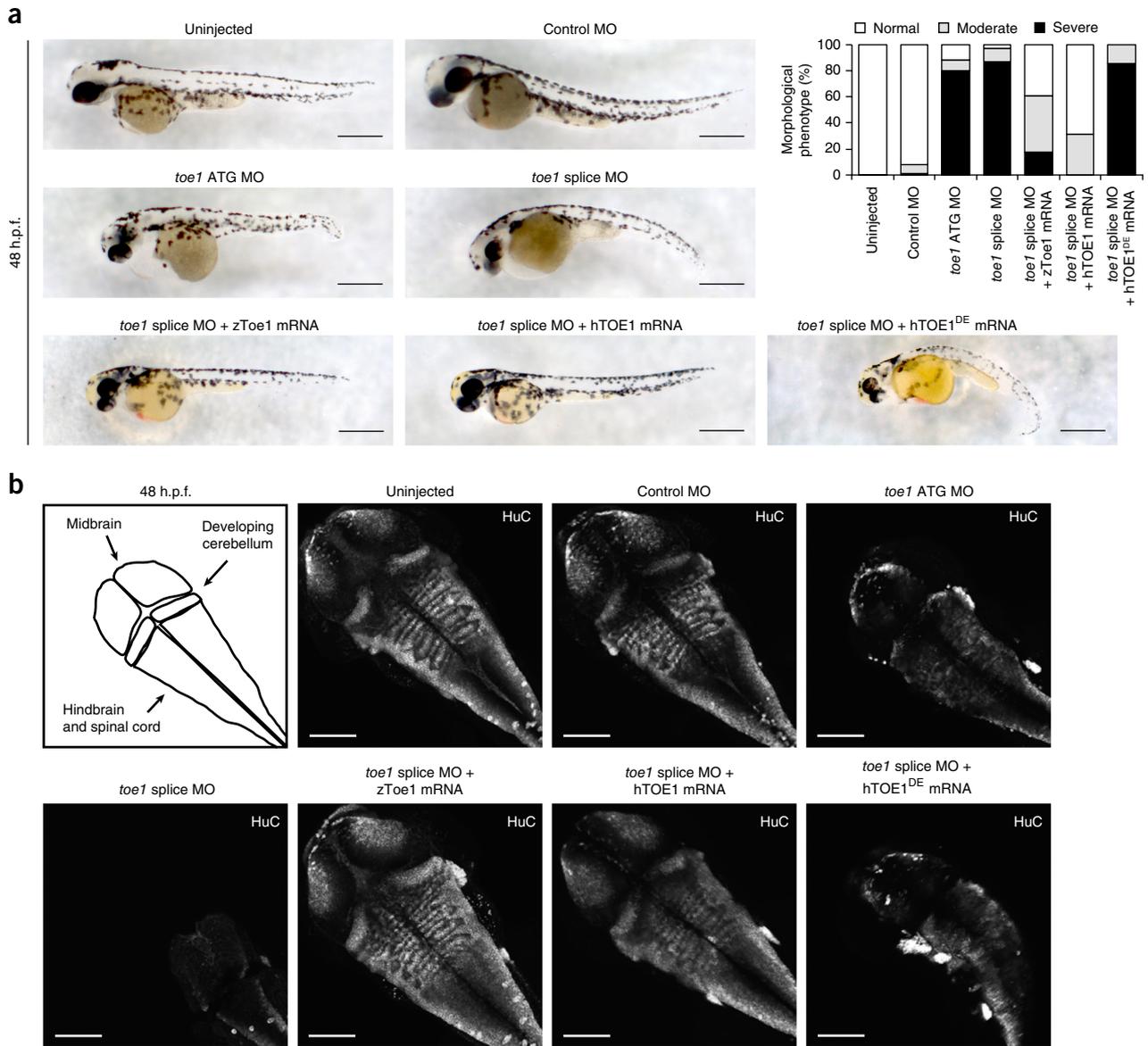
Figure 3 *toe1* depletion in zebrafish results in structural brain defects, mimicking human PCH7 pathology. (**a**) Comparison of zebrafish at 48 h.p.f. injected with 6 ng of control MO, *toe1* ATG MO, *toe1* splice-blocking MO, or *toe1* splice-blocking MO together with 0.1 pg of mRNA encoding zebrafish (z) Toe1 or human (h) TOE1. Zebrafish injected with *toe1* MO have abnormal head shape and thin, curved tails; this phenotype is rescued by addition of wild-type mRNA for zebrafish or human TOE1 but not mRNA for the human DE mutant. Scale bars, 500 μm. Quantification is shown at the top right. Normal, no observable phenotype; moderate, small eyes, slight reduction in head size; severe, small eyes, small head, thin curved tail. $n = 100$ embryos per. (**b**) Maximum confocal projection of whole-mount pan-neuronal HuC immunofluorescence at 48 h.p.f. A diagram of zebrafish anatomy is shown to the upper left. The midbrain, cerebellum and hindbrain regions of zebrafish injected with *toe1* MO have reduced HuC protein levels, which were restored with mRNA encoding wild-type zebrafish or human TOE1 but not the human DE mutant. Scale bars, 100 μm. $n = 6$ embryos per condition.

curly tail in 90% of embryos by 48 hours post-fertilization (h.p.f.), which was rescued by co-injection with human *TOE1* mRNA but not mRNA encoding the catalytically inactive DE mutant or the mutants identified in patients (**Fig. 3a** and **Supplementary Fig. 6a**). To visualize neurons, we performed whole-mount immunofluorescence for the neuronal marker HuC. Like the human patients, zebrafish injected with *toe1* MO showed structural defects of the developing midbrain, cerebellum and hindbrain (**Fig. 3b**). This phenotype was largely rescued by co-injection with wild-type zebrafish *toe1* or human *TOE1*, but not the mutant mRNAs (**Fig. 3b** and **Supplementary Figs. 5b** and **6b**). To determine whether neuronal loss in the zebrafish injected with *toe1* MO was due to cell death, we performed staining for cleaved

caspase-3 (Casp3) at 24 h.p.f. While zebrafish injected with control MO showed few Casp3-positive cells, those injected with *toe1* MO showed a dramatic apoptotic response that was consistent with neurodegeneration (**Supplementary Fig. 5c**). We conclude that reduced expression of TOE1 leads to neurodegeneration and PCH-like structural brain defects *in vivo*.

