Genome Medicine

RESEARCH

Open Access

Variant-specific pathophysiological mechanisms of *AFF3* differently influence transcriptome profiles

Sissy Bassani[1,36], Jacqueline Chrast[1], Giovanna Ambrosini[2,3], Norine Voisin[1,40], Frédéric Schütz[4], Alfredo Brusco[5,6], Fabio Sirchia[5,6,37,38], Lydia Turban[7], Susanna Schubert[7], Rami Abou Jamra[7], Jan-Ulrich Schlump[8], Desiree DeMille[9], Pinar Bayrak-Toydemir[10], Gary Rex Nelson[10], Kristen Nicole Wong[10], Laura Duncan[11,39], Mackenzie Mosera[11], Christian Gilissen[12], Lisenka E. L. M. Vissers[12], Rolph Pfundt[12], Rogier Kersseboom[13], Hilde Yttervik[14], Geir Åsmund Myge Hansen[14], Marie Falkenberg Smeland[15], Kameryn M. Butler[16], Michael J. Lyons[16], Claudia M. B. Carvalho[17,18], Chaofan Zhang[18], James R. Lupski[18,19,20,21], Lorraine Potocki[18,21], Leticia Flores-Gallegos[22], Rodrigo Morales-Toquero[22], Florence Petit[23], Binnaz Yalcin[24], Annabelle Tuttle[25], Houda Zghal Elloumi[25], Lane McCormick[26], Mary Kukolich[26], Oliver Klaas[27], Judit Horvath[27], Marcello Scala[28,29], Michele Iacomino[29], Francesca Operto[30], Federico Zara[28,29], Karin Writzl[31,32], Aleš Maver[31], Maria K. Haanpää[33], Pia Pohjola[33], Harri Arikka[34], Anneke J. A. Kievit[35], Camilla Calandrini[35], Christian Iseli[2,3], Nicolas Guex[2,3] and Alexandre Reymond[1*]

Abstract

Background We previously described the KINSSHIP syndrome, an autosomal dominant disorder associated with intellectual disability (ID), mesomelic dysplasia and horseshoe kidney, caused by de novo variants in the degron of AFF3. Mouse knock-ins and overexpression in zebrafish provided evidence for a dominant-negative mode of action, wherein an increased level of AFF3 resulted in pathological effects.

Methods Evolutionary constraints suggest that other modes-of-inheritance could be at play. We challenged this hypothesis by screening ID cohorts for individuals with predicted-to-be damaging variants in *AFF3*. We used both animal and cellular models to assess the deleteriousness of the identified variants.

Results We identified an individual with a KINSSHIP-like phenotype carrying a de novo partial duplication of *AFF3* further strengthening the hypothesis that an increased level of AFF3 is pathological. We also detected seventeen individuals displaying a milder syndrome with either heterozygous Loss-of-Function (LoF) or biallelic missense variants in *AFF3*. Consistent with semi-dominance, we discovered three patients with homozygous LoF and one compound heterozygote for a LoF and a missense variant, who presented more severe phenotypes than their heterozygous parents. Matching zebrafish knockdowns exhibit neurological defects that could be rescued by expressing human *AFF3* mRNA, confirming their association with the ablation of *aff3*. Conversely, some of the human *AFF3* mRNAs carrying missense variants identified in affected individuals did not rescue these phenotypes. Overexpression of mutated *AFF3*

*Correspondence:
Alexandre Reymond
alexandre.reymond@unil.ch
Full list of author information is available at the end of the article

mRNAs in zebrafish embryos produced a significant increase of abnormal larvae compared to wild-type overexpression further demonstrating deleteriousness.

To further assess the effect of *AFF3* variation, we profiled the transcriptome of fibroblasts from affected individuals and engineered isogenic cells harboring +/+, KINSSHIP/KINSSHIP, LoF/+, LoF/LoF or KINSSHIP/LoF *AFF3* genotypes. The expression of more than a third of the AFF3 bound loci is modified in either the KINSSHIP/KINSSHIP or the LoF/LoF lines. While the same pathways are affected, only about one third of the differentially expressed genes are common to the homozygote datasets, indicating that *AFF3* LoF and KINSSHIP variants largely modulate transcriptomes differently, e.g. the DNA repair pathway displayed opposite modulation.

Conclusions Our results and the high pleiotropy shown by variation at this locus suggest that minute changes in *AFF3* function are deleterious.

Keywords Mesomelic dysplasia, Horseshoe kidney, Intellectual disability, Transcriptome, Zebrafish model

Background

AFF3 encodes the ALF Transcription Elongation Factor 3 (MIM*601464), a member of a gene family with four paralogs (*AFF1-4*) in mammals. These nuclear proteins function as transcriptional activators, promoting RNA elongation [1–3]. They share conserved N-terminal (NHD) and C-terminal homology domains (CHD) [4], an AF4-LAF4-FMR2 (ALF) domain [2, 3, 5], which contains the degron motif, a Serine-rich transactivation domain (TAD) [6], and a nuclear/nucleolar localization sequence (NLS) (Fig. 1A). AFF proteins are integral components of transcriptional super elongation complexes (SECs) that include positive transcription elongation factor (P-TEFb) [2, 3]. SECs are made of an AFF family member as scaffold, YEATS domain-containing MLLT proteins (myeloid/lymphoid or mixed-lineage leukemia; translocated to), and an ELL (Elongation Factor for RNA Polymerase II) protein [2]. By phosphorylating the C-terminal domain of RNA polymerase II, these complexes regulate the RNA transcription elongation process [3, 7]. Distinct combinations of components yield different SECs providing gene target specificity [2, 3]. AFF3 regulates the expression of genes involved in mesoderm and ectoderm development, as well as mesenchymal cell proliferation, cell adhesion, angiogenesis, cartilage and lens development, and immunoglobulin class switch recombination [8, 9]. It was recently linked with the establishment of biological rhythms, e.g. somitogenesis progression and niche switching [10, 11].

The SEC-L3 complex, which incorporates AFF3, is enriched at imprinted loci, for example at the lncRNA *XIST* locus that initiate X chromosome inactivation [16, 17]. AFF3 binds both silent and active chromatin regions to modulate expression of imprinted regions. For example, within the Dlk1-Dio3 interval, it is recruited by ZFP281 to the *Meg3* enhancer region to maintain an active chromatin state through H3K27ac modification and an allele-specific expression [8, 18].

We previously reported the association of *AFF3* alterations with KINSSHIP (horseshoe KIdney, Nievergelt/Savarirayan mesomelic dysplasia, Seizures, Hypertrichosis, Intellectual disability, and Pulmonary involvement) syndrome [11] (OMIM# 619297). Twenty-one affected individuals allowed delineation of its cardinal characteristics. Such individuals presented with developmental delay/intellectual disability (DD/ID), brain atrophy, epileptic encephalopathy, failure to thrive, horseshoe kidney, a specific mesomelic dysplasia, fibular hypoplasia, scoliosis, hypertrichosis, dysmorphic facial features, gastrointestinal, and pulmonary symptoms [11, 19]. This

(See figure on next page.)

Fig. 1 *AFF3* variants. **A** Schematic protein structure of AFF3 (NM_002285.3) with its N-terminal homology domain (NHD, cyan), the AF4-LAF4-FMR2 (ALF, pink) domain [2, 3, 5] containing the degron motif, a Serine-rich transactivation domain (TAD, yellow) [6], a bipartite nuclear/nucleolar localization sequence (NLS, green), and the C-terminal homology domain (CHD, blue) [4] showing positioning of all *AFF3* coding variants mentioned in the text. Missense and truncating variants are shown above, while extents of microdeletions and microduplications are depicted below the structure with continuous and dashed lines, respectively. The variants are color coded: loss-of-function (truncation and deletion) in red, biallelic missense outside the degron in blue and KINSSHIP-associated missense variants, deletion, and duplication in purple. The de novo missense identified in individuals M1 and M2 are shown in green. While the p.(Arg947Pro) shown in black was shown to segregate with isolated syndactyly [12], we have also identified it in individuals with no digit abnormalities. The 500 kb KINSSHIP deletion was originally wrongly defined as encompassing the entirety of the gene [13] but was later shown to only remove exons 4 to 13 of *AFF3* [11, 14, 15]. **B** UCSC genome browser snapshot of the Chr2 99.5 to 101.4-Mb region showing the genes mapping to this interval. The extent of the duplication identified in the DUP1 individual that encompasses exon 10 to exon 24 of *AFF3* and exon 1 to 3 of *REV1* is indicated by the black bar and the light blue shadow. **C** Examples of pedigrees of transmitting affected families suggesting semi-dominance

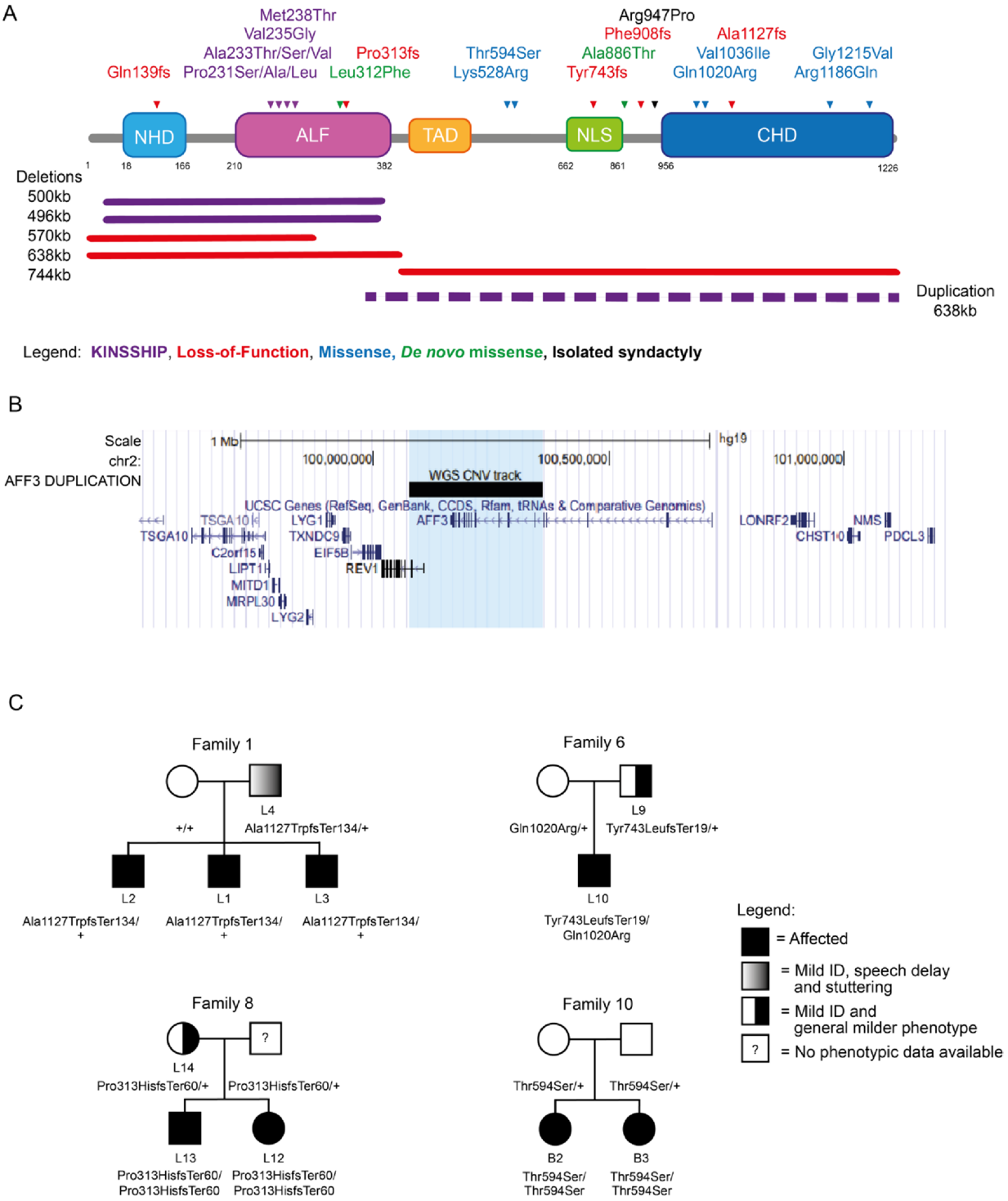


Fig. 1 (See legend on previous page.)

autosomal dominant disease is associated with de novo germline missense variants and deletions, as well as mosaic variants, in the conserved degron motif of AFF3 that are predicted to weaken or prevent binding to a SIAH ubiquitin ligase [11, 19](Figs. 1A and 2A,B). Both mouse knock-ins and overexpression in zebrafish suggested a dominant-negative mode of action, wherein an increased level of AFF3 resulted in pathological effects.

