De Novo Variants Disrupting the HX Repeat Motif of ATN1 Cause a Recognizable Non-Progressive Neurocognitive Syndrome

Elizabeth E. Palmer,[1,2,3,4,25] Seungbeom Hong,[5,25] Fatema Al Zahrani,[6] Mais O. Hashem,[6] Fajr A. Aleisa,[5] Heba M. Jalal Ahmed,[5] Tejaswi Kandula,[1,2] Rebecca Macintosh,[1] Andre E. Minoche,[3] Clare Puttick,[3] Velimir Gayevskiy,[3] Alexander P. Drew,[3] Mark J. Cowley,[3,7] Marcel Dinger,[3,7] Jill A. Rosenfeld,[8] Rui Xiao,[8,9] Megan T. Cho,[10] Suliat F. Yakubu,[5] Lindsay B. Henderson,[10] Maria J. Guillen Sacoto,[10] Amber Begtrup,[10] Muddathir Hamad,[11] Marwan Shinawi,[12] Marisa V. Andrews,[12] Marilyn C. Jones,[13] Kristin Lindstrom,[14] Ruth E. Bristol,[15] Saima Kayani,[16] Molly Snyder,[17] María Mercedes Villanueva,[18] Angeles Schteinschnaider,[18] Laurence Faivre,[19,20] Christel Thauvin,[19] Antonio Vitobello,[19] Tony Roscioli,[1,21,22] Edwin P. Kirk,[1,2,21] Ann Bye,[1,2] Jasmeen Merzaban,[23] Łukas Jaremko,[5] Mariusz Jaremko,[23] Rani K. Sachdev,[1,2] Fowzan S. Alkuraya,[6,24,25,*] and Stefan T. Arold[5,25,*]

Polyglutamine expansions in the transcriptional co-repressor Atrophin-1, encoded by *ATN1,* cause the neurodegenerative condition dentatorubral-pallidoluysian atrophy (DRPLA) via a proposed novel toxic gain of function. We present detailed phenotypic information on eight unrelated individuals who have *de novo* missense and insertion variants within a conserved 16-amino-acid "HX repeat" motif of ATN1. Each of the affected individuals has severe cognitive impairment and hypotonia, a recognizable facial gestalt, and variable congenital anomalies. However, they lack the progressive symptoms typical of DRPLA neurodegeneration. To distinguish this subset of affected individuals from the DRPLA diagnosis, we suggest using the term CHEDDA (congenital hypotonia, epilepsy, developmental delay, digit abnormalities) to classify the condition. CHEDDA-related variants alter the particular structural features of the HX repeat motif, suggesting that CHEDDA results from perturbation of the structural and functional integrity of the HX repeat. We found several non-homologous human genes containing similar motifs of eight to 10 HX repeat sequences, including *RERE*, where disruptive variants in this motif have also been linked to a separate condition that causes neurocognitive and congenital anomalies. These findings suggest that perturbation of the HX motif might explain other Mendelian human conditions.

The combination of unbiased chromosomal analysis (chromosomal microarray) and next-generation sequencing approaches (exome sequencing [ES] and whole-genome sequencing [WGS]), along with the use of databases that promote sharing of information on genotype and phenotype, is enabling the identification and validation of genetic conditions and improved diagnostic rates for complex congenital conditions.[1] Such unbiased genetic approaches can also unveil the complexity of how different types of genetic variation in a particular gene can result in varied and sometimes distinct phenotypic presentations.[2–6]

ATN1 (MIM: 607462), located at chromosomal region 12p13.31, comprises 10 exons and has two transcript variants (GenBank: NM_001007026.1, NM_001940.3) that differ only in their untranslated exons. *ATN1* encodes atrophin-1 (ATN1), a member of a class of evolutionarily conserved transcriptional corepressors involved in nuclear signaling.[7] The normal roles of ATN1 are incompletely understood; however, converging evidence supports a role for this protein as a nuclear transcriptional regulator important in the control of brain and other organ system development.[8–10] Although *Atn1*$^{-/-}$ mice are neurologically