TOE1 was originally identified as a growth suppressor and a direct target gene of ERG1, an immediate-early transcription factor[3]. We reported TOE1, which is alternatively called CAF1z because of homology with CAF1 deadenylases, as a 3′-to-5′ exonuclease with a preference for adenosines *in vitro*[4]. Unlike characterized mRNA deadenylases, TOE1 is concentrated in nuclear Cajal bodies[4]. Cajal bodies
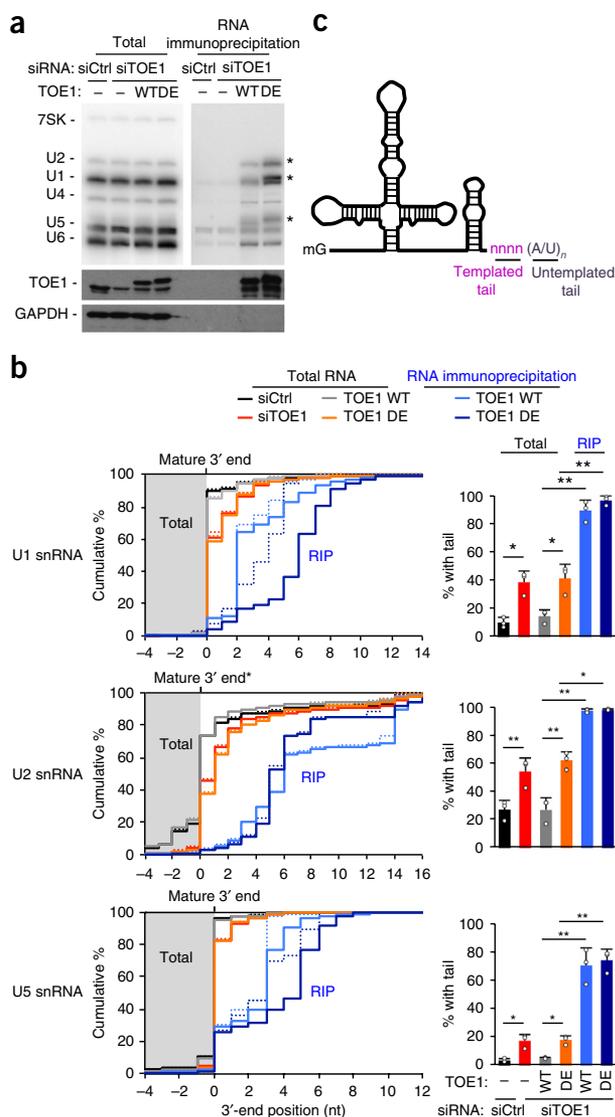
Figure 4 TOE1 targets 3′-extended pre-snRNAs. (**a**) RNA (top) and protein (middle and bottom; cropped) blots for cells treated with control (siCtrl) or *TOE1* (siTOE1) siRNA and then induced to express either wild-type (WT) or DE FLAG-TOE1 at near-endogenous levels. Input samples are shown to the left, and samples from RNA immunoprecipitation with antibody to FLAG are shown to the right. Asterisks mark slower-migrating snRNA species associated with TOE1 DE, suggesting impaired processing. (**b**) Left, cumulative plots of sequence reads for snRNA 3′ ends corresponding to the average of three independent experiments. Position 0 refers to the mature 3′ end of the snRNAs[33], with shaded areas corresponding to 3′-end positions within the mature snRNAs. Reads terminating at position −4 or further downstream are represented (see **Supplementary Fig. 8b** for all reads). Dotted lines represent the 3′-end positions of genome-templated snRNA sequences, and solid lines mark the 3′-end positions of snRNA sequences, including untemplated tails. RIP, RNA immunoprecipitation. Right, bar graphs showing the average percentage of snRNA reads with 3′ tails from three independent experiments. Independent experiments are represented by dots. Error bars represent s.d. from three independent experiments, and *P* values were determined by Student's two-tailed paired *t* test: *$P < 0.05$, **$P < 0.01$. The cumulative plots and bar graphs for the U1 and U5 snRNAs represent reads from all snRNA variants, whereas for the U2 snRNA reads are only from the *RNU2-1* gene (see **Supplementary Fig. 8b** for *RNU2-2P*) and the 3′ adenosine added to mature RNU2-1 snRNAs was left out of the analysis to allow visualization of exonucleolytic processing. (**c**) Schematic of a U1 snRNA processing intermediate with 3′-end templated tail (encoded nucleotides) and untemplated tail (unencoded nucleotides).

are rich in enzymes that process RNAs not known to have poly(A) tails, suggesting that TOE1 might target non-poly(A) RNA substrates. More recently, TOE1 was shown to associate with spliceosomal proteins[7,8], which are known to localize to Cajal bodies, and TOE1 depletion caused defective splicing of a pre-mRNA reporter[7]. To validate TOE1 association with the spliceosome, we performed *TOE1* knockdown together with expression at near-endogenous levels of FLAG-tagged wild-type or DE mutant TOE1 in T-REx-293 cells followed by immunoprecipitation (IP). As assessed by RNA blotting (**Fig. 4a**) and tandem mass spectrometry (MS) (**Supplementary Fig. 7** and **Supplementary Table 3**), TOE1 assembled with snRNAs and, in near-complete overlap with previously reported immunoprecipitation and mass spectrometry (IP–MS) results[7], with spliceosomal proteins. The CCR4-like protein ANGEL2 (also known as CCR4D), which we previously reported in complex with TOE1 (ref. 4), was not detected in our IP–MS analysis, possibly owing to low cellular abundance of this protein[9].