Consistent with this hypothesis, the corresponding de novo degron variants of AFF4 that are associated with CHOPS (Cognitive impairment and Coarse facies, Heart defects, Obesity, Pulmonary involvement, Short stature, and Skeletal dysplasia) syndrome (OMIM#616368) were shown to be more resistant to degradation upon co-transfection with the SIAH1 ubiquitin ligase [20, 21].

According to population metrics presented in GnomAD v4.0.0 [22], *AFF3* is under constraint with a pLI = 1 and a pLOEUF = 0.32, which suggests that *AFF3* haploinsufficiency could also be deleterious. Consistent with this hypothesis, mosaic CGG trinucleotide-repeat expansions in the promoter of *AFF3* that result in its hypermethylation and silencing were associated with mild ID, speech and motor delays, seizures, behavioral disturbances, generalized hypotonia, dysmorphic features, and congenital anomalies [23, 24].

Here we describe novel *AFF3* genetic alterations associated with an overexpression disease mechanism, as well as the effect of decreased AFF3 function through haploinsufficiency, homozygous truncation, and autosomal recessive inheritance. The affected individuals present symptoms that partially overlap those of KINSSHIP. The different mode of inheritance and their associated phenotypes are summarized in Fig. 2A.

Methods

Patient cohort

The affected individuals within our cohort were enrolled through collaboration of multiple diagnostics laboratories and data aggregation from the DECIPHER database. This article includes a comprehensive documentation of all the individuals presenting pathogenic causative variants in the open reading frame of *AFF3* we have identified up to April 2024.

Sample and variant identification

Informed consent forms were obtained for all affected individuals or their guardians participating in this study. The current study was approved by the CER ("Commission d'éthique de la recherche") of the canton of Vaud (Protocol number: CER-VD 2021–01400). This research complies with the principles of the Declaration of Helsinki. Affected individuals underwent genetic counselling and clinical examination followed by exome sequencing as described [25, 26] and/or array comparative genome hybridization, made exception of proband B1 who was sequenced with the Illumina TruSight One Expanded panel covering about one third of the exome. Affected individuals do not carry potentially causative variants in known ID genes. Genome sequencing was performed on the DUP1 individual (Table S1) to characterize the breakpoints of his duplication.

Protein model

Alignment of multiple AFF3 orthologous sequences was performed with the Clustal Omega tool [27, 28]. 3D modeling for AFF3 (UniProt: P51826) and SIAH1 (UniProt:Q8IUQ4) interaction was built using the Swiss-Pdb Viewer [29] as previously described [11].

Zebrafish husbandry

Zebrafish (*Danio rerio*, Oregon AB) were maintained at 28.5 °C on a 14:10 h light/dark cycle. Zebrafish are staged by hours (h) or days (d) post fertilization (pf). Adult zebrafish were housed in Active Blue racks (Tecniplast, Buguggiate, Italy) with a maximum of 20 fish per tank. All procedures complied with the European Convention for the Protection of Animals for Experimental and Scientific Purposes (ETS number 123) and the National Institutes of Health guide for the care and use of Laboratory animals. Housing and experiments were approved by the Vaud cantonal authority (authorization VD-H21).

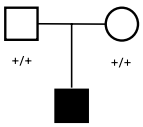
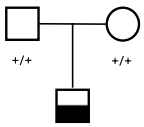
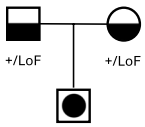
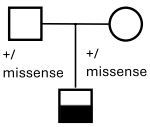
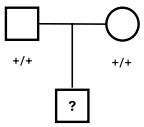
Zebrafish CRISPR-Cas9 model

The one-to-one ortholog of *AFF3* in zebrafish, *aff3*, encodes the five AFF domains. We generated F0 mutant zebrafish depleted for *aff3* by CRISPR/Cas9 genome editing. Two single synthetic guide RNAs (sgRNAs) targeting the coding sequence in *aff3* exon 6 of both isoforms annotated by Ensembl (Zebrafish GRCz11) (sgRNA_r2 5′- TCCAAAGCAGTACCCAGCCAAGG -3′; sgRNA_r19 5′- GCACCTGAGAATATATACCTTGG -3′) were designed with the CHOPCHOP tool [30, 31] and ordered

(See figure on next page.)

Fig. 2 **A** Variant-specific pathophysiological mechanisms of *AFF3*. Schematic summary of the different type of *AFF3* identified variants, their mode of inheritance, postulated mechanisms, and associated clinical phenotypes (see text for details). The number of patients we identified in each category and their identifiers are indicated. All symptoms are reported in Tables S1 and S2. Filled form = KINSSHIP syndrome, Half-filled form = milder syndrome, Form with included filled disk = more severe syndrome associated with semi-dominance, Question mark = possible association warranting further investigation. **B** Summary of in vivo and in vitro experiments. Schematic summary of traits associated with diminished expression and increased stabilization of AFF3 in mouse, zebrafish, HEK293 human cell, and affected individuals. The results previously published in Voisin et al., *AJHG* 2021 [11] are in blue, while results of this report are in black. The p.(A233T) variant is the most common de novo KINSSHIP variant [11], while the p.(M238T) and p.(M238V) variants described in this report present a milder phenotype (see text for details). Abbreviations: DD, developmental disorder; del, deletion; dup, duplication; ex, exon, ID, intellectual disability; LoF, loss-of-function; +, wild-type allele

A

Mechanism	Dominant negative	Haploinsufficiency	Semi-dominance	Recessive disorder	Dominant disorder
Mode of Inheritance	 +/missense in degron +/deletion of degron +/duplication	 +/LoF	 LoF/LoF	 missense outside degron/missense outside degron	 +/missense outside degron
Patients	28 patients (K1-K27 and DUP1)	10 patients (L1-L9 and L14)	3 patients (L11-L13)	7 patients (B1-B7)	2 patients (M1-M2)
Main Phenotype	**KINSSHIP syndrome** Severe DD/ID Muscle disorders Facial dysmorphisms Brain atrophy Epileptic encephalopathy Horseshoe kidney Mesomelic dysplasia Skeletal defects Hypertrichosis	Mild ID Muscle disorders Facial dysmorphisms Brain atrophy Abnormal corpus callosum Speech impairment	Moderate-severe DD/ID Facial dysmorphisms Anxiety Speech impairment Skeletal defects Respiratory defects	Mild DD/ID Epileptic encephalopathy No/Short attention span	DD Possibly ASD Mild speech impairment
Suggestive of semi-dominance, we also identified a patient compound heterozygous for a LoF and a missense *AFF3* variant (L10) presenting more severe phenotypes than his heterozygous parents (see Figure 1C).					

B

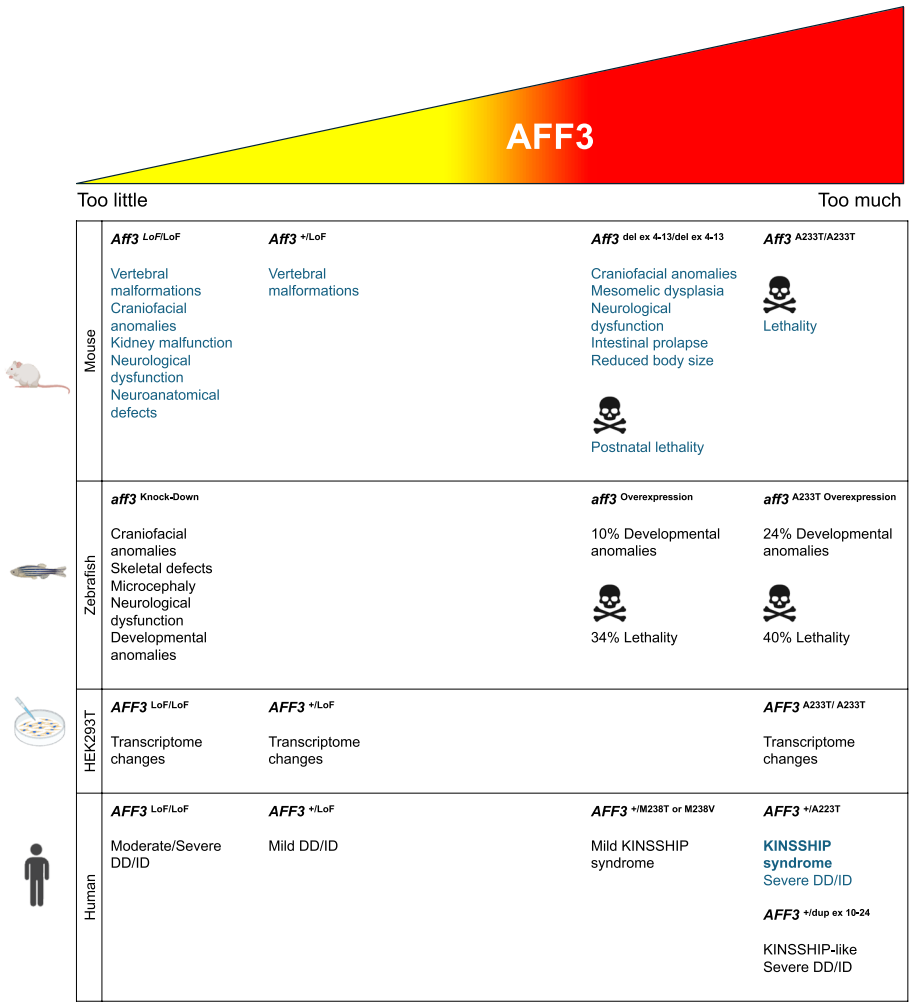


Fig. 2 (See legend on previous page.)

from Synthego, Redwood City, CA, USA. A total of 1 nl of a cocktail containing 50 ng/µl of gRNA_r2, 50 ng/µl of gRNA_r19, and 200 ng/µl of TrueCut™Cas9 v2 (Invitrogen) was injected into one- to two-cell stage embryos. In mock-injected larvae, the Cas9 was replaced by the same volume of water. KCl (200 mM) was added to increase efficiency. To determine the CRISPR-Cas9 targeting efficiency of each sgRNA in 5dpf founder (F0) mutants, a mismatch detection assay using T7 endonuclease 1 (New England Biolabs, Ipswich, MA, United States) was performed. Briefly, DNA was extracted, and PCR amplified with primers flanking the sgRNAs target site (5′- TCC AAAGCAGTACCCAGCCAAGGTATATATTCTCAG GTGC -3′). PCR products were denatured, reannealed, and incubated with T7 for 15 min at 37 °C. The reaction was stopped by adding 1.5 µl of 0.25 M EDTA. The products were then separated on 2% agarose gel to determine rearrangements at the targeted site.