[1]Sydney Children's Hospital, Randwick, NSW 2031, Australia; [2]School of Women's and Children's Health, University of New South Wales, Randwick, NSW 2031, Australia; [3]The Kinghorn Centre for Clinical Genomics, Garvan Institute of Medical Research, Darlinghurst, NSW 2010, Australia; [4]Genetics of Learning Disability Service, Hunter Genetics, Waratah, NSW 2298, Australia; [5]King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Division of Biological and Environmental Sciences and Engineering (BESE), Thuwal 23955-6900, Saudi Arabia; [6]Department of Genetics, King Faisal Specialist Hospital and Research Center, Riyadh 11211, Saudi Arabia; [7]St. Vincent's Clinical School, University of New South Wales, Darlinghurst, NSW 2010, Australia; [8]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas 77030, USA; [9]Baylor Genetics, Houston, Texas 77021, USA; [10]GeneDx, Gaithersburg, Maryland 20877, USA; [11]King Khalid University Hospital, King Saud University, Riyadh 11472, Saudi Arabia; [12]Department of Pediatrics, Division of Genetics and Genomic Medicine, Washington University School of Medicine, St. Louis, Missouri 63110, USA; [13]Division of Genetics, Department of Pediatrics, University of California, San Diego and Rady Children's Hospital, San Diego, California 92123, USA; [14]Division of Genetics and Metabolism, Phoenix Children's Hospital, Phoenix, Arizona 85016, USA; [15]Division of Pediatric Neurosurgery, Phoenix Children's Hospital, Phoenix, AZ 85016, USA; [16]University of Texas Southwestern Medical Center, Dallas, Texas 75390, USA; [17]Department of Neurology, Children's Health, Dallas, Texas 75235, USA; [18]Fundación para la Lucha contra las Enfermedades Neurológicas de la Infancia, Montañeses, Buenos Aires 2325, Argentina; [19]Inserm U1231, Lipides, Nutrition, Cancer UMR 1231 Génétique des Anomalies du Développement, Burgundy University, Dijon 21079, France; [20]Reference Center for Developmental Anomalies, Department of Medical Genetics, Dijon University Hospital, Dijon 21079, France; [21]New South Wales Health Pathology Genomic Laboratory, Prince of Wales Hospital, Randwick 2031, Australia; [22]Neuroscience Research Australia, University of New South Wales 2031, Australia; [23]King Abdullah University of Science and Technology, Division of Biological and Environmental Sciences and Engineering, Thuwal 23955-6900, Saudi Arabia; [24]Saudi Human Genome Program, King Abdulaziz City for Science and Technology, Riyadh 11442, Saudi Arabia
[25]These authors contributed equally
*Correspondence: falkuraya@kfshrc.edu.sa (F.S.A.), stefan.arold@kaust.edu.sa (S.T.A.)
https://doi.org/10.1016/j.ajhg.2019.01.013.

normal,[8] Zhang et al.[9] demonstrated that knockdown of *Atn1* in rat neuronal progenitor cells (NPCs) led to significant abnormalities in brain development; these abnormalities could be largely rescued by co-transfection with a human *ATN1* construct. That study also demonstrated that ATN1 is a direct target of the lysine-specific histone demethylase 1A (LSD1), a protein known to have key developmental roles, including controlling embryonic stem cell differentiation, cortical neuronal migration, and adult NPC proliferation.[9] *ATN1* transcripts are widely expressed, including in brain, heart, lung, kidney, and skeletal muscle; expression is higher in fetal tissues, especially in the brain.[11] In the human adult brain, *ATN1* is broadly expressed in multiple regions, including the amygdala, corpus callosum, hippocampus, hypothalamus, caudate nucleus, substantia nigra, subthalamic nucleus, and thalamus, consistent with a role for ATN1 in central nervous system development and function.[11]

The only human condition definitively associated with *ATN1* to date is the autosomal-dominant neurodegenerative condition dentatorubral-pallidoluysian atrophy (DRPLA, MIM: 125370)[12,13] caused by a polyglutamine expansion in exon 5. DRPLA is characterized by the progressive neurological features of choreoathetosis, myoclonus, epilepsy, ataxia, and dementia. Age of onset ranges from infancy to late adulthood, dependent on size of the expansion.[14,15] Congenital anomalies are not a feature. The underlying pathogenic mechanism whereby polyglutamine expansion of ATN1 causes DRPLA is incompletely understood: it is postulated that a toxic gain-of-function effect of the expanded polyglutamine tract causes neurotoxicity rather than simple loss of function. These toxic effects might include formation of peri- and intranuclear inclusions; abnormal protein cleavage or abnormal phosphorylation of ATN1; and downstream suppression of cAMP-response-element-binding protein (CREB)-dependent transcriptional activation, which is required for neuronal plasticity and survival.[16–19]