Inspection of RNA blots for snRNAs associated with the DE TOE1 catalytic mutant revealed slower-migrating U1, U2 and U5 snRNA species (**Fig. 4a**). Processing of RNA polymerase II (Pol II)-transcribed snRNAs (U1, U2, U4 and U5) initiates with co-transcriptional cleavage by the Integrator complex downstream of the mature 3′ end[10], but the mechanism mediating the removal of the 3′ tail to produce a mature-length snRNA is unknown[11,12]. To characterize the snRNAs that migrated slower in the RNA blots, we performed 3′-end sequencing of TOE1-associated snRNAs and found that TOE1 bound to Pol II–transcribed snRNAs that were incompletely processed at the 3′ end (**Fig. 4b,c** and **Supplementary Fig. 8**), suggesting that TOE1 may mediate snRNA 3′-tail processing.

The 3′-tail sequences of TOE1-associated pre-snRNAs consisted of both genome-encoded and post-transcriptionally added nucleotides, henceforth referred to as templated and untemplated tails, respectively (**Fig. 4b,c** and **Supplementary Fig. 9a**). The untemplated snRNA tails, which were previously observed by global 3′-end sequencing of noncoding RNA[13], were found almost exclusively on snRNAs that were either longer or shorter than their annotated mature length and consisted primarily of uridines and adenosines, with snRNAs associated with DE TOE1 predominantly enriched for untemplated 3′ adenosines (**Supplementary Fig. 9a,b**). Interestingly, 3′ maturation of U2 snRNA finishes with the addition of an untemplated 3′ adenosine[14], but this modification was entirely absent from TOE1-associated U2 snRNAs (**Supplementary Fig. 9b**). Taken together, these observations suggest that TOE1 associates specifically with pre-snRNAs that are not fully processed at the 3′ end and have often acquired untemplated tails.

snRNAs associated with DE TOE1 contained longer tails than those associated with wild-type TOE1, suggesting that TOE1 may catalytically process these tails. In accordance with this hypothesis, 3′-end sequencing of the total snRNA pool showed an increased fraction of 3′-end-extended snRNAs upon depletion of *TOE1* (**Fig. 4b** and **Supplementary Fig. 8**). Notably, complementation with exogenous wild-type TOE1 rescued the snRNA 3′-end defect, while addition of DE TOE1 failed to rescue the defect (**Fig. 4b** and **Supplementary Fig. 8**). There was little to no accumulation of 3′ tails for C/D-box U3 and H/ACA-box SNORA63 small nucleolar RNAs (snoRNAs), 5.8S rRNA, tRNAs and U6 snRNA (**Supplementary Figs. 8** and **10**), which, like Pol II–transcribed snRNAs, are processed from 3′-extended precursor RNAs[15–18]. These results support a catalytic role for TOE1 as a 3′-to-5′ exonuclease with specificity for snRNA processing.

TOE1 could promote either maturation or degradation of pre-snRNAs. To distinguish between these possibilities, we tested the activity of immunoisolated TOE1 on co-purifying pre-snRNA

Figure 5 TOE1 enzymatic activity processes snRNA 3′ ends. (**a**) RNA from TOE1 activity assays. Left, input RNA samples from cells treated with siCtrl or siTOE1 and then induced to express either wild-type or DE FLAG-TOE1 at near-endogenous levels. Right, samples from RNA immunoprecipitation with antibody to FLAG were divided and treated with either 2 mM Mg^{2+} or 2 mM EDTA on beads after IP; asterisks mark TOE1-processed snRNAs. (**b**) Cumulative plots showing the 3′-end positions of wild type– or DE TOE1–associated U1 and U5 snRNAs after incubation with Mg^{2+} or EDTA. The shaded area corresponds to reads terminating within the mature snRNAs. For U5 snRNA, an alternative 3′ end (alt 3′ end) is indicated. (**c**) Primary Pol II–transcribed snRNAs (pri-snRNAs) are co-transcriptionally cleaved by the Integrator complex. Remaining 3′ tails are subsequently processed to mature length by TOE1, in a process that might involve a 3′ terminal nucleotidyltransferase (3′ TnT).

Figure 6 Patient-derived fibroblasts and neuronal progenitor cells show defects in snRNA 3′ ends. (**a**) Top, cumulative U1 and U2 snRNA 3′-end sequence reads for fibroblasts from affected (A) and unaffected (U) individuals, showing that *TOE1* mutations result in snRNA tails. Bottom, bar graph showing the percentage of reads with 3′ tails. (**b**) Top, cumulative U1 and U2 snRNA 3′-end sequence reads for NPCs from affected and unaffected individuals showing that *TOE1* mutations result in snRNA tails in NPCs. Bottom, percentage of reads with 3′-end tails.

ribonucleoprotein complexes (pre-snRNPs) *in vitro*. RNA blot analyses showed Mg^{2+}- and active-site-dependent shortening of U1, U2 and U5 pre-snRNAs by TOE1 with no overall loss of snRNA levels (**Fig. 5a**),

suggesting that TOE1 has a function in maturation rather than degradation. 3′-end sequencing confirmed that the *in vitro* processing of U1, U2 and U5 by TOE1 halted at or before the mature 3′ end (**Fig. 5b** and **Supplementary Fig. 11a**), including at an alternative U5 3′ end also present in total cellular snRNA (**Supplementary Fig. 8b**). In accordance with TOE1 having a function in maturation rather than degradation, *TOE1* depletion did not affect the levels of unstable endogenous variants of U1 snRNA (**Supplementary Fig. 11b**)[19], nor did it affect the levels or 3′ processing of unstable U1 snRNAs mutated in the Sm-binding site (**Supplementary Fig. 11c,d**)[20], suggesting that the latter are targeted for degradation before 3′ processing by TOE1. We conclude that TOE1 acts in snRNA 3′-end maturation (**Fig. 5c**).