Locomotion assays

At 72 hpf, the escape response test was performed to evaluate the swimming ability of the fish upon a slight touch stimulation. The motion of each larva was examined and scored as « normal swimming», « pause», « looping swimming», « pinwheel swimming», or « motionless» due to malformations. At 5 dpf, we analyzed spontaneous zebrafish motility using the Zebrabox® recording system (Viewpoint, Lissieu, France) equipped with infrared illumination for imaging in the dark. Locomotion was recorded for each larva on a 96-well plate for 30 min (15-min adaptation phase in the light followed by a 15-min phase in the dark) and presented as slow (3–6 mm/s) and high velocities (>6 mm/s) [32]. The velocity of the fish was tracked with the Viewpoint software, and experiments were performed a minimum of three times. The resulting data were pooled together for statistical analysis. Fisher's exact test or one-way ANOVA analysis was performed based on the data with Prism10.

Immunofluorescence

PTU (1-phenyl 2-thiourea—75 µM) treatment was used on 24hpf zebrafish to prevent pigmentation. At appropriate developmental stages, embryos were dechorionated and euthanized with 0.0168% tricaine (Sigma-Aldrich) and immediately fixed in 4% PFA for 1 h at RT. Permeabilization of larvae was achieved with 1X phosphate saline buffer (PBS), 0.5% Triton X-100, for 90 min at RT and subsequently in 1X PBS, 1%Triton X-100, for 2 h at RT on a slow shaker. Embryos were then incubated in blocking buffer (1% BSA in 1X PBS) for 1 h at RT and incubated in primary antibodies, mouse anti-synaptotagmin 2 (Znp-1, diluted 1:100 in blocking solution – DSHB, Iowa City, IA, United States), or mouse anti-islet 1 and 2 (39.4D5, diluted 1:100 in blocking solution – DSHB, Iowa City, IA, United States), overnight at 4 °C on a slow shaker. After 3 washes in 1X PBS, the embryos were incubated with a secondary antibody, Alexa Fluor™ 488 conjugated (diluted 1:500 in blocking solution, Invitrogen), overnight at 4 °C. Nuclei were stained with DAPI (diluted 1:8000, Sigma-Aldrich) for 5/10 min at RT. After washing in PBS, zebrafish larvae were mounted onto microscopic slides with Mowiol 4–88 (Sigma-Aldrich). Imaging was performed using LSM880 airyscan confocal microscope (Carl Zeiss). Evaluation of motor neurons' structure and hindbrain spinal cord projecting neurons' development was performed [33].

Morphological analyses

Images of 5dpf zebrafish were acquired with a Leica microscope (M165 FC) and Leica CMOS camera (IC90E, Leica Camera AG, Wetz-lar, Germany) for morphological inspection. Inter-ocular distance and head width were quantified using the Fiji software [33]. Fisher's exact test or one-way ANOVA analysis was performed with Prism10.

Staining of cartilaginous structure

At 5dpf, embryos were euthanized with 0.0168% tricaine (Sigma-Aldrich) and fixed overnight in 4% PFA at RT. Fixed embryos were washed four times with 1X PBS and 0.1% Tween-20 (PBST) and bleached with 30% hydrogen peroxide for 2 h at RT. After three wash cycles with PBST, specimens were transferred into an Alcian Blue solution (1% concentrated hydrochloric acid, 70% ethanol, 0.1% Alcian blue) and stained overnight at 4 °C. Embryos were rinsed a few times with acidic ethanol (5% concentrated hydrochloric acid, 70% ethanol, HCl-EtOH) and incubated in acidic ethanol for 20 min at RT on a slow shaker. Specimens were then re-hydrated as follows: (i) 5/10 min at RT in 1 mL of 75% HCl-EtOH / 25% H_2O_d; (ii) 5/10 min at RT in 1 mL of 50% HCl-EtOH / 50% H_2O_d; (iii) 5/10 min at RT in 1 mL of 25% HCl-EtOH / 75% H_2O_d; and (iv) 5/10 min at RT in 1 mL of 100% H_2O_d. Specimens were stored in 1 mL of 50% Glycerol and 50% (1%) KOH or kept in 100% Glycerol for extended storage. Stained embryos were positioned in 50% Glycerol and 50% (1%) KOH solution in a Petri dish. The head was photographed in a ventral–dorsal and a lateral view using a stereo microscope (Motic SMZ-171) with the Motic Image Plus software (version 3.0). *T*-test or one-way ANOVA analysis was performed with Prism10.

Overexpression analysis in zebrafish

Tagged human *AFF3* wild-type mRNA (GenBank: NM_002285.3) was cloned into pEZ-M13 vector [11]. The variants of interest, i.e., the two KINSSHIP variants

Val235Gly and Ala233Thr and the three newly-identified missense variants Gln179Glu, Lys528Arg and Thr594Ser, were engineered using the QuikChange II XL Site-Directed Mutagenesis Kit following the manufacturer's instructions (Agilent Technologies). Positive clones were confirmed by Sanger sequencing. *AFF3* mRNA was transcribed from the linearized vector pEZ-M13 + AFF3-FLAG Wt [11] or containing each of the studied variants using the mMESSAGE mMACHINE T7 transcription kit (Ambion) and purified using RNeasy Mini Kit (Qiagen) following the manufacturers' instructions. The injection mix consisted of the mRNAs at three different concentrations (180 ng, 360 ng, and 720 ng) diluted in RNAse-free water. One nanoliter of each diluted mRNA was injected inside the yolk, below the cell, in AB wild-type zebrafish embryos at the one- to two-cell stage. Distilled water was injected as vehicle control in a similar volume. Depending on RNA amounts, experiments were repeated twice or three times.

Phenotypic rescue in zebrafish

We engineered F0 zebrafish depleted for *aff3* by CRISPR/Cas9 genome editing and expressing the human *AFF3* mRNA of interest. The rescue experiment was conducted by evaluating the spontaneous zebrafish motility in the dark using the Zebrabox® recording system (Viewpoint, Lissieu, France). The injection mix consisted of sgRNAs/Cas9 complex and human *AFF3* mRNA Wt (for the phenotypic rescue) or the human AFF3 mRNA carrying each variant of interest. Different concentrations of human AFF3 mRNA Wt (25 ng, 50 ng, 75 ng, 100 ng, 150 ng, 200 ng) were tested to reach the phenotypic rescue. Phenotypic rescue by the variants was assessed by injecting 1 nl of mix containing sgRNAs/Cas9 complex + 75 ng (75 pg/µl) of each mRNA into one- or two-cell stage embryos. The sgRNAs were injected alone in mock-injected larvae with Cas9 replaced by the same volume of water. The GraphPad Prism software (version 10.0) was used to perform statistical analysis of the data. A Kruskal–Wallis test was adopted to determine differences between experimental groups. Experiments were performed seven times.

HEK293T isogenic cell lines

HEK293T cells were used to engineer *AFF3* knock-ins (KINSSHIP) and knockouts (LoF) cell lines by CRISPR/Cas9 genome editing. Four guides were used to create the LoF lines: one targeting the coding sequence of exon 6, designed with the Thermo Fisher Scientific tool, and three targeting exon 5 with the Gene Knockout Kit v2 of Synthego. To engineer the KINSSHIP lines, one sgRNA targeting the coding sequence in exon 6 designed using the Thermo Fisher Scientific tool, was combined with a DNA donor template to knock-in the Ala233Thr variant. The sgRNAs, DNA donor template, and corresponding sequencing primer pairs were ordered at Invitrogen, Synthego, or Sigma-Aldrich. The cocktail to induce AFF3 knockout contained 7.2 µg of the four combined sgRNAs and 36.2 µg of TrueCut™Cas9 v2 protein (Invitrogen). In the KINSSHIP model, 7.2 µg of the sgRNA by Thermo Fisher Scientific, combined with 36.2 µg of TrueCut™Cas9 v2 protein and 14.5 µg ds DNA donor, was used. Each mix was transfected using the Lipofectamine™ CRISPRMAX™ Cas9 Transfection Reagent Kit (Invitrogen) on 10-cm HEK293T cell plates according to the manufacturer's protocol. Forty-eight hours after transfection, cells were collected, resuspended post-counting, and diluted at a density of 8 cells/ml. One hundred microliters of this resuspension was transferred to each well of a 96-well plate. At the desired cell confluency, clones were screened with the QIAprep&amp CRISPR kit (QIAstock, QIAGEN, AG). Variants were confirmed by Sanger sequencing. Heterozygotes were further confirmed by cloning and sequencing of both alleles. We engineered five biallelic LoF HEK293T lines (LoF/LoF) with different combinations of variants (lines No.20 and 98: stop-gain/stop-gain; No.15: stop-gain/20 bp deletion; No.4: 4 bp deletion/114 bp deletion; No.216: 94 bp deletion/94 bp deletion), one heterozygous LoF stop-gain/+line (No.1), two homozygous Ala233Thr/Ala-233Thr KINSSHIP/KINSSHIP lines (No.54 and 90), and two compound heterozygous KINSSHIP and LoF lines (No.51 and 86: Ala233Thr/stop-gain). These ten lines and three unmutated HEK293T lines were grown simultaneously in biological triplicate before RNA extraction with RNeasy Mini Kit (QIAstock, QIAGEN AG). The nomenclature of the engineered variants is:

(A) *Stop-gain (through A insertion)*: GRCh37:2:100623265:A:AT, NM_002285.3:c.701dup, NP_002276.2:p.(Tyr234Ter) NC_000002.11:g.100623266dup
(B) *4 bp del*: GRCh37:2:100623262:CACAT:C, NM_002285.3:c.701_704del, NP_002276.2:p.(Tyr234Ter) NC_000002.11:g.100623265_100623268del
(C) *20 bp del*: GRCh37:2:100623247:GCCGTCCATTGGCCTCACATA:G, NM_002285.3:c.700_719del, NP_002276.2:p.(Tyr234ProfsTer) NC_000002.11:g.100623249_100623268del
(D) *94 bp del*: GRCh37:2:100623727:ACGAGGGCTGGTTCTGGGCTCTTGAATCTGCAACAAAATGTTCATCGATCTTGTTCACAGGAGTCTGAGGAACCCCAGGTTTGGGAACTCCAACG:A, NM_002285.3:c.276_369del, NP_002276.2:p.(Val93LeufsTer97) NC_000002.11:g.100623728_100623821del
(E) *114 bp del*: GRCh37:2:100623258:GCCTCACATACGCGGTCGGTTTCTGCTGGACCAGGC

TGGGTTTTGAAGCTAGGGATGGAGGAA AGTTCTGAACACAGTGTCCGCTGCTGC TGTGCTTGGCCGCCATGGCAGGTGGC:G, NM_002285.3:c.594_708del, NP_002276.2:p.(Arg198SerfsTer16) NC_000002.11:g.100623260_100623374del

(F) *Ala233Thr KINSSHIP variant*: GRCh37:2:100623268:CGC:GGT, NM_002285.3:c.697_699delinsACC, NP_002276.2:p.(Ala233Thr) NC_000002.11:g.100623268_100623270delinsGGT

Fibroblasts

Fibroblast cells from two patients' skin biopsies and three healthy age-matched control individuals (2–16 years of age) were grown simultaneously. At the desired cell confluency, RNA was extracted with RNeasy Mini Kit (QIAstock, QIAGEN AG).