We report here on a cohort of eight individuals affected with overlapping severe primarily neurocognitive phenotypes. All of these individuals harbored *de novo* variants in a specific and highly evolutionarily conserved and invariant 16 amino acid motif, consisting of a histidine-rich 16 amino acid motif encoded by exon 7 of *ATN1* (Figure 2). This motif is distal to the Gln-rich region involved in DRPLA, and the affected individuals lacked the progressive neurodegenerative features characteristic of DRPLA.[14]

For all affected individuals, within the first three months of life, there arose concerns regarding significant hypotonia, feeding difficulties, seizures, congenital malformations, and distinctive facial features, and all have severe to profound global developmental delay and/or intellectual disability, truncal hypotonia, global motor disability, and very limited verbal communication. (See Table 1 for an overview of the clinical data and Table S2 for further clinical details.) Five have a seizure disorder; for four of these, the seizure disorder could be described as a neonatal or infantile-onset developmental encephalopathy. Hearing and visual impairments and functional gastrointestinal disorders were common and frequently severe; four individuals required orogastric feeding or total parenteral nutrition. Other than suboptimal weight gain in those with more significant feeding difficulties, growth parameters were within the normal range. Individual 8 was born prematurely at 33 weeks and died at 2 months of age as a result of respiratory distress in the setting of severe multiple congenital anomalies.

Congenital structural anomalies were common but variable between individuals: four individuals had cardiac malformations, including atrial and ventricular septal defects, plus abnormalities of the aorta and superior vena cava; two individuals had palatal clefts; three individuals had congenital renal anomalies; and two had an anteriorly placed anus. Common neuroanatomical abnormalities were evident on examination of available MRI in one center (individuals 2, 5, 7, and 8) (Figure S1). Several individuals had cranio-skeletal abnormalities: in particular, two individuals had stenosis of the craniocervical junction, which prompted screening for this complication in all individuals. When the individuals were assessed as a group, a similarity in facial features was apparent (Figure 1), a particularly striking feature being sparsity of the lateral forehead hair and low-set, posteriorly rotated ears. Characteristic hand and foot features were overlapping toes, camptodactyly, persistent fetal fingertip pads, and abnormalities of the palmar creases (Figure 1).

This clinical cohort was collated through the identification of individuals with *de novo ATN1* variants listed in the ClinVar database, as well as via contact with individual diagnostic laboratories and networking at human genetics conferences. Prior to ES or WGS, all individuals with a *de novo ATN1* variant were undiagnosed despite clinical genetic assessment and prior genetic screening, which in all individuals included chromosomal microarray. (Further details of prior genetic studies are provided in Table S1.) Individual 5 was previously included in a large exome sequencing study,[21] and individual 8 was previously described clinically without a molecular diagnosis.[22] All *ATN1* variants in the clinical cohort were rare in that they were absent from the 125,748 exomes and 15,708 genomes listed in the gnomAD 2.1 database (a database depleted of individuals with severe pediatric disease[23]) and from the BRAVO database of 62,785 healthy individuals. All missense variants were predicted to be pathogenic by the majority of *in silico* pathogenicity scoring tools (see Table 2). No affected individual had another plausible cause for their neurocognitive condition or congenital anomalies after ES or WGS variant filtering and prioritization (see Table S2). Segregation analysis was consistent with the variant's being *de novo* for all individuals and provided no evidence of mosaicism. Details of the methods for sequencing, variant filtering, and prioritization are provided in the Supplemental Data. Genetic studies were approved by local ethics committees, and written informed consent for

Table 1. Comparison of Clinical Features of Affected Individuals with Missense Variants in the Poly HX Domain of *ATN1*