To determine whether *TOE1* mutations affect snRNA processing in the individuals in whom they are present, we performed snRNA 3′-end sequencing in the patient-derived cell lines. This identified an increased fraction of U1 and U2 snRNAs, as well as U5 snRNA to a lesser extent, that contained tails in fibroblasts derived from patients with *TOE1* mutations when compared to those derived from their unaffected relatives, which were indistinguishable from the control (**Fig. 6a** and **Supplementary Fig. 12a,b**). Similarly, patient-derived NPC lines showed enrichment for snRNA extensions (**Fig. 6b** and **Supplementary Fig. 12a,b**). We conclude that cells derived from patients with PCH7 accumulate incompletely processed snRNAs.

In summary, we have shown that individuals with PCH7 harbor biallelic, damaging mutations in *TOE1* that result in the accumulation of incompletely processed snRNAs. The involvement of TOE1 in snRNA maturation was surprising, given the homology of TOE1 with CAF1 mRNA deadenylases and its preference for poly(A) RNA *in vitro*[4]. However, this result could provide a mechanistic explanation for

the observed defect in pre-mRNA splicing upon *TOE1* depletion[7]. The enrichment of snRNAs with untemplated 3′ adenosines associated with catalytically inactive TOE1 suggests that 3′ adenylation by an unknown 3′-terminal nucleotidyltransferase (3′ TnT) serves as a mechanism to recruit TOE1 to pre-snRNAs (**Fig. 5c**). The abundance of Sm subunits in the TOE1 IP–MS analysis (**Supplementary Fig. 7** and **Supplementary Table 3**) and the absence of processing of U1 snRNA mutated in the Sm-binding site (**Supplementary Fig. 11d**) suggest that the Sm complex is another component important for TOE1 recruitment. PARN, which has been linked to familial pulmonary fibrosis, and PARN-like DEDD deadenylases were recently found to promote 3′ processing of snoR-NAs and PIWI-interacting RNAs (piRNAs), respectively[15,21,22], and, like snRNAs, RNA targets had templated and untemplated tails upon depletion of *PARN*[15]. Thus, the adaptation of deadenylases to function in 3′ processing of noncoding RNAs in conjunction with poly(A) polymerases might be a general principle in RNA metabolism.

Our results suggest that perturbations in snRNA pools may contribute to the features manifested by patients with PCH7, although other currently unknown functions of TOE1 cannot be ruled out as being causal in the disease. A defining feature of pontocerebellar degeneration is loss of motor neurons, akin to spinal muscular atrophy (SMA). The gene associated with SMA, *SMN1* (survival of motor neuron 1), encodes a well-established snRNP assembly factor[23,24], and this association, together with our results, suggests that maintaining proper snRNP biogenesis is important for neuronal survival. *EXOSC3*, which encodes a core component of the 3′-RNA-exonuclease exosome complex, is mutated in PCH type 1 (ref. 25), suggesting that there may be shared RNA targets for TOE1 and the exosome, perhaps snRNAs. In accordance with the idea that snRNAs are important for neuronal survival, a recent study identified recessive cerebellar degeneration resulting from a mutation in the *Rnu2-8* U2 snRNA gene in mice[26]. We and others recently described mutations in genes involved in protein synthesis, including *CLP1* and other tRNA splicing endonuclease (TSEN) complex components, leading to misprocessing of pre-tRNAs in cells derived from patients with PCH[27–30]. While we found no defect in tRNA 3′-end processing in *TOE1*-mutant cells (**Supplementary Fig. 10**), CLP1 has also been implicated in snRNA 3′-end processing[31], suggesting a possible link between *TOE1* and *CLP1* mutations in PCH. Our data, along with data from other recent studies[26,32], suggest defects in the processing of small noncoding RNAs as a common cause of severe, early-onset neurodegenerative conditions.

URLs. NHLBI Exome Sequencing Project (ESP), http://evs.gs.washington.edu/EVS/; SeattleSeq, http://snp.gs.washington.edu/SeattleSeqAnnotation137/; OMIM, http://www.omim.org/; PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/; Greater Middle Eastern Variome, http://igm.ucsd.edu/gme/.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

AUTHOR CONTRIBUTIONS

A.E.S., S.M., N.A., A.G.-G., L.D.H. and A.G. performed fibroblast, iPSC, NPC and knockout mouse experiments. R.M.L. generated stable cell lines and performed all snRNA/P experiments. V.R.C.E., N.A., E.S., D. Musaev, R.M., A.W., A.E.S., J.L.S., E.D., R.O.R. and H.T.C. analyzed clinical and exome results. S.G., A.E.S., Z.S., B.R., A.G.-G., I.M.-V. and L.D.H. generated zebrafish data, and N.C.C. and D.T. provided resources for zebrafish experimentation. M.S.Z., N.F., W.B.D., L.S., S.M.B., E.M.V., J.H.D., L.d.M., H.K., U.A., M.L.F., L.W., D.C., S.B., C.F., M.K., K.A.A., D. Manchester, N.M., A.O.Ç., K.B., H.P., M.-C.N., H.S., M.S., K.O. and K.M. conceived of the genetic investigation. S.B.G. and M.G. supported exome sequencing. E.L.V.N., S.S., T.L.S. and G.W.Y. supported RNA sequencing and computational analysis. E.J.B. performed mass spectrometry. R.L.M., A.E.S., J.L.-A. and J.G.G. wrote the manuscript. R.M.L., A.E.S., V.R.C.E., Z.S., N.C.C., D.T., G.W.Y., F.B., J.L.-A. and J.G.G. edited the manuscript. J.L.-A., F.B. and J.G.G. directed the project.