Transcriptome profiling

RNA quality was assessed on a Fragment Analyzer (Agilent Technologies). The RNAs had RQNs between 9.0 and 10.0. RNA-seq libraries were prepared from 500 ng of total RNA with the Illumina TruSeq Stranded mRNA reagents (Illumina) using a unique dual indexing strategy, and following the official protocol automated on a Sciclone liquid handling robot (PerkinElmer). Libraries were quantified by a fluorometric method (QubIT, Life Technologies) and their quality assessed on a Fragment Analyzer (Agilent Technologies). Sequencing was performed on an Illumina NovaSeq 6000 for 100 cycles single read. Sequencing data were demultiplexed using the bcl2fastq2 Conversion Software (version 2.20, Illumina). We profiled transcriptomes with a minimum of 17.9 and 52.9 million mapped reads for HEK293T and fibroblasts, respectively. The HEK293T and fibroblast reads are deposited in GEO under accession GSE241621 and GSE246554, respectively. Raw reads were aligned to the human (hg38) genome using STAR (2.7.10b), the exact parameters are as follows: STAR [34] –runMode alignReads –twopassMode Basic –outSAMtype BAM SortedByCoordinate –outSAMattributes All –readFilesCommand "gzip -dc" –quantMode GeneCounts. Gene counts were generated using FeatureCounts [35] and differential expression analysis was performed with the DESeq2 (v.1.36.0) [36] package from Bioconductor (v3.15) [37]. Genes were considered differentially expressed based on an adjusted p-value cutoff of < 0.05. Pathway enrichment analysis was carried out using clusterProfiler (v.4.4.4) [38, 39] from Bioconductor using the enricher function. GSEA [40] analysis was carried out using the GSEA function in ClusterProfiler, and the following annotated gene sets from MSigDB v6.2 [41]: the Hallmark gene set [42]. For comparison with ChIP-seq studies in human HEK293T [17] and ES mouse cell lines [18], external sequencing data in bigWig format were acquired from GEO. UCSC bigWig files were created at 1 bp resolution and normalized to total alignable reads (reads-per-million). Peak detection was performed with MACS v3.0 [43, 44] using the bdgpeakcall function (with cut-offs 0.4 and 0.6 respectively). The AFF3 peak regions in mice were lifted over to the hg38 human genome assembly. The peak regions were annotated in R using the ChIPseeker [44] package, in particular the "annotate Peak" function.

Results

KINSSHIP probands

Through data aggregation, we identified four more KINSSHIP individuals (K22-K25). Affected individuals K22, K23, and K24 harbor the previously unreported de novo variants p.(Met238Thr), p.(Met238Val), and p.(Pro231Ser), respectively, whereas individual K25 carries the most commonly described p.(Ala233Thr) variant [11] (Fig. 1A; Table S1). Two more affected individuals (K26-K27) who carry the previously detected variants p.(Ala233Thr) and p.(Ala233Ser) [11], respectively, are mentioned in reference [45]. The genotypes and phenotypes of patients K22-K27 are described and compared in Table S1. Their variants fall within the nine-amino-acid-long 230-KPTAYVRPM-238 degron motif and further expand the number of its residues whose modification is associated with KINSSHIP (i.e. Pro231, Ala233, Val235, and Met238; numbering according to NM_002285.3 throughout) (Fig. 1A). Pathogenicity of the previously undescribed missense variants is supported by the 3D representation of the encoded degron peptide (Figure S1). Whereas changes at Pro231 were previously suggested to affect the backbone kink conferred by this conserved residue [11], the Met238 sidechain is pointing outward, forward-facing the Ser154 sidechain of a SIAH ubiquitin ligase loop. Modeling suggests that variants at this position should only slightly alter binding, predicting a less severe phenotype. Consistent with this hypothesis, probands K22 and K23 had phenotypes milder than that of typical KINSSHIP individuals with variants of the Pro231, Ala233, or Val235 residues that dock in the ubiquitin ligase binding pocket [11] (Table S1, Fig. 2B). Proband K22 presented with mild DD, mild speech impairment, facial dysmorphisms, skeletal malformations, mild hypertrichosis, and mild hypotonia and proband K23 with ID, autism, obesity, short but proportionate stature, some dysmorphisms, mild scoliosis, obstructive sleep apnea, and hypotonia (Table S1).

Duplication proband

Data aggregation also enabled ascertainment of an individual with a KINSSHIP-like phenotype carrying a de novo partial duplication of *AFF3* further strengthening the hypothesis that an increased level of AFF3 is pathological. This DUP1 proband presented with severe failure to thrive with postnatal onset, severe DD with poor eye contact, poor head control, inability to sit and speak, epilepsy, corpus callosum hypoplasia, facial dysmorphism, hypertrichosis, hypotonia, hip, knee, ankle and wrist flexion contractures, and severe scoliosis (Table S1). Whole genome sequencing revealed a tandem duplication of the interval encompassing exon 10 to exon 24 of *AFF3* encoding part of the ALF domain, the TAD, NLS, and CHD domains, and exon 1 to 3 (up to intron 3–4) of the same orientation ubiquitous *REV1* (chr2:g.100,077,649_100,359,928dup (hg19), NC_000002.11:g.100,077,649_100,359,928dup) (Figs. 1A,B and 2B, Figure S2). The expression of the partially duplicated copy of *AFF3* is then under the control of the *REV1* promoter, which could result in the expression of a degron-less AFF3, a hypothesis that we could not further test directly due to lack of available sample.

Heterozygous LoF and biallelic probands

To further challenge the hypothesis that diminished expression of *AFF3* is deleterious, we searched for individuals with loss-of-function (LoF) variants in *AFF3*. Using data aggregation of multiple laboratories and clinical centers, e.g., GeneMatcher [46] and DECIPHER [47], we identified ten affected individuals with monoallelic (individuals L1-L9 and L14) and three (L11-L13) with biallelic *AFF3* truncating variants, as well as a proband compound heterozygous for a LoF and a rare missense variant (L10; Figs. 1A, C and 2A, B; Table S2). Of note, one of the affected individuals described in reference [24] was a compound heterozygote for a CGG expansion and a deletion of the *AFF3* promoter. The identified truncating variants are either microdeletions (L5, L7-L8) or frameshifts (L1-L4, L6, L9-L14) not described in GnomAD (v4.0.0) (Fig. 1A, Table S2). Consistent with the deleteriousness of diminished or absence of *AFF3* expression, these fourteen individuals (10 males and 4 females) shared common phenotypes such as global DD/ID (11 out of 11), abnormal corpus callosum (4/6), speech impairment (10/11), muscle disorders/hypotonia (7/9), facial dysmorphisms (6/7), mild cranial dysmorphisms (3/8), and skeletal defects (4/7). All symptoms are reported in Table S2 and summarized in Fig. 2A. The siblings L12 and L13, who are homozygous for a truncating variant, and L10, who is compound heterozygote for a LoF variant and missense p.(Gln1020Arg) present a more severe phenotype than their parents who are heterozygotes for the LoF variant (families 6 and 8; Figs. 1A, C and 2A and B; Table S2). Our search also identified four affected individuals with biallelic homozygous (B1-B3) or compound heterozygous (B7) missense variants in *AFF3* (Figs. 1A, C and 2A; Table S2). A consanguineous family with three affected individuals was previously described in reference [48] (B4-B6). Contrary to the KINSSHIP missense variants that map to the degron, these missense variants modify either the CHD domain or the interval between the TAD and the NLS domains (Fig. 1A). They are rare or not described in GnomAD and/or affect the expression of *AFF3* (see below and Table S2). They present overlapping symptoms such as DD/ID (6/7) and ADHD (attention deficit hyperactivity disorder) (2/3), epileptic encephalopathy or abnormal sleep EEG (electroencephalography) (3/4), short/no attention span (3/4), speech impairment (3/4), heart defects (2/4), and vision impairment (2/4), and other defects detailed in Table S2 and summarized in Fig. 2A. These two cohorts showed a milder phenotype than KINSSHIP probands, suggesting they might represent a new syndrome.

In silico modeling of most of the identified missense variants is hampered by the lack of reliable AFF structural information with the exception of the CHD that is important for dimerization and the ALF that contains the degron and the ELL-binding domains (ELLbow, see below) [49, 50]. The p.(Gln1020Arg), p.(Val1036Ile), p.(Arg1186Gln), and p.(Gly1215Val) variants fall within the CHD (Fig. 1A). A bulky sidechain at position 1215 will collide with Leu1063 and/or Leu1192. Likewise, Val1036 is optimally surrounded by the hydrophobic sidechains of Leu1068, Leu1071, Tyr1072, and Met1075, and cannot accommodate the bulkier p.(Val1036Ile) variant without affecting local packing. Gln1020 and Arg1186 are located at the domain surface, where changes in the local charge might affect binding specificity.

Lastly, it is possible that some missense variants outside of the degron are linked to an autosomal dominant disease, as we identified an individual carrying a de novo p.(Ala886Thr) variant presenting with DD, speech impairment, and ASD (autism spectrum disorder) symptoms and as an individual with a de novo p.(Leu312Phe) variant presenting with DD was described in reference [45] (M1-M2, Figs. 1A and 2A, Table S2). While the p.(Ala886Thr) variant, which maps just carboxy-terminally from the NLS (Fig. 1A), cannot be 3D modeled, the p.(Leu312Phe) variant maps to the ELLbow and its models suggest that four of the five possible Phe rotamers will severely clash with either the AFF3 Phe329 or the ELL2 His618 residue.

Animal models

To further assess if the diminished expression of *AFF3* was deleterious to organismal phenotypes we knocked down (KD) *aff3*, the one-to-one zebrafish ortholog of *AFF3*, using CRISPR-Cas9 genome editing. We used two single-guide RNAs targeting exon 6 each providing more than 90% efficiency. At 5 days post fertilization (dpf), we observed malformations in 10% of KD larvae, including incomplete eye pigmentation, altered head structure, lateral belly edema, pericardial edema, and skeletomuscular dysmorphology (Fig. 3A). Staining of the cartilaginous cranial structure revealed malformations in 75% of KD larvae (Fig. 3B). Inter-ocular distance (IOD) and head width (HW) were significantly decreased in *aff3* KD compared to uninjected (Un) ($p=0.011$ IOD and $p=0.001$ HW) and mock (M) injected larvae ($p=0.041$ IOD; $p=0.004$ HW) (Fig. 3C–E). The escape response test upon a tactile stimulus performed at 3 dpf showed that while none of the mock-injected zebrafish showed perturbed escape responses, about a third of the *aff3* KD larvae were affected ($p<0.0001$). The majority presented either looping (22.5%) or pinwheel swimming (4.7%), behaviors linked to neurological and mechano-sensory system impairment [32, 51, 52] and 5.6% were motionless due to extensive malformations (Fig. 3F). At 5dpf, the locomotion ability was quantitatively evaluated with an automated tracking device. The *aff3* KD larvae showed a statistically significant decrease in global swimming velocity in the dark compared to Un ($p<0.0001$) and M larvae ($p=0.0025$) (Fig. 3G). As such hypo-locomotion is often associated with neuromotor deficits and akinesia [51], we immunostained *aff3* KD larvae hindbrain and motoneurons. Hindbrain Mauthner cells [52] presented a general developmental delay, and while normal in growth and architecture, motoneurons were disorganized in deformed larvae (Figure S3). Our zebrafish and previously published mouse results [11, 53] support the causativeness of *AFF3* LoF variants.

Assessing variants

We previously showed that overexpression in zebrafish embryos of human *AFF3* leads to a dose-dependent increase of developmental anomalies [11], a phenotype that was further exacerbated upon overexpression of the p.(Ala233Thr) KINSSHIP isoform that are predicted to be more resistant to degradation [12]. To assess the pathogenicity of the missense variants identified in the biallelic individuals, we injected zebrafish with human *AFF3* mRNA wild-type (Wt), two selected missense variants present in homozygous state in probands B1 and B2 and his affected sister B3 and mapping outside of crystalized domains (Lys528Arg and Thr594Ser), two KINSSHIP variants (Ala233Thr and Val235Gly) and as control Gln179Glu (Chr2 (GRCh37) g.100623432:G>C, c.535C>G), a variant not described in GnomAD, which we identified in homozygosity in a healthy individual. The resulting 5dpf larvae were cataloged as described [12]: (i) normal phenotype, (ii) Class 1 with skeletomuscular dysmorphology and small dimension, (iii) Class 2 with a more severe phenotype including at least three of skeletomuscular dysmorphology, small dimensions, head malformations, eyes' alteration, pericardial edema, and lateral belly edema, or (iv) deceased. Consistent with previously published observations, accumulation of *AFF3* Wt mRNA significantly increased the number of larvae with debilitating traits ($p=0.0002$). Compared to *AFF3* Wt mRNA accumulation, both Ala233Thr and Val235Gly isoforms led to a further significant increase in the number of malformed larvae and mortality at all doses ($p<0.0001$) (Fig. 3H). Overexpression of the two missense variants identified in proband B1 and B2 (and his affected sister B3) similarly caused higher malformations and mortality rates than overexpression of AFF3

(See figure on next page.)