Affected Individual	1	2	3	4	5	6	7	8[22]
Current Age	3 years	1 year	5 years	7 years	9 years	4 years	5 years	2 months
Gender	M	M	F	F	F	F	F	F
Ethnicity	Argentinian	Hispanic	Hispanic	Hispanic	Saudi	Mexican	Australian	French
Variant cDNA[1]	c.3160C>A	c.3172C>T	c.3177_3178insAACCTG	c.3177_3178insGACCTG	c.3178C>T	c.3184C>G	c.3188T>G	c.3185A>G
AA Change[2]	p.His1054Asn	p.His1058Tyr	p.Ser1059_His1060insAsnLeu	p.Ser1059_His1060insAspLeu	p.His1060Tyr	p.His1062Asp	p.Leu1063Arg	p.His1062Arg
Antenatal Findings	increased nuchal translucency (karyotype N)	oligohydramnios and partial urinary obstruction	normal antenatal USS	normal antenatal USS	no	ambiguous genitalia and cardiac malformation	normal antenatal USS	normal antenatal USS, breech, PROM
Gestation	term	term	term	31^{+6} weeks	term	term	term	33 weeks
Birth Centiles: Length	NA	>90%	NA	50%–75%	NA	50%	15%	50%
Birth Centiles: Weight	3%–15%	50%	3%	25%–50%	NA	15%	3%	50%
Birth Centiles: Head Circumference	NA	NA	15%–50%	50%	NA	50%	15%–50%	75%
Current Centiles: Length	3%	85%	25%	5%	25%	30%	3%–10%	NK
Current Centiles: Weight	3%	40%	85%	3%	2%	85%	3%	NK
Current Centiles: Head Circumference	50%	45%	60%	85%	75%	20%	25%	NK
Infantile Hypotonia	yes	yes	yes	yes	yes	yes	yes	yes
Current Neurology	global hypotonia	central hypotonia, appendicular spasticity	global hypotonia	global hypotonia, hyperkinetic UL movts	central hypotonia and appendicular spasticity	global hypotonia	global hypotonia	axial hypotonia and appendicular hypertonia
Overt Seizure Disorder	yes; Lennox Gestaut onset 1yr	no; no clinical seizures	no	yes; inf spasms controlled 2 AED	yes; controlled MT	yes; intractable neonatal onset EE (mixed types)	yes; EE onset 7 mos controlled MT	no
EEG	MEA with slow spike and wave and periods of voltage attenuation	bitemporal epileptiform discharges	focal theta slowing	hypsarrhythmia + bg slowing	diffuse slowing	diffuse slowing	MEA	ND

(*Continued on next page*)

Table 1. ***Continued***

Affected Individual	1	2	3	4	5	6	7	8[22]
Level of DD or ID	severe-profound GDD	GDD	severe GDD	profound	severe	severe-profound	severe	severe
Visual Impairment	yes; does not fix or follow, corneal leukoma	yes; CVI	no	yes; CVI	no	yes; CVI	yes; CVI	yes; microphthalmia
Hearing Impairment	yes	yes; bl mod (hearing aids)	yes; OME (grommets)	no; OME (grommets)	no; OME (grommets)	yes; bl sn	yes; bl mod (hearing aids and grommets)	NK
Verbal Ability	non-verbal	coos	single words	non-verbal	non-verbal	non-verbal	babbles	none
Gross Motor Ability	no head control, cannot roll	rolls to side	walks few steps unsupported	sits with support	immobile	sits with support	sits with support	none
Fine Motor Ability			holds small objects	grasps objects			grasps objects	
MRI Brain	parenchymal atrophy, unilateral PVL, left cerebellar hyperintensity	peri-sylvian polymicrogyria, parenchymal atrophy, thin CC, absent falx cerebri	normal	normal	peri-sylvian polymicrogryri, thin CC, partial absence falx cerebri, parenchymal atrophy	vermian hypoplasia	peri-sylvian polymicrogyria, thin CC, absent falx cerebri	polymicrogyria of the rt Sylvian fissure, vermian hypoplasia, thin CC
MRI Cervical Spine	ND	craniocervical stenosis	craniocervical stenosis	ND	ND	normal	normal	ND
Respiratory Symptoms	no	yes; severe OSA	no	no	yes; asthma	yes; ul choanal stenosis, O+CSA (tracheostomy)	yes; OSA (CPAP)	yes; respiratory distress
Orofacial Clefting	no; high, narrow palate	yes; small, hard palate cleft	no	no	no	no	no; high, narrow palate	yes; cleft palate and gingiva
GI Abnormalities	yes; pyloric hypertrophy, dysphagia GERD	yes; GERD NEC	yes; dysphagia constipation	no	yes; GERD	yes; dysphagia GERD, ant anus	yes; dysphagia GERD, constipation	yes; ant. anus
Nutrition	oral feeding	TPN	self feeds	oral feeding	G tube	G tube	oral feeding	orogastric feeding
Congenital Heart Disease	no	yes; ASD+	no	no	yes; VSD + ASD	yes; CoA + hypoplasia LV and AA	no	yes; large foramen ovale, persistence left SVC
Genitourinary Disease	right renal agenesis	cryptorchidism, VUR, ul hydrouteronephrosis rUTI	no	no; normal USS	yes; r UTI normal USS	no; normal USS	no; normal USS	yes; left non-dysplastic renal hypoplasia