1. Goldstrohm, A.C. & Wickens, M. Multifunctional deadenylase complexes diversify mRNA control. *Nat. Rev. Mol. Cell Biol.* **9**, 337–344 (2008).
2. Anderson, C., Davies, J.H., Lamont, L. & Foulds, N. Early pontocerebellar hypoplasia with vanishing testes: a new syndrome? *Am. J. Med. Genet. A.* **155A**, 667–672 (2011).
3. De Belle, I., Wu, J.X., Sperandio, S., Mercola, D. & Adamson, E.D. *In vivo* cloning and characterization of a new growth suppressor protein TOE1 as a direct target gene of Egr1. *J. Biol. Chem.* **278**, 14306–14312 (2003).
4. Wagner, E., Clement, S.L. & Lykke-Andersen, J. An unconventional human Ccr4–Caf1 deadenylase complex in nuclear cajal bodies. *Mol. Cell. Biol.* **27**, 1686–1695 (2007).
5. Namavar, Y. *et al.* Clinical, neuroradiological and genetic findings in pontocerebellar hypoplasia. *Brain* **134**, 143–156 (2011).
6. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248–249 (2010).
7. Fong, K.W. *et al.* Whole-genome screening identifies proteins localized to distinct nuclear bodies. *J. Cell Biol.* **203**, 149–164 (2013).
8. Will, C.L. *et al.* The human 18S U11/U12 snRNP contains a set of novel proteins not found in the U2-dependent spliceosome. *RNA* **10**, 929–941 (2004).
9. Nagaraj, N. *et al.* Deep proteome and transcriptome mapping of a human cancer cell line. *Mol. Syst. Biol.* **7**, 548 (2011).
10. Baillat, D. *et al.* Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**, 265–276 (2005).
11. Madore, S.J., Wieben, E.D. & Pederson, T. Intracellular site of U1 small nuclear RNA processing and ribonucleoprotein assembly. *J. Cell Biol.* **98**, 188–192 (1984).
12. Wieben, E.D., Nenninger, J.M. & Pederson, T. Ribonucleoprotein organization of eukaryotic RNA. XXXII. U2 small nuclear RNA precursors and their accurate 3′ processing *in vitro* as ribonucleoprotein particles. *J. Mol. Biol.* **183**, 69–78 (1985).
13. Goldfarb, K.C. & Cech, T.R. 3′ terminal diversity of MRP RNA and other human noncoding RNAs revealed by deep sequencing. *BMC Mol. Biol.* **14**, 23 (2013).

14. Cho, H.D., Tomita, K., Suzuki, T. & Weiner, A.M. U2 small nuclear RNA is a substrate for the CCA-adding enzyme (tRNA nucleotidyltransferase). *J. Biol. Chem.* **277**, 3447–3455 (2002).

15. Berndt, H. *et al.* Maturation of mammalian H/ACA box snoRNAs: PAPD5-dependent adenylation and PARN-dependent trimming. *RNA* **18**, 958–972 (2012).

16. Lund, E. & Dahlberg, J.E. Cyclic 2′,3′-phosphates and nontemplated nucleotides at the 3′ end of spliceosomal U6 small nuclear RNA's. *Science* **255**, 327–330 (1992).

17. Granneman, S. & Baserga, S.J. Crosstalk in gene expression: coupling and co-regulation of rDNA transcription, pre-ribosome assembly and pre-rRNA processing. *Curr. Opin. Cell Biol.* **17**, 281–286 (2005).

18. Phizicky, E.M. & Hopper, A.K. tRNA biology charges to the front. *Genes Dev.* **24**, 1832–1860 (2010).

19. O'Reilly, D. *et al.* Differentially expressed, variant U1 snRNAs regulate gene expression in human cells. *Genome Res.* **23**, 281–291 (2013).

20. Shukla, S. & Parker, R. Quality control of assembly-defective U1 snRNAs by decapping and 5′-to-3′ exonucleolytic digestion. *Proc. Natl. Acad. Sci. USA* **111**, E3277–E3286 (2014).

21. Izumi, N. *et al.* Identification and functional analysis of the pre-piRNA 3′ trimmer in silkworms. *Cell* **164**, 962–973 (2016).

22. Tang, W., Tu, S., Lee, H.C., Weng, Z. & Mello, C.C. The RNase PARN-1 trims piRNA 3′ ends to promote transcriptome surveillance in *C. elegans. Cell* **164**, 974–984 (2016).

23. Fischer, U., Liu, Q. & Dreyfuss, G. The SMN–SIP1 complex has an essential role in spliceosomal snRNP biogenesis. *Cell* **90**, 1023–1029 (1997).

24. Liu, Q., Fischer, U., Wang, F. & Dreyfuss, G. The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell* **90**, 1013–1021 (1997).

25. Wan, J. *et al.* Mutations in the RNA exosome component gene *EXOSC3* cause pontocerebellar hypoplasia and spinal motor neuron degeneration. *Nat. Genet.* **44**, 704–708 (2012).

26. Jia, Y., Mu, J.C. & Ackerman, S.L. Mutation of a U2 snRNA gene causes global disruption of alternative splicing and neurodegeneration. *Cell* **148**, 296–308 (2012).

27. Schaffer, A.E. *et al.* *CLP1* founder mutation links tRNA splicing and maturation to cerebellar development and neurodegeneration. *Cell* **157**, 651–663 (2014).

28. Hanada, T. *et al.* CLP1 links tRNA metabolism to progressive motor-neuron loss. *Nature* **495**, 474–480 (2013).

29. Budde, B.S. *et al.* tRNA splicing endonuclease mutations cause pontocerebellar hypoplasia. *Nat. Genet.* **40**, 1113–1118 (2008).

30. Breuss, M.W. *et al.* Autosomal-recessive mutations in the tRNA splicing endonuclease subunit *TSEN15* cause pontocerebellar hypoplasia and progressive microcephaly. *Am. J. Hum. Genet.* **99**, 228–235 (2016).

31. Hallais, M. *et al.* CBC–ARS2 stimulates 3′-end maturation of multiple RNA families and favors cap-proximal processing. *Nat. Struct. Mol. Biol.* **20**, 1358–1366 (2013).

32. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323**, 1208–1211 (2009).