Fig. 3 *aff3* knocked down zebrafish larvae display altered behavior and morphological anomalies. The conditions analyzed are the following: Uninjected (Un), Mock-injected (M), and *aff3* knockdown (KD). **A** Proportions of normal and developmentally defective 5 dpf embryos. In 10% of *aff3* KD zebrafish, we identified several morphological anomalies such as head malformations, belly and heart edema, skeleton-muscular dysmorphologies, and alteration of eye pigmentation. **B** Alcian blue staining at 5 dpf revealed jaw malformation in 57% of *aff3* KD zebrafish. **C** Visualization of morphological inter-ocular distance (IOD) and head width (HW) measurements from dorsoventral zebrafish image. Quantification of IOD (**D**) and HW (**E**) indicates a significative decrease in IOD and HW in *aff3* KD larvae; $p^{*}<0.04$; $p^{**}<0.0049$. **F** Touch test response assay at 3 dpf. Upon a touch stimulus, we classified the larvae swimming behavior in «normal swimming», «pause», «looping swimming», «pinwheel swimming», or «motionless» due to malformations. **G** Swimming global velocity analysis at 5dpf in the dark of Un, M, and *aff3* KD and *aff3* KD co-injected with human *AFF3* (*hAFF3*) mRNA wild-type (Wt), which recovered 57% of the locomotion function compared to *aff3* KD, or harboring the indicated missense variant, i.e., the KINSSHIP variants Ala233Thr and Val235Gly or the biallelic variants identified in this report in a healthy (Gln179Glu) or affected individuals (Lys528Arg and Thr592Ser). **H** Proportions of normal and developmentally defective 5 dpf embryos uninjected (Un), injected with water as control (H_20) or with 360 ng of h*AFF3* mRNA Wt or the indicated missense variant. Larvae were cataloged as described: (i) normal phenotype, (ii) Class 1 with skeletomuscular dysmorphology and/or small dimension, (iii) Class 2 with a more severe phenotype including at least three of the following characteristics: skeletomuscular dysmorphology, small dimensions, head malformations, eyes' alteration, pericardial edema, and lateral belly edema or (iv) dead. Injections of 180 and 720 ng of h*AFF3* mRNA showed similar results

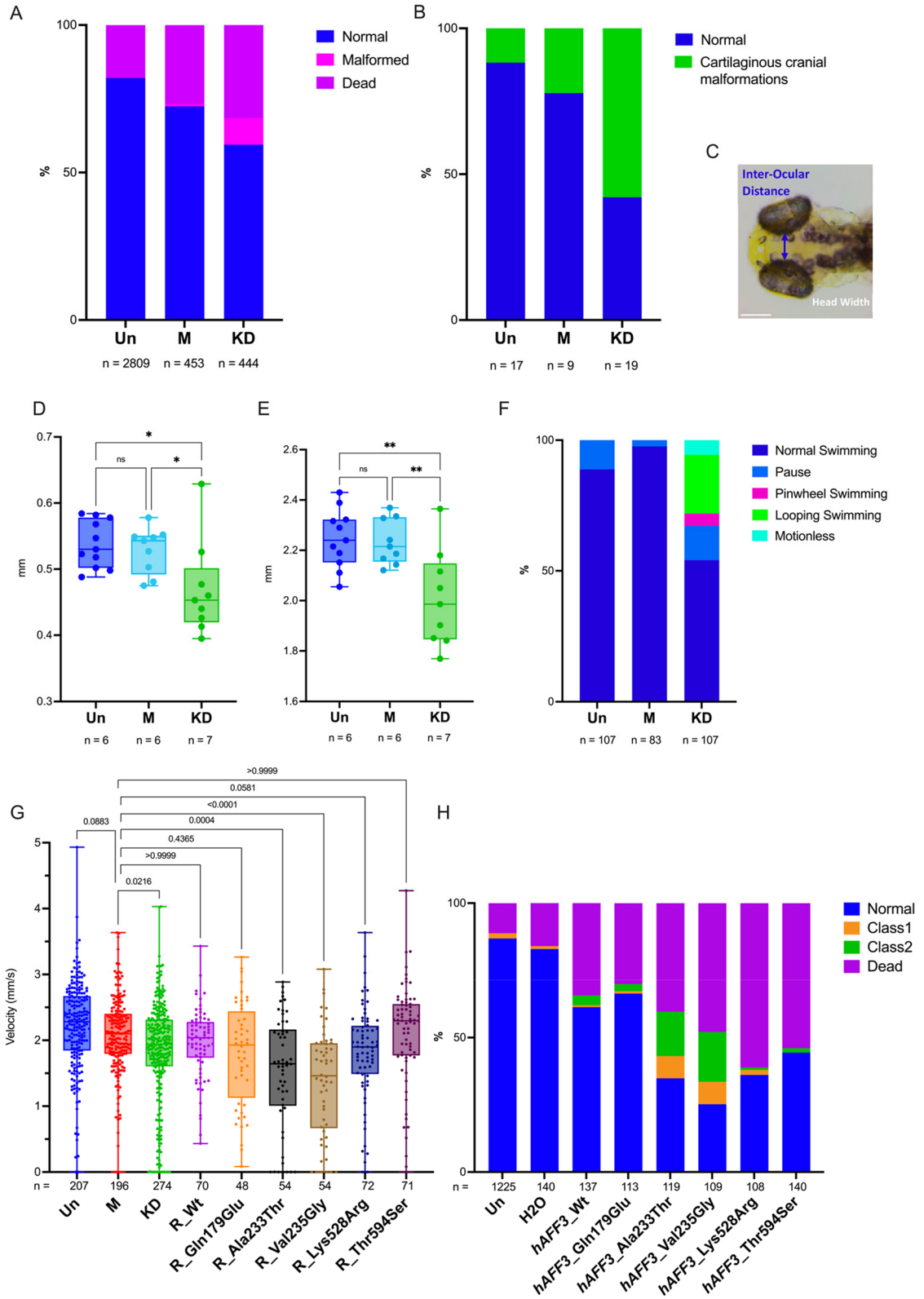


Fig. 3 (See legend on previous page.)

Wt (Lys528Arg $p<0.0001$ and Thr594Ser $p=0.0018$) albeit not at the rate of the KINSSHIP variants. On the contrary, the control variant p.(Gln179Glu) had an effect similar to that of Wt overexpression ($p=0.7$; Fig. 3H).

These results suggest that like truncating variants, at least some of the missense variants identified in the affected individuals could be causative. To challenge this hypothesis further, we performed a phenotypic rescue experiment [54]. As described above, depletion of *aff3* in 5 dpf zebrafish larvae resulted in decreased swimming velocity. That decrease could be rescued by co-injection of human *AFF3* Wt mRNA demonstrating first that human AFF3 can compensate for the loss of its zebrafish ortholog and second that ablation of *aff3* activity was causative of the phenotype (Fig. 3G). Consistent with the detrimental effect of the overexpression of the predicted to be more resistant to degradation KINSSHIP variants, we observed an even lower average velocity upon co-injection of Val235Gly and Ala233Thr mRNAs (both $p<0.001$ compared to injection of Wt; and respectively $p<0.0001$ and $p=0.004$ compared to M), while co-injection of Lys528Arg mRNA partially rescued *aff3* ablation ($p=0.0581$). Co-injection of Thr594Ser mRNA resulted in increased activity with injected larvae presenting an increased velocity compared to mock (Fig. 3G). The control variant p.(Gln179Glu) had an intermediate effect halfway between Wt and Lys528Arg mRNA injections. Together these results suggest that these missense variants impact the activity of AFF3 and that biallelic *AFF3* variants could be associated with an autosomal recessive disease. Consistent with the latter hypothesis, neither homozygous nor compound heterozygous classified as "weak missense variant or worse," i.e., with a MAF ≤ 1% and REVEL score ≥ 0.644 [55], were identified in GnomAD v2.1.1.

Transcriptome profiling

To compare the transcriptional consequences of AFF3 loss and overexpression, we used CRISPR-Cas9 genome editing to engineer multiple KINSSHIP and LoF variants in an isogenic cell model, the human embryonic kidney 293 T line. HEK293T was chosen (i) as KINSSHIP individuals often present with a horseshoe kidney [54], (ii) as *AFF3* is expressed in this cell line, and (iii) as in this line both transcriptome profiles of *AFFs* shRNAs knockdowns [3], and (iv) ChiP-seq of AFF3 have been published [17]. We engineered five biallelic LoF HEK293T lines (LoF/LoF) with different combinations of variants (lines No.20 and 98: stop-gain/stop-gain; No.15: stop-gain/20 bp deletion; No.4: 4 bp deletion/114 bp deletion; No.216: 94 bp deletion/94 bp deletion), one heterozygous LoF stop-gain/+line (No.1), two homozygous Ala233Thr/Ala233Thr KINSSHIP/KINSSHIP lines (No.54 and 90), and two compound heterozygous KINSSHIP and LoF lines (No.51 and 86: Ala233Thr KINSSHIP/stop-gain). We profiled the transcriptomes of three biological replicates of each of these lines by RNA-sequencing and compared them to those of three biological replicates of three wild-type lines (Wt1, Wt2, and Wt4), for a total of 39 profiles. While *AFF3* mRNA levels are significantly decreased in the five biallelic LoF/LoF lines (*padj* = 3.5E-53), the Ala233Thr/Ala233Thr KINSSHIP/KINSSHIP and the LoF/+lines present *AFF3* transcript levels comparable and intermediate (*padj* = 0.011) to that found in control Wt lines, respectively (Figure S4). We first compared the transcriptome of homozygous LoF/LoF and KINSSHIP/KINSSHIP lines to that of +/+lines and identified 3553 and 4177 differentially expressed genes (DEG) at an adjusted *p*-value threshold of 0.05, respectively (Fig. 4A and B, Table S3-S4). We observed an overlap of 23% of DEGs with previous transcriptome profiling of HEK293T cells in which *AFF3* was knocked down with shRNAs [3].

The LoF/LoF and KINSSHIP/KINSSHIP lines present with significant repression of genes involved in the G2-M transition, oxidative phosphorylation, targets of E2F and MYC-related genes, most markedly in the KINSSHIP/KINSSHIP model. Both lines showed an upregulation

(See figure on next page.)