(*Continued on next page*)

Table 1. *Continued*

Affected Individual	1	2	3	4	5	6	7	8[22]
Skeletal System	joint hypermobility, scoliosis	hip dysplasia	short trunk DDH, joint laxity	mild scoliosis	DDH	sagittal craniosynostosis	normal	phalangeal hypoplasia
Facial Gestalt	yes	yes	yes	yes	yes	yes	yes	yes
Hands and Feet	overlapping digits	overlapping digits	overlapping digits	overlapping toes	abnormal PC	overlapping toes, single PC, fetal pads	overlapping toes, fetal pads	overlapping toes, proximally implanted thumbs, single PC
Others	-	-	-	-	-	inv nipples	inv nipples	

Abbreviations are as follows: AA, aortic arch; abn, abnormal; AED, antiepileptic drug; Ant. anus, anteriorly placed anus; AmA, amino acid; ASD, atrial septal defect; bg, background; bl, bilateral; CC, corpus callosum; C+OSA, central and obstructive sleep apnoea; CVI, cortical visual impairment; CoA, coarctation of the aorta; DDH, developmental disorder of the hips; EE, epileptic encephalopathy; F, female; GDD, global developmental delay; GERD, gastresophageal reflux disease; G tube, gastrostomy tube; inf spasms, infantile spasms; inv, inverted; LV, left ventricle; MEA, multifocal epileptiform activity; MT, monotherapy (i.e., controlled on monotherapy); M: male; mo: month; mod: moderate; movts: movements; NA: not available; ND: not done; NK: not known; NEC: acute necrotising enterocolitis; PC: palmar crease; PROM: premature rupture of membranes; OSA: obstructive sleep apnoea; OME: otitis media with effusions PVL: periventricular leukomalacia; sn: sensorineural; SVC: superior vena cava; TPN: total parenteral nutrition; rUTI: recurrent urinary tract infections; r: recurrent; UL: upper limb; ul: unilateral; USS: ultrasound; VUR: vesicoureteric reflux; VSD: ventriculoseptal defect; year: year.
[1]Based on transcript GenBank: NM_001007026.1.
[2]Protein ID GenBank: NP_001007027.1.

molecular genetic analysis and for the publication of clinical and radiological data and photographs, which were obtained as part of standard diagnostic procedures, was obtained from the participants' legal guardians.

The histidine-rich motif that is perturbed in all individuals of our cohort is located in the C-terminal part of ATN1 (residues 1049–1065). It consists of eight HX repeats, where H is a histidine and X is any amino acid (Figures 2A–2C). The DNA region encoding for this HX repeat motif was covered in the sequencing of all individuals to a depth of at least 20 reads (Table S2). A search for the $(HX)_8$ pattern in the human proteome (via PatternSearch[29]) yielded 71 sequences (see Supplemental Data), 20 of which appeared to be distinct from ATN1's HX repeat in that the "X" position only contained histidines or prolines (see PHLDA1, Figure 2B). The remaining 51 sequences were isoforms of ATN1 and its paralogous arginine-glutamic-acid dipeptide-repeat protein RERE (which is also a transcriptional repressor); were isoforms of the paralogous autism-susceptibility gene 2 protein (AUTS2, a component of the PRC1-like complex involved in maintaining the transcriptional repressive state of many genes during development), fibrosin (FBRS), and fibrosin-like proteins (FBRSL1); or belonged to the ZIP family of zinc transporters (Figure 2). Within these 51 sequences, the "X" position showed limited variability and mostly consisted of Gln and Thr (Figure 2D).