33. Guthrie, C. & Patterson, B. Spliceosomal snRNAs. *Annu. Rev. Genet.* **22**, 387–419 (1988).

[1]University of California San Diego, La Jolla, California, USA. [2]Laboratory of Pediatric Brain Disease and Howard Hughes Medical Institute, The Rockefeller University, New York, New York, USA. [3]Department of Cellular and Molecular Medicine, Stem Cell Program and Institute for Genomic Medicine, University of California San Diego, La Jolla, California, USA. [4]Department of Clinical Genetics, Academic Medical Center, Amsterdam, the Netherlands. [5]Clinical Genetics Department, Human Genetics and Genome Research Division, National Research Centre, Cairo, Egypt. [6]Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, California, USA. [7]Department of Paediatric Medicine, University Hospital Southampton NHS Foundation Trust, Southampton, UK. [8]Pediatric Neurology and Metabolic Diseases, Universitair Ziekenhuis Brussels, Vrije Universiteit Brussel, Brussels, Belgium. [9]Medical Genetics Department, Koc University School of Medicine, Istanbul, Turkey. [10]Medical Genetics Department, Istanbul Medical Faculty, Istanbul University, Istanbul Turkey. [11]Department of Clinical Genetics, The Canberra Hospital, Woden, Australian Capital Territory, Australia. [12]Australian Capital Territory Genetic Service, The Canberra Hospital, Canberra City, Australian Capital Territory, Australia. [13]Department of Pediatrics, Division of Clinical and Metabolic Genetics, The Hospital for Sick Children, Toronto, Ontario, Canada. [14]The Prenatal Diagnosis and Medical Genetics Program, Department of Obstetrics and Gynecology, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada. [15]Division of Neuroradiology, Department of Diagnostic Imaging, The Hospital for Sick Children, Toronto, Ontario, Canada. [16]Department of Medical Genetics, School of Medicine, Istanbul Bilim University, Istanbul, Turkey. [17]Yale Program on Neurogenetics, Departments of Neurosurgery, Neurobiology and Genetics, Yale University School of Medicine, New Haven, Connecticut, USA. [18]Department of Genetics, Yale Center for Genome Analysis, Yale University School of Medicine, New Haven, Connecticut, USA. [19]Division of Pediatric Neurology, Department of Pediatrics, Erciyes University School of Medicine, Kayseri, Turkey. [20]Department of Clinical Genetics, Odense University Hospital, Odense, Denmark. [21]Hans Christian Andersen Children's Hospital, Odense University Hospital, Odense, Denmark. [22]Center for Integrative Brain Research, Seattle Children's Research Institute, Seattle, Washington, USA. [23]Department of Pediatrics, Clinical Genetics and Metabolism, University of Colorado School of Medicine, Children's Hospital Colorado, Aurora, Colorado, USA. [24]Department of Human Genetics, Yokohama City University, Graduate School of Medicine, Yokohama, Japan. [25]Department of Pediatrics, Gunma University School of Medicine, Showa-machi, Maebashi City, Japan. [26]Department of Biochemistry, Hamamatsu University School of Medicine, Hamamatsu, Japan. [27]Department of Biochemistry, Yokohama City University Graduate School of Medicine, Yokohama, Japan. [28]Southampton University Hospitals Trust, Southampton, UK. [29]UCSD Cardiology, University of California San Diego, La Jolla, California, USA. [30]Clinical Genetics Unit, Department of Women, Mother and Neonates, "Vittore Buzzi" Children's Hospital, Istituti Clinici di Perfezionamento, Milan, Italy. [31]Child Neurology Unit, Department of Pediatrics, "Vittore Buzzi" Children Hospital, Istituti Clinici di Perfezionamento, Milan, Italy. [32]Broad Institute of Harvard and MIT, Cambridge, Massachusetts, USA. [33]Section of Neurosciences, Department of Medicine and Surgery, University of Salerno, Salerno, Italy. [34]Pediatric Neurology, Université Catholique de Louvain, Cliniques Universitaires Saint-Luc, Brussels, Belgium. [35]Department of Physiology, National University of Singapore and Molecular Engineering Laboratory, A*STAR, Singapore. [36]These authors contributed equally to this work. Correspondence should be addressed to J.L.-A. (jlykkeandersen@ucsd.edu), F.B. (f.baas@amc.uva.nl) or J.G.G. (jogleeson@rockefeller.edu).

ONLINE METHODS

Patient recruitment. Patients were enrolled and sampled according to standard local practice in approved human subjects protocols at the University of California, The Rockefeller University and The Academic Medical Center (AMC) in Amsterdam for blood, saliva and skin biopsy sampling.

Sequencing. Blood was acquired from informed, consenting individuals according to institutional guidelines, and DNA was extracted using established protocols. Exome sequencing was performed on both the parents and affected member(s) from each family as previously described[34]. All variants were prioritized by allele frequency, conservation and predicted effect on protein function, and were tested by Sanger sequencing for segregation with disease. Sequence data were analyzed with Sequencer 4.9 (Gene Codes).

Genetic mapping. Chromosomal ideogram plots were generated using the Bioconductor package quantsmooth. Red represents homozygous regions that segregate between affected and unaffected family members.

Fibroblast culture, and iPSC and NPC generation. Fibroblasts were generated from unaffected and affected dermal biopsy explants. Induced pluripotent stem cells (iPSCs) and NPCs were obtained as previously described[35]. Mycoplasma testing was routinely performed, and all cell lines were negative.

cDNA synthesis and RT–PCR. cDNA was synthesized with the Superscript III First-Strand cDNA synthesis system for RT–PCR (Life Technologies) and used for real-time PCR or cloning. qRT–PCR for intron-containing tRNAs was performed in triplicate on 10 ng of human cDNA[27]. RT–PCR to assess *toe1* intron inclusion in zebrafish embryos morphant for *toe1* splice-blocking MO was performed with 10 ng of zebrafish cDNA and primers flanking the first intron of zebrafish *toe1* (**Supplementary Table 4**).

Plasmid constructs and stable cell line generation. The ORF of human *TOE1* was subcloned into pcDNA5/FRT/TO-FLAG with BamHI and NotI restriction sites from pcDNA3-FLAG[4]. The ORF of zebrafish *toe1* or human *TOE1* was subcloned into pCS2+ (ref. 36) with BamHI and XhoI. Missense mutations were generated using the QuikChange Mutagenesis kit (Agilent). For expression of N-terminally FLAG-tagged proteins, stable HEK FLp-In T-REx-293 cell lines were generated (Invitrogen). Mycoplasma testing was routinely performed, and all cell lines were negative.