Fig. 4 Transcriptome profiles of engineered isogenic HEK293T cells. **A** Four-way Venn diagram of differentially expressed genes (DEGs) in biallelic loss-of function (LoF/LoF) *AFF3* lines and biallelic KINSSHIP/KINSSHIP (DN (dominant negative)/DN) *AFF3* lines upon comparison with unmutated wild-type lines. DEG counts are stratified in genes up- (UP) and downregulated (DOWN). **B** Volcano plots of DEGs in biallelic loss-of function (LoF/LoF) *AFF3* lines (left panel) and biallelic KINSSHIP/KINSSHIP (DN/DN) AFF3 KINSSHIP lines (right panel) upon comparison with unmutated wild-type lines. The top 30 most significant DEGs in LoF/LoF that are dysregulated in an opposite manner in KINSSHIP/KINSSHIP (DN/DN) are indicated, together with some of the most differentially expressed genes (-log10(Padj) > 20 and abs(log2FoldChange) > 0.5). **C** Four-way Venn diagram of differentially expressed genes (DEGs) in biallelic loss-of function (LoF/LoF) *AFF3* lines and biallelic KINSSHIP/KINSSHIP (DN/DN) *AFF3* KINSSHIP lines upon comparison with unmutated wild-type lines and AFF3 ChIP-seq peaks identified in HEK293T cells (HEK293T) and in *Mus musculus* ES cells (mmES). DEGs bound by AFF3 discussed in the text are indicated. **D** Gene set enrichment analysis (GSEA) for hallmark pathways of DEGs in biallelic loss-of function (LoF/LoF) *AFF3* lines (left panel) and biallelic dominant-negative KINSSHIP/KINSSHIP (DN/DN) *AFF3* KINSSHIP lines (right panel) upon comparison with unmutated wild-type lines. **E** Examples of DEGs *NRC31* (top) and *DDX17* (bottom) loci bound by AFF3. UCSC genome browser snapshot showing from to top to bottom AFF3 ChIP-seq HEK293T results, UCSC and REFSeq curated gene structure and vertebrate PhyloP conservation scores (left panels). Expression level of *NRC31* (top) and *DDX17* (bottom) in +/+(blue), LoF/LoF (yellow), and KINSSHIP/KINSSHIP (DN/DN; green) HEK293T engineered lines (right panels)

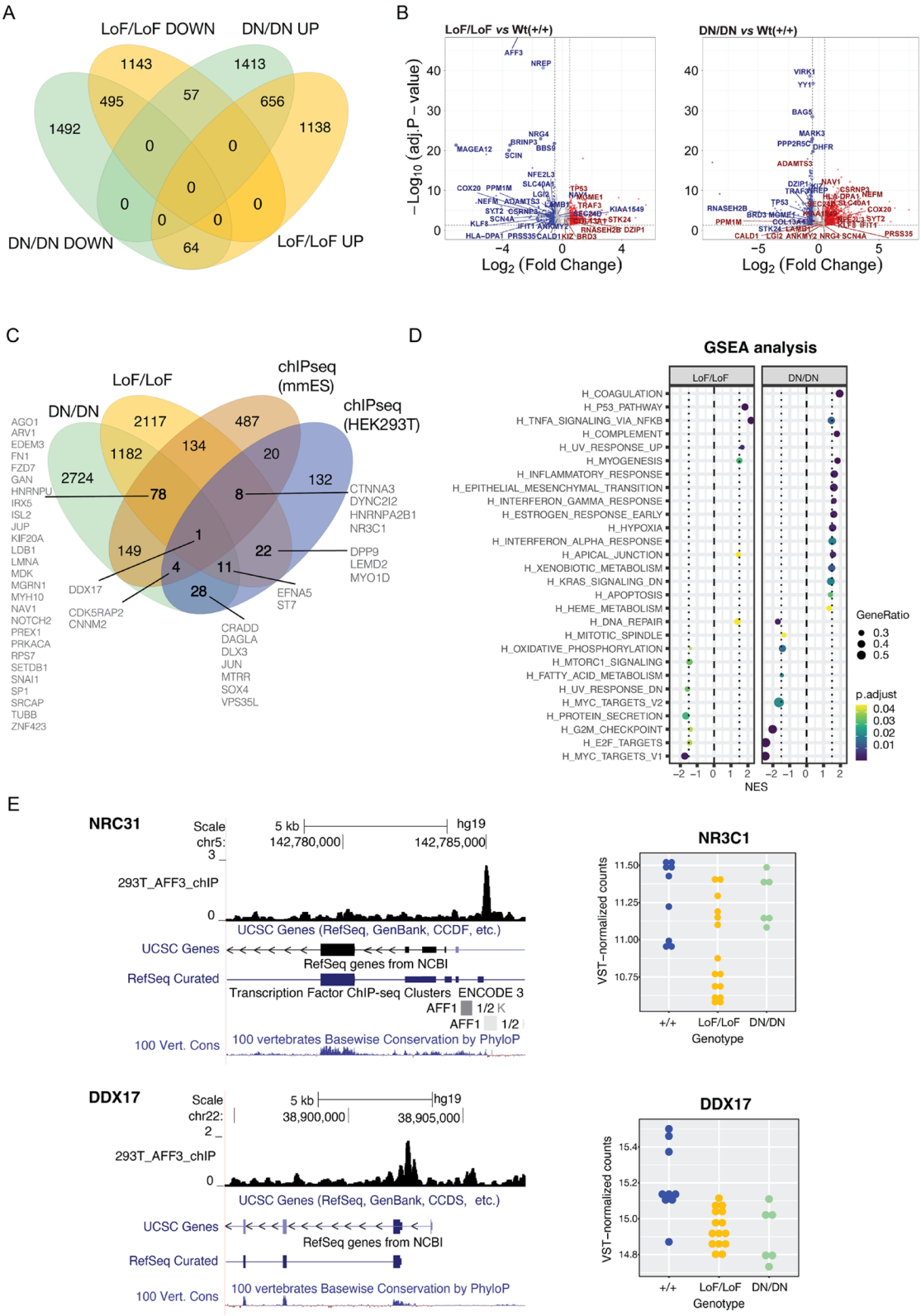


Fig. 4 (See legend on previous page.)

of the inflammatory response (e.g., TNFA-signaling via NFKB) and pathways important for myogenesis and apical junction (Fig. 4B,C). Only about a third of the DEGs are common to both datasets suggesting that *AFF3* LoF and KINSSHIP variants largely modulate transcriptomes differently (Fig. 4A,B). For example, pathways involved in the epithelium-mesenchyme transition, early response to estrogen, hypoxia, xenobiotic metabolism, and apoptosis, as well as genes that are downregulated by KRAS are specifically upregulated in the KINSSHIP/KINSSHIP lines (Fig. 4B,C). Within the set of 1272 common DEGs, 121 genes present opposite effects in both strains, i.e., they are upregulated in one genotype and downregulated in the other (Fig. 4A). They are enriched for DNA repair genes, a pathway activated in LoF/LoF but not in KINSSHIP/KINSSHIP cells (Fig. 4C). A core set of 20 DEGs are similarly modified upon *AFF2*, *AFF3*, and *AFF4* knockdown [3] or when *AFF3* harbors homozygous KINSSHIP or LoF variants suggesting that they are sensitive to any SECs' perturbation.

We then assessed if DEGs were direct or indirect targets of AFF3. While in excess of 3500 genes are dysregulated in each genotype, only 226 genes presented with a neighboring AFF3 ChIP-seq peak using a FDR of 0.05% [17] (Table S5), suggesting that many of the observed transcriptome changes are downstream effects. However, 32% (74 out of 226) of the bound loci were dysregulated in either the LoF/LoF and/or the KINSSHIP/KINSSHIP lines (Fig. 4D). The binding sites of the orthologous mouse Aff3 were determined in ES cells by ChiP-seq [18]. Upon lifting Aff3 ChIP-seq peaks to the human genome, we similarly found that 42% of genes with a binding site (374 out of 881) were DEGs in either the LoF/LoF and/or the KINSSHIP/KINSSHIP lines (Fig. 4D; Table S6). While such inter-clade binding comparisons have caveats, our HEK293T and ES results suggest that a substantial proportion of bound loci are dysregulated upon changes in the expression level of *AFF3* and/or stability of AFF3 (Fig. 4D,E; Figure S5, Table S7). While we cannot reject the null hypothesis of independence between being differentially expressed and being bound by AFF3 in mouse ES cells ($p=0.1422$; Fisher's exact test), in HEK293 these two categorical variables are significantly related ($p=0.0002013$). The latter result demonstrates the biologically relevance of the identified DEGs. These dysregulated direct targets include genes associated with traits present in *AFF3* variants carriers such as neurodevelopmental disorders (e.g., *AGO1, ARV1, CDK5RAP2, CNNM2, CRADD, DPP9, EDEM3, GAN, HNRNPA2B1, HNRNPU, IRX5, MGRN1, MTRR, PREPL, SOX4*, *SRCAP, TOR1A TUBB, VPS35L*) and autism (i.e., part of SFARI gene list, e.g., *CDK5RAP2, CTNNA3, DAGLA, DLX3, LDB1, MYH10, PREX1, PRKACA, SETDB1, SRCAP,* and *ST7*), ossification and limb defects (*DLX3*, *DPP9*, *DYNC2I2, FN1, IRX5, RPS7SRCAP*, *VPS35L*), pilosity abnormalities (*DLX3*, *GAN, JUP, NR3C1*), renal diseases (*CNNM2, RPS7, ZNF423*), cardiac disorders (*CTNNA3, JUP, KIF20A, LMNA*, *VPS35L*), and dysmorphisms (*LEMD2*, *IRX5, LMNA, RPS7, SOX4*). They also comprise key genes implicated in axon guidance, cell migration, and cell fate (e.g., *DDX17*, *EFNA5, FZD7*, *GAN, ISL2, JUN, MDK, MYO1D, NAV1, NOTCH2, SNAI1*, and *SP1*). Importantly, some direct targets are upregulated, while others are downregulated. For example, *DDX17* is downregulated in both LoF/LoF and KINSSHIP/KINSSHIP lines, whereas *CTNNA3* and *NR3C1* are only downregulated in LoF/LoF lines and *CDK5RAP2* only in KINSSHIP/KINSSHIP lines. On the contrary, *CNNM2* is upregulated in KINSSHIP/KINSSHIP lines (Fig. 4E, Figure S5).

As many of our affected individuals present heterozygous LoF variants, we then compared the transcriptome profiles of the LoF/+ lines with that of the LoF/LoF lines and observed that, while the same pathways are affected (Figure S6), only 22% of the DEGs of the homozygous line were also dysregulated in the heterozygous line, suggesting a dose-dependent modification (Table S8). We similarly compared the transcriptomes of the compound heterozygote KINSSHIP/LoF lines to those of both the LoF/LoF and KINSSHIP/KINSSHIP lines. While 37% of the DEGs common to LoF/LoF and KINSSHIP/KINSSHIP lines are modified in the KINSSHIP/LoF lines, we also observe in this compound heterozygous line modifications that are specific to one or the other of these homozygous line, i.e. 22 and 27% of the DEGs specifically modified in each group of lines, respectively (Table S9). This suggests a co-dominance of the Ala233Thr and stop-gain variants where the increased stability of the first allele does partially compensate the decreased expression level of the second allele in some instances and over-compensate in others.

In parallel, we compared the transcriptome of primary fibroblasts from two probands with biallelic missense alterations of *AFF3* (B1: homozygous p.(Lys528Arg); B7: compound heterozygote p.(Val1036Ile)/p.(Arg1186Gln (Fig. 1A; Table S2) to those of three healthy controls by RNA-sequencing. We found 142 DEGs at an adjp-value threshold of 0.05 (Table S10). *AFF3* mRNA levels are significantly decreased in both probands ($p<0.0002$; Table S10) and a comparable number of distinctive reads corresponding to both alleles of the compound heterozygote were identified (Table S11), which is consistent with the notion that the three *AFF3* missense variants harbored by these probands are deleterious. While only 19 and 16% of the fibroblasts DEGs are also DEGs in the LoF/LoF and KINSSHIP/KINSSHIP HEK293T lines, respectively,

the same hallmark pathways are dysregulated. For example, genes involved in the G2-M transition, targets of E2F and MYC-related genes and interferon alpha-response are enriched within their list of respective DEGs (Table S11, Figure S6). These results suggest that similar pathological mechanisms are at play when AFF3 is haploinsufficient and when it harbors biallelic missense variants.

Discussion

We present evidence suggesting that multiple *AFF3* variant-specific mechanisms are associated with cognitive impairment. While repeat expansion in the promoter and de novo dominant-negative variants in the degron of this gene were previously linked to mild ID [23, 24] and KINSSHIP syndrome [11], respectively, we show that duplication, truncation, deletion, ablation, and biallelic variants in *AFF3* are also associated with ID. The mode of inheritance and associated phenotypes are summarized in Fig. 2A. Our zebrafish, mouse and cellular results summarized in Fig. 2B support these hypotheses. The orthologous mouse knockouts, *Aff3*$^{+/-}$ and *Aff3*$^{-/-}$ C57BL/6N, exhibited skeletal defects, an abnormal skull shape, kidney defects, and neurological dysfunction [11, 53]. Homozygous *Aff3*$^{-/-}$ also presented with significantly enlarged lateral ventricles and decreased corpus callosum size [11] (Fig. 2B). Suggestive of semi-dominance, homozygous LoF (L12 and L13) and compound heterozygous LoF/missense (L10) individuals present more severe phenotypes than their heterozygous parents (Figs. 1C and 2A). The hypothesis that non-degron de novo missense variants are also linked to DD/ID warrants further investigation and the identification of more affected individuals (Fig. 2A). Common variants in this locus were similarly GWAS- or MTAG-associated (multi-trait analysis of GWAS) with cognition proxies such as fluid intelligence, educational attainment, and mathematical ability, or with correlated traits such as household income, occupational attainment, and brain morphology [56]. Consistent with these findings, *AFF3* and its macaque, mouse, rat, rabbit, and chicken orthologs are expressed during the early stages of brain and cerebellum development in particular in late neurons [57, 58], where it plays a direct role in the migration of cortical neurons [59]. Likewise, common variants in this locus are associated with scoliosis, anthropometric traits (BMI, height), and pulmonary involvement (vital capacity, asthma, chronic obstructive pulmonary disease), three cardinal features of KINSSHIP syndrome. *AFF3* is also GWAS-/MTAG-associated with diabetes (type 1, type 2, diabetic nephropathy, and HDL cholesterol), addictions (smoking initiation, alcohol consumption, cannabis dependence, television watching), autoimmunity (lupus, celiac disease, rheumatoid and juvenile idiopathic arthritis), sexual development and dimorphism (age at menarche, endometriosis, mammographic density, male baldness, biological sex), blood measurements (e.g., hematocrit, hemoglobin measurement), eye diseases (e.g., astigmatism, intraocular pressure, corneal measurements), and insomnia [56]. This high pleiotropy is consistent with the large and diversified transcriptional role of AFF3. It suggests that any perturbation of its expression level might be deleterious. Consistent with this hypothesis we identify multiple modes-of-action and observe variant-specific/expression level modulation of the phenotype. Firstly, untimely (over)expression of KINSSHIP variants that are less sensitive to SIAH regulation leads to extremely severe phenotypes in human, zebrafish, and rodents, e.g., homozygous Ala233Thr knock-in and homozygous ablation of *Aff3* exon 4–13 mimicking the original KINSSHIP deletion of 500 kb lead to mouse lethality [11] (Fig. 2B). Secondly, C57BL/6N and CD1 genetic backgrounds modulate the phenotypes presented by *Aff3* mouse knockouts [11, 14, 15]. Thirdly, knockdown and overexpression of mouse *Aff3* in dermal cells impair niche switching, which is required for hair reconstitution [10]. Fourthly, *AFF3* overexpression in HeLa cells perturbed the dynamics of the nuclear speckles [1]. Our RNA-seq experiments further demonstrate that changes in the amount and/or function of AFF3 dramatically alter transcriptome profiles (Fig. 2B). We show that the expression of about one third of the AFF3 targets (bound loci) are differentially expressed upon *AFF3* modification and observe a progression in the extent of transcriptome alterations with those linked to haploinsufficiency being less drastic than that of homozygous LoF cells, which in turn are less impacted than cells harboring homozygous KINSSHIP variants.

Conclusions

In conclusion, we are adding to a growing list of variant-specific neurodevelopmental mechanisms and their associated genotype–phenotype correlations [60–63] (Fig. 2A). We demonstrate that beside degron variants that impair the degradation of the encoded protein [11] and downregulation due to promoter hypermethylation [23, 24], dysregulation of *AFF3* through gene duplication, heterozygous and biallelic truncating variants, biallelic missense variants, and compound heterozygous truncating/missense variants are associated with cognitive impairment.

Abbreviations

ID	Intellectual disabilities
DD	Developmental delay
DN	Dominant-negative
LoF	Loss-of-function
+	Wild-type allele

NHD	N-terminal homology domain
CHD	C-terminal homology domains
ALF	AF4-LAF4-FMR2 domain
TAD	Transactivation domain
NLS	Nuclear/nucleolar localization sequence
SECs	Transcriptional super elongation complexes
P-TEF	Positive transcription elongation factor
MLLT	Myeloid/lymphoid or mixed-lineage leukemia; translocated to
ELL	Elongation Factor for RNA Polymerase II
ELLbow	ELL-binding domain
pLI	Probability of being loss-of-function intolerant
pLOEUF	Loss-of-function observed/expected upper bound fraction
PTU	1-Phenyl 2-thiourea
K22-K26	KINSSHIP patients 22 to 26
DUP	Patient with partial duplication of *AFF3*
L1-L14	Patients 1 to 14 with loss-of-function variants
B1-B7	Patients 1 to 7 with biallelic missense variants
ADHD	Attention deficit hyperactivity disorder
EEG	Electroencephalography
ASD	Autism spectrum disorder
KD	Knockdown
Dpf	Days post fertilization
IOD	Inter-ocular distance
HW	Head width
Un	Uninjected
M	Mock
Wt	Wild-type
MAF	Minor allele frequency
DEG	Differentially expressed genes
MTAG-associated	Associated by multi-trait analysis of GWAS
BMI	Body mass index

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13073-024-01339-y.

Additional file 1: Figure S1. 3D protein modelling of the human AFF3 degron region bound to SIAH ubiquitin ligase. **Figure S2**. Sequencing read profiles of the DUP1 individual. **Figure S3**. Immunostaining of hindbrain neurons and motoneurons in 3dpf zebrafish. **Figure S4**. AFF3 expression levels in engineered isogenic HEK293T cells. **Figure S5**. Examples of differential expressed genes (DEGs) loci bound by AFF3. **Figure S6**. Gene Set Enrichment Analysis for hallmark pathways of DEGs in biallelic loss-of function (LoF/LoF) *AFF3* lines and heterozygote loss-of function (Lof/+) *AFF3* lines upon comparison with unmutated wildtype lines.

Additional file 2. In Table S1 and Table S2 are described the genotypes and phenotypes of KINSSHIP individuals and of other carriers of AFF3 variants, respectively.

Additional file 3. This Excel file contains nine supplementary tables cataloging transcriptomic (S3-S4 and S7-S10) and Chromatin-IP (S5-S6) results and their intersection (S7) **Table S3**. DEGs (Differentially Expressed Genes) in biallelic loss-of-function (LoF/LoF) AFF3 HEK293T lines. **Table S4**. DEGs in homozygous dominant negative (DN/DN) KINSSHIP AFF3 HEK293T lines. **Table S5**. AFF3 ChIP-seq peaks in HEK293T cells. **Table S6**. AFF3 ChIP-seq peaks in *M. musculus* ES cells. **Table S7**. DEGs and AFF3 ChIP-seq peaks intersections. **Table S8**. DEGs in heterozygous loss-of-function (LoF/+) AFF3 HEK293T lines. **Table S9**. DEGs in compound heterozygote dominant negative/ loss-of-function (DN/LoF) AFF3 HEK293T lines. **Table S10**. DEGs in patients' fibroblasts. **Table S11**. number of AFF3 reads at different location (in patients' fibroblasts).

Acknowledgements
We thank Jacques Beckmann for comments.

Authors' contributions
SB, JC, NV, and BY engineered and phenotyped animal and cell models. GA, FS, and CI analyzed transcriptomes profiles. NG 3D-modeled missense variants. AB, FS, LT, SS, RAJ, J-US, DD, PB-T, GRN, KNW, LD, MM, CG, LELMV, RP, RK, HY, GÅMH, CJ, MFS, KMB, MJL, CMBC, CZ, JRL, LP, LF-G, RM-T, FP, AT, HZE, LM, MK, OK, JH, MS, MI, FO, FZ, KW, AM, MKH, PP, HA, AJAK, and CC collected clinical information and genomic DNAs, sequenced and analyzed exomes and/or genomes. AR and SB conceived the study and wrote the manuscript. All other authors commented on the manuscript. They read and approved the final manuscript.

Funding
Open access funding provided by University of Lausanne This work was supported by grants from the Swiss National Science Foundation (31003A_182632 and IZSTZ0_216615 to AR), the Lejeune Foundation (#1838-2019A to AR), the Blackswan Foundation (to AR), a PRIN 2020 grant from the Italian Ministry of Universities and Research (20203P8C3X to AB), and the US National Institutes of Health (NS105078 and HG011758 to JRL). This study makes use of data generated by the DECIPHER community. Funding for the DECIPHER project was provided by the Wellcome Trust [grant number WT223718/Z/21/Z]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Availability of data and materials
The HEK293T and fibroblast RNA-seq reads generated and analyzed in the current study are accessible at NCBI GEO website under the accession number GSE241621 [64] and GSE246554 [65].

Declarations

Ethics approval and consent to participate
Informed consent forms were obtained for all affected individuals or their guardians participating in this study. The current study was approved by the CER ("Commission d'éthique de la recherche") of the canton of Vaud (Protocol number: CER-VD 2021–01400). This research complies with the principles of the Declaration of Helsinki.

Consent for publication
Informed consent forms for data publication were obtained for all affected individuals or their guardians participating in this study.

Competing interests
Annabelle Tuttle, Houda Zghal Elloumi and Chaofan Zhang are employees of GeneDx and Desiree DeMille works for ARUP Laboratories. James R. Lupski has stock ownership in 23andMe and is a paid consultant for Genome International. Claudia M.B. Carvalho provides consulting service for Ionis Pharmaceuticals. The other authors have no competing interests to declare.

Author details
[1]Center for Integrative Genomics, University of Lausanne, Genopode Building, Lausanne, CH 1015, Switzerland. [2]Bioinformatics Competence Center, University of Lausanne, Lausanne, Switzerland. [3]Bioinformatics Competence Center, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland. [4]Biostatistics Platform, University of Lausanne, Lausanne, Switzerland. [5]Department of Neurosciences Rita Levi-Montalcini, University of Turin, 10126 Turin, Italy. [6]Medical Genetics Unit, Città Della Salute E Della Scienza University Hospital, 10126 Turin, Italy. [7]Institute of Human Genetics, University of Leipzig Medical Center, Leipzig, Germany. [8]Department of Pediatrics, Centre for Neuromedicine, Gemeinschaftskrankenhaus Herdecke Gerhard-Kienle-Weg, Herdecke, Germany. [9]Genomics Analysis 396, ARUP Laboratories, Salt Lake City, UT, USA. [10]Pediatric Neurology, University of Utah School of Medicine, Salt Lake City, UT, USA. [11]Department of Pediatrics, Medical Center North, Vanderbilt University Medical Center, Nashville, TN, USA. [12]Department of Human Genetics, Research Institute for Medical Innovation, Radboud University Medical Center, Nijmegen, The Netherlands. [13]Center for Genetic Developmental Disorders Southwest, Zuidwester, Middelharnis, The Netherlands. [14]Department of Medical Genetics, University Hospital of North Norway, Tromsø, Norway. [15]Department of Pediatric Rehabilitation, University Hospital of North Norway, Tromsø, Norway. [16]Greenwood Genetic Center, Greenwood, SC, USA. [17]Pacific Northwest Research Institute (PNRI), Broadway, Seattle, WA, USA. [18]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA. [19]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA. [20]Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA. [21]Texas Children's Hospital, Houston, TX, USA.