RNA blotting. RNA was extracted using TRIzol and separated by electrophoresis on 9% 19:1 polyacrylamide, 0.6× Tris/borate/EDTA (TBE), 8 M urea gels at 20 mA per gel for 2 h. RNA was transferred in 0.5× TBE to nylon membrane at 25 V for 16 h. Membranes were UV cross-linked, blocked with Ultrahyb Oligo (Life Technologies) and hybridized with 5′-end-radiolabeled DNA oligonucleotides (**Supplementary Table 4**).

3′ RNA tail sequencing and analysis. RNA adaptors containing barcodes and 10- to 11-nt randomers (**Supplementary Table 4**) were ligated to the 3′ ends of total and eluted RNA with T4 RNA ligase (New England BioLabs) at 16 °C for 16 h, treated with RQ1 RNase-free DNase (Promega) for 30 min at 37 °C and extracted with PCA (Affymetrix). cDNA was generated using AR-17 primer with Superscript III (Life Technologies). 3′ ends were amplified by Q5 DNA polymerase (New England BioLabs) with snRNA-gene-specific primers and AR-17, and then with primers D501 and D702 (Illumina; **Supplementary Table 5**). For tRNA 3′-end sequencing, tRNAs were gel purified from a 9% polyacrylamide urea gel and libraries were prepared using the eCLIP input RNA protocol (days 3 and 4)[37] with the AG10N and AG11N 3′ RNA adaptors and the 3Tr3 5′ DNA adaptor (**Supplementary Table 5**). All libraries were purified on AMPure XP magnetic beads (Beckman) and validated by 2200 TapeStation (Agilent). Libraries were pooled at a 4 nM concentration, denatured with NaOH and diluted in HT1 buffer (Illumina) to 7 pM and were then sequenced with MiSeq Reagent Kit V2, Nano configuration (Illumina). snRNA 3′ sequence reads were decomplexed, and duplicates were removed based on randomer sequences. Sequences were aligned with snRNA genes (including downstream regions) from Ensembl. snRNA 3′-end positions were identified based on perfect alignment with one or more snRNA genes (reads with

nucleotide mismatches were removed from the analyses), and 3′ untemplated tails were identified as 3′ sequence additions that did not align with the genomic sequence. Each experiment generally generated 5,000–35,000 unique mapped reads, with the lowest number being 1,173 (**Supplementary Table 4**). Cumulative plots were generated from reads mapping to all U1, U4, U5 and U6 variants, whereas for U2 separate cumulative plots were generated for reads mapping to the *RNU2-1* and *RNU2-2P* genes and reads were trimmed for the A and CA tails that are post-transcriptionally added to mature RNU2-1 and RNU2-2P RNAs, respectively, to allow discrimination of the mature U2 snRNAs from those that are incompletely 3′-end exonucleolytically processed (**Figs. 4** and **6**, and **Supplementary Figs. 8b** and **11a**).

RNA immunoprecipitation and TOE1 activity assays. One 10-cm dish of stable cell lines expressing siRNA-resistant FLAG-TOE1 variants was transfected twice with 20 nM siRNA (Dharmacon) targeting either luciferase (siCtrl) or the 3′ UTR of *TOE1* (siTOE1) using siLentFect (Bio-Rad) 72 h and 24 h before cell collection. Treatment with siTOE1 generally resulted in ~10–20% remaining protein, as determined by protein blot (**Supplementary Fig. 3d**). At ~50% density, cells were induced for 24 h with tetracycline (1 ng/ml), collected in PBS and lysed in isotonic buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.2 mM EDTA, 0.1% Triton X-100, 1 mM PMSF, 2 μg/ml aprotinin, 2 μg/ml leupeptin, 0.1 μg/ml yeast total RNA (Roche) and 80 U/ml RNaseOUT (Invitrogen) for 10 min on ice. Lysates were cleared at 20,000g for 15 min at 4 °C. FLAG peptide was added to 1 μg/ml, and lysates were incubated for 2 h with anti-FLAG M2 beads (Sigma) at 4 °C. Beads were washed eight times with NET2 (10 mM Tris, pH 7.5, 150 mM NaCl and 0.1% Triton X-100). One-fifth of the beads were eluted in protein loading buffer, and the rest were resuspended in TRIzol (Life Technologies). RNA was extracted according to the manufacturer's protocol. Activity assays were performed after washing and on the beads for 30 min at 25 °C with either 2 mM EDTA or 2 mM MgCl$_2$.

Protein blotting. Protein blotting was performed with rabbit polyclonal antibody to Caf1z/Toe1 (ref. 4), rabbit polyclonal antibody to GAPDH (Cell Signaling, 2118s) and mouse antibody to vinculin (Sigma, V9264) at 1:1,000 dilutions in 5% nonfat milk in PBST. The secondary antibodies were HRP-conjugated anti-mouse and anti-rabbit secondary antibodies used at 1:20,000 dilutions in 5% nonfat milk in PBST.

Animals. All animal experiments complied with the Institutional Animal Care and Use Committee at the University of California San Diego and were carried out in a non-blinded fashion. The *Toe1*-mutant mice were generated using CRISPR/SpCas9 technology. Briefly, pronuclear co-injection with 5 ng of *Cas9* mRNA and 2.5 ng of sgRNA (targeting sequence, 5′-CTGTGTGAGATGTTCCCAGC-3′) was performed in 143 embryos. A total of 126 embryos were transferred into host dams for implantation, resulting in 23 live pups. Sanger sequence genotyping identified only one mouse with a heterozygous single-base-pair frameshift mutation (chr. 4: 116806688–89insA; c.668_669insA), resulting in a *Toe1* null allele. Mutagenesis was performed on C57BL/6J single-cell blastocysts. Positive founders were bred to establish lines transmitting the *Toe1* null allele. Male and female carriers were intercrossed at 6–8 weeks of age to assess the embryonic phenotype associated with *Toe1* mutation.