[22]Hospital Ángeles Puebla, Puebla, Mexico. [23]CHU Lille, Clinique de Génétique, 59000 Lille, France. [24]Inserm UMR1231, University of Burgundy, 21000 Dijon, France. [25]GeneDx, Gaithersburg, MD, USA. [26]Department of Genetics, Cook Children's Medical Center, Cook Children's Health Care System, Fort Worth, TX, USA. [27]Institute for Human Genetics, University Hospital Muenster, Muenster, Germany. [28]Department of Neuroscience, Rehabilitation, Ophthalmology, Genetics, Maternal and Child Health (DINOGMI), University of Genoa, Genoa 16132, Italy. [29]Medical Genetics Unit, IRCCS Istituto Giannina Gaslini, Genoa, Italy. [30]Department of Medicine, Child and Adolescent Neuropsychiatry Unit, Surgery and Dentistry, University of Salerno, Salerno, Italy. [31]Clinical Institute of Genomic Medicine, University Medical Centre Ljubljana, Ljubljana, Slovenia. [32]Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia. [33]Department of Genomics, Turku University Hospital, Turku, Finland; University of Turku, Turku, Finland. [34]Department of Pediatric Neurology, Turku University Hospital, Turku, Finland; University of Turku, Turku, Finland. [35]Department of Clinical Genetics, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands. [36]Present address: Institute of Medical Genetics, University of Zurich, Zurich, Switzerland. [37]Present address: Department of Molecular Medicine, University of Pavia, Pavia, Italy. [38]Present address: Medical Genetics Unit, IRCCS San Matteo Foundation, Pavia, Italy. [39]Present address: Mayo Clinic, Rochester, MN, USA. [40]Present address: Sophia Genetics, St Sulpice, Switzerland.

Received: 12 December 2023 Accepted: 19 April 2024
Published: 30 May 2024

References

1. Melko M, et al. Functional characterization of the AFF (AF4/FMR2) family of RNA-binding proteins: insights into the molecular pathology of FRAXE intellectual disability. Hum Mol Genet. 2011;20:1873–85.
2. Guo C, et al. The super elongation complex (SEC) mediates phase transition of SPT5 during transcriptional pause release. EMBO Rep. 2023;24:e55699.
3. Luo Z, et al. The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output. Mol Cell Biol. 2012;32:2608–17.
4. Chen Y, Cramer P. Structure of the super-elongation complex subunit AFF4 C-terminal homology domain reveals requirements for AFF homo- and heterodimerization. J Biol Chem. 2019;294:10663–73.
5. Nilson I, et al. Exon/intron structure of the human AF-4 gene, a member of the AF-4/LAF-4/FMR-2 gene family coding for a nuclear protein with structural alterations in acute leukaemia. Br J Haematol. 1997;98:157–69.
6. House CM, et al. A binding motif for Siah ubiquitin ligase. Proc Natl Acad Sci U S A. 2003;100:3101–6.
7. Jonkers I, Lis JT. Getting up to speed with transcription elongation by RNA polymerase II. Nat Rev Mol Cell Biol. 2015;16:167–77.
8. Wang Y, et al. A permissive chromatin state regulated by ZFP281-AFF3 in controlling the imprinted Meg3 polycistron. Nucleic Acids Res. 2017;45:1177–85.
9. Tsukumo SI, et al. AFF3, a susceptibility factor for autoimmune diseases, is a molecular facilitator of immunoglobulin class switch recombination. Sci Adv. 2022;8:eabq0008.
10. Takeo M, et al. Cyclical dermal micro-niche switching governs the morphological infradian rhythm of mouse zigzag hair. Nat Commun. 2023;14:4478.
11. Voisin N, et al. Variants in the degron of AFF3 are associated with intellectual disability, mesomelic dysplasia, horseshoe kidney, and epileptic encephalopathy. Am J Hum Genet. 2021;108:857–73.
12. Khan H, et al. A novel variant in AFF3 underlying isolated syndactyly. Clin Genet. 2023;103:341–5.
13. Steichen-Gersdorf E, et al. Triangular tibia with fibular aplasia associated with a microdeletion on 2q112 encompassing LAF4. Clin Genet. 2008;74:560–5.
14. Kraft K, et al. Deletions, Inversions, Duplications: Engineering of Structural Variants using CRISPR/Cas in Mice. Cell Rep. 2015;10:833–9.
15. Birling MC, et al. A resource of targeted mutant mouse lines for 5,061 genes. Nat Genet. 2021;53:416–9.
16. Veitia RA. AFF3: a new player in maintaining XIST monoallelic expression. J Mol Cell Biol. 2019;11:723–4.
17. Zhang Y, et al. AFF3-DNA methylation interplay in maintaining the monoallelic expression pattern of XIST in terminally differentiated cells. J Mol Cell Biol. 2019;11:761–9.
18. Luo Z, et al. Regulation of the imprinted Dlk1-Dio3 locus by allele-specific enhancer activity. Genes Dev. 2016;30:92–101.
19. Inoue Y, et al. Three KINSSHIP syndrome patients with mosaic and germline AFF3 variants. Clin Genet. 2023;103:590–5.
20. Raible SE, et al. Clinical and molecular spectrum of CHOPS syndrome. Am J Med Genet A. 2019;179:1126–38.
21. Izumi K, et al. Germline gain-of-function mutations in AFF4 cause a developmental syndrome functionally linking the super elongation complex and cohesin. Nat Genet. 2015;47:338–44.
22. Karczewski KJ, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020;581:434–43.
23. Jadhav B, et al. A phenome-wide association study of methylated GC-rich repeats identifies a GCC repeat expansion in AFF3 as a significant cause of intellectual disability. medRxiv [Preprint]. 2023:2023.05.03.23289461. https://doi.org/10.1101/2023.05.03.23289461.
24. Metsu S, et al. FRA2A is a CGG repeat expansion associated with silencing of AFF3. PLoS Genet. 2014;10:e1004242.
25. Mattioli F, et al. Biallelic truncation variants in ATP9A are associated with a novel autosomal recessive neurodevelopmental disorder. NPJ Genom Med. 2021;6:94.
26. Alfaiz AA, et al. TBC1D7 mutations are associated with intellectual disability, macrocrania, patellar dislocation, and celiac disease. Hum Mutat. 2014;35:447–51.
27. Goujon M, et al. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res. 2010;38:W695–9.
28. Sievers F, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.
29. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis. 1997;18:2714–23.
30. Labun K, et al. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res. 2019;47:W171–4.
31. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. Nucleic Acids Res. 2014;42:W401–7.
32. Bassani S, et al. Variants in USP48 encoding ubiquitin hydrolase are associated with autosomal dominant non-syndromic hereditary hearing loss. Hum Mol Genet. 2021;30:1785–96.
33. Schindelin J, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9:676–82.
34. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21.
35. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–30.
36. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.
37. Huber W, et al. Orchestrating high-throughput genomic analysis with Bioconductor. Nat Methods. 2015;12:115–21.
38. Wu T, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb). 2021;2:100141.
39. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7.
40. Subramanian A, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545–50.
41. Liberzon A, et al. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011;27:1739–40.
42. Liberzon A, et al. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst. 2015;1:417–25.
43. Zhang Y, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9:R137.
44. Yu G, Wang LG, He QY. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics. 2015;31:2382–3.
45. Deciphering Developmental Disorders S. Prevalence and architecture of de novo mutations in developmental disorders. Nature. 2017;542:433–8.

46. Sobreira N, Schiettecatte F, Valle D, Hamosh A. GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. Hum Mutat. 2015;36:928–30.
47. Firth HV, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. Am J Hum Genet. 2009;84:524–33.
48. Harripaul R, et al. Mapping autosomal recessive intellectual disability: combined microarray and exome sequencing identifies 26 novel candidate genes in 192 consanguineous families. Mol Psychiatry. 2018;23:973–84.
49. Tang D, et al. Structural and functional insight into the effect of AFF4 dimerization on activation of HIV-1 proviral transcription. Cell Discov. 2020;6:7.
50. Qi S, et al. Structural basis for ELL2 and AFF4 activation of HIV-1 proviral transcription. Nat Commun. 2017;8:14076.
51. Kalueff AV, et al. Towards a comprehensive catalog of zebrafish behavior 10 and beyond. Zebrafish. 2013;10:70–86.
52. Sillar KT. Mauthner cells. Curr Biol. 2009;19:R353–5.
53. Skarnes WC, et al. A conditional knockout resource for the genome-wide study of mouse gene function. Nature. 2011;474:337–42.
54. Niederriter AR, et al. In vivo modeling of the morbid human genome using Danio rerio. J Vis Exp. 2013;(78):e50338. https://doi.org/10.3791/50338.
55. Ioannidis NM, et al. REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. Am J Hum Genet. 2016;99:877–85.
56. Sollis E, et al. The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. Nucleic Acids Res. 2023;51:D977–85.
57. Telley L, et al. Sequential transcriptional waves direct the differentiation of newborn neurons in the mouse neocortex. Science. 2016;351:1443–6.
58. Cardoso-Moreira M, et al. Gene expression across mammalian organ development. Nature. 2019;571:505–9.
59. Moore JM, et al. Laf4/Aff3, a gene involved in intellectual disability, is required for cellular migration in the mouse cerebral cortex. PLoS ONE. 2014;9:e105933.
60. den Hoed J, et al. Mutation-specific pathophysiological mechanisms define different neurodevelopmental disorders associated with SATB1 dysfunction. Am J Hum Genet. 2021;108:346–56.
61. O'Donnell-Luria AH, et al. Heterozygous Variants in KMT2E Cause a Spectrum of Neurodevelopmental Disorders and Epilepsy. Am J Hum Genet. 2019;104:1210–22.
62. Haijes HA, et al. De Novo Heterozygous POLR2A Variants Cause a Neurodevelopmental Syndrome with Profound Infantile-Onset Hypotonia. Am J Hum Genet. 2019;105:283–301.
63. Paul MS, et al. Rare EIF4A2 variants are associated with a neurodevelopmental disorder characterized by intellectual disability, hypotonia, and epilepsy. Am J Hum Genet. 2023;110:548.
64. Bassani S, A.G., Voisin N, Sirchia F, Turban L, Schubert S, Rami AJ, Schlump J, DeMille D, Bayrak-Todemir P, Nelson GR, Wong KN, Duncan L, Mosera M, Gilissen C, Vissers L, Pfundt R, Kersseboom R, Yttervik H, Hansen GA, Jonsrud C, Smeland MF, Lyons MJ, Carvalho CM, Zhang C, Lupski JR, Potocki L, Flores-Gallegos L, Morales-Toquero R, Petit F, Yalcin B, Tuttle A, Mccormick L, Kukolich M, Scala M, Iacomino M, Zara F, Writzl K, Maver A, Iseli C, Guex N, Reymond A. Variant-specific pathophysiological mechanisms of AFF3 differently influence transcriptome profiles. *NCBI Gene Expression Omnibus* 2024 GSE241621, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE241621.
65. Bassani S, C.J., Ambrosini G, Voisin N, Sirchia F, Turban L, Schubert S, Rami AJ, Schlump J, DeMille D, Bayrak-Todemir P, Nelson GR, Wong KN, Duncan L, Mosera M, Gilissen C, Vissers L, Pfundt R, Kersseboom R, Yttervik H, Hansen GA, Jonsrud C, Smeland MF, Lyons MJ, Carvalho CM, Zhang C, Lupski JR, Potocki L, Flores-Gallegos L, Morales-Toquero R, Petit F, Yalcin B, Tuttle A, Mccormick L, Kukolich M, Scala M, Iacomino M, Zara F, Writzl K, Maver A, Iseli C, Guex N, Reymond A. Variant-specific pathophysiological mechanisms of AFF3 differently influence transcriptome profiles. *NCBI Gene Expression Omnibus*. 2024 GSE246554, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE246554.

Publisher's Note