Adult male and female zebrafish (<18 months old) from wild-type (AB Tubingen) and transgenic strains were maintained under standard laboratory conditions. At least three adult pairs were used to generate embryos at 0–48 h.p.f. for each experiment. Translation-blocking antisense MO (7 ng), splice-blocking antisense MO (6 ng) and control non-targeting MO (7 ng) (**Supplementary Table 5**) (Gene Tools) were injected into one-cell-stage embryos. Gross morphology of zebrafish was assessed at 48 h.p.f., and embryos were defined as affected if they had an obviously misshapen head, small eyes and a curly tail (>10% reduction in head/eye size and >10° change in the angle of the tail). HuC/HuD whole-mount immunostaining was used to assess the presence of neuronal tissue. For mRNA rescue experiments, 0.1 pg of *in vitro*–transcribed zebrafish *toe1* or human *TOE1* mRNA (mMESSAGE mMACHINE, Ambion) was co-injected with 6 ng of splice-blocking MO into one-cell-stage embryos, and embryos were assessed at 48 h.p.f. for morphological differences and neuronal tissue (minimum of 100 embryos per condition). Immunofluorescent staining was performed as previously described[38] with

primary rabbit polyclonal antibody to caspase-3 (Abcam, ab13847) or primary mouse monoclonal antibody to HuC/HuD (Thermo Fisher, A-21271) and Alexa Fluor 488–conjugated anti-rabbit or anti-mouse secondary antibody. Zebrafish were immobilized in agarose, and fluorescent images were acquired with a Zeiss LSM 780 and Olympus FV1000 confocal microscope. Cells were counted using ImageJ. No statistical method was used to predetermine sample size. The experiments were not randomized.

Immunocytochemistry. Cells were seeded on coverslips and fixed in 4% paraformaldehyde or 100% cold methanol, permeabilized with 0.1% Triton X-100, and blocked with 0.1% BSA and 0.5% gelatin from cold water fish skin in PBS. Cells were incubated with primary antibody in blocking solution overnight at 4 °C, washed and incubated with secondary antibody in blocking solution and 0.4 μg of Hoechst for 1 h at room temperature. The primary antibodies used were mouse antibody to nestin (EMD Millipore, MAB5326) and rabbit antibody to Pax6 (Covance, PRB-278P) at 1:1,000 dilutions. The secondary antibodies were Alexa Fluor 555–conjugated anti-rabbit and Alexa Fluor 488–conjugated anti-mouse antibody at 1:500 dilutions. Images were taken with an Olympus IX51 inverted fluorescent microscope.

Mass spectrometry. Liquid chromatography and tandem mass spectrometry (LC–MS/MS) assays were performed as described previously[39] using anti-FLAG IP samples from extracts of T-Rex-293 cell lines stably expressing FLAG-tagged wild-type or DE TOE1 at near-endogenous TOE1 levels, with a parental cell line serving as a control.

Protein modeling. Needleman–Wunsch alignment of human TOE1 and CNOT7 was performed with protein–protein BLAST (pBLAST). Homologous *TOE1* missense mutations, and previously published inactivating mutations[6], were modeled onto the predicted protein structures of human CNOT7 (PDB 2D5R) and TOE1 zinc-finger domain (PDB 2FC6). Patient mutations predicted to result in truncated protein (for example, nonsense, splice and frameshift mutations) were excluded from analysis. Residues not in alignment were excluded (Phe148). PyMOL was used to create 3D renderings.

Statistics. Student's two-tailed paired *t* test was employed to test the significance of accumulation of extended snRNAs as indicated (**Fig. 4b** and **Supplementary Fig. 8a**) (U1 snRNA: siCtrl vs. siTOE1, $P = 0.0054$; WT vs. DE, $P = 0.0030$; WT vs. RIP WT, $P = 0.00069$; DE vs. RIP DE, $P = 0.0050$; U2 snRNA: siCtrl vs. siTOE1, $P = 0.014$; WT vs. DE, $P = 0.014$; WT vs. RIP WT, $P = 0.0058$; DE vs. RIP DE, $P = 0.012$; U5 snRNA: siCtrl vs. siTOE1, $P = 0.032$; WT vs. DE, $P = 0.014$; WT vs. RIP WT, $P = 0.011$; DE vs. RIP DE, $P = 0.0027$; U4 snRNA: siCtrl vs. siTOE1, $P = 0.033$; WT vs. DE, $P = 0.016$; WT vs. RIP WT, $P = 0.12$; DE vs. RIP DE, $P = 0.042$; U6 snRNA: siCtrl vs. siTOE1, $P = 0.57$; WT vs. DE, $P = 0.015$; WT vs. RIP WT, $P = 0.021$; DE vs. RIP DE, $P = 0.19$). Student's two-tailed unpaired *t* test was employed to test the significance of cleaved caspase-3-positive cells in **Supplementary Figure 5c** (control MO vs. *toe1* ATG MO, $P = 0.015$).

Data availability. The exome sequencing data for all individuals consented for public release of data in this study have been deposited to the database of Genotypes and Phenotypes (dbGaP) under accession phs000288.v1.p1. RNA sequencing data have been deposited into Gene Expression Omnibus (GEO) under accession GSE71536.

34. Novarino, G. *et al.* Exome sequencing links corticospinal motor neuron disease to common neurodegenerative disorders. *Science* **343**, 506–511 (2014).
35. Chambers, S.M. *et al.* Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat. Biotechnol.* **27**, 275–280 (2009).
36. Turner, D.L. & Weintraub, H. Expression of achaete-scute homolog 3 in *Xenopus* embryos converts ectodermal cells to a neural fate. *Genes Dev.* **8**, 1434–1447 (1994).
37. Van Nostrand, E.L. *et al.* Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016).
38. Akizu, N. *et al.* Biallelic mutations in *SNX14* cause a syndromic form of cerebellar atrophy and lysosome–autophagosome dysfunction. *Nat. Genet.* **47**, 528–534 (2015).
39. Erickson, S.L. *et al.* Competition between decapping complex formation and ubiquitin-mediated proteasomal degradation controls human Dcp2 decapping activity. *Mol. Cell. Biol.* **35**, 2144–2153 (2015).