To probe the impact of a pathological variant on the molecular behavior of the ATN1 protein, we studied two ATN1-derived polypeptides containing the HX repeat motif in solution by nuclear magnetic resonance (NMR). The peptides 1046-NVTPHHHQHSHIHSHLHLHQQD-1067 ($ATN1_{1046–1067}$) and another bearing the variant p.His1060Tyr 1046-NVTPHHHQHSHIHS**Y**LHLHQQD-1067 ($ATN1_{1046–1067}^{p.His1060Tyr}$) were commercially synthesized and dissolved in 500 μL of 100% D_2O at a concentration of 2.2 mg/mL. NMR experiments were performed on a 700 MHz Bruker spectrometer at 25°C and a pD (the pH for D_2O) between 5.51 and 6.4. The NMR data were processed by NMRPipe[30] and analyzed with SPARKY. ^1H and ^{13}C resonances for the two peptides were analyzed via standard procedures[31] on the basis of 2D homonuclear 2D ^1H-^1H TOCSY (mixing times were 10 and 80 ms) and 2D ^1H-^1H ROESY, as well as 2D ^1H-^1H NOESY (mixing times were 300 and 500 ms); these analyses were based on 2D ^1H-^{13}C HSQC experiments, which made use of the natural abundance of ^{13}C, were carried out using separate tuning for the aliphatic and aromatic regions. For details, see Supplemental Data.

For all histidines, the cross peaks corresponding to the Hδ2/Cδ2 atoms and those corresponding to the Hε1/Cε1 atoms were each clustered at indistinguishably close ^1H and ^{13}C chemical shifts (Figure 2E). Moreover, 2D homonuclear ^1H -^1H ROESY and 2D ^1H -^1H spectra showed that the trivial ($|j - i| = 0$) and short-range ($|j - i| < 2$) NOE cross peaks were absent or of very weak intensity (Figure 2F). Both types of experiments indicated that the regularly spaced occurrence of histidines introduces a spatial and

Figure 1. Clinical Images of Affected Individuals with CHEDDA
(A) Facial images of affected individuals 1–7. Common facial features include tall foreheads with bitemporal narrowing; deeply set eyes; sparsity of the lateral forehead hair; low-set posteriorly rotated ears; a bulbous, slightly overhanging nasal tip; longer philtrum; prominent columella; and a thin upper lip.
(B) Hand and foot images of affected individuals 1, 2, 3, 5, 6 and 7. Common features include abnormalities of the palmar creases, bulbous endings to the fingers and toes, and overlapping toes.

dynamical synchronization of the histidines. This synchronization was lost in $ATN1_{1046–1067}^{p.His1060Tyr}$, which displayed dispersed cross peaks of the side chain aromatic region and stronger intensity of the NOE cross peaks.

Side-chain imidazole rings of histidines are known to coordinate metal ions.[32,33] Zn^{2+} binding was assessed by the stepwise addition of $ZnCl_2$ stock solution (500 mM in 100% D_2O) to the solution of peptides reaching the peptide:Zn^{2+} molar ratios of 1:0.5, 1:1, 1:4, 1:8, 1:16, 1:32, and 1:48. For each peptide:Zn^{2+} ratio, the pD was checked and corrected if required, and the same 1D and 2D NMR spectra as for the free peptides were recorded (see Supplemental

Table 2. ***De novo ATN1*** **Variants Reported in this Clinical Cohort**

Affected Individual	Genomic Location (GRCh37)	Variant cDNA Change (NM_001007026.1)	Amino Acid Change (GenBank: NP_001007027.1)	Present in gnomAD Database?	Segregation	SIFT[24] Score	PROVEAN[25] Score	DANN[26] Score	CADD[27] Score	Classification as per ACMG Guidelines[28]
1	GenBank: NC_000012.11: g.7048286C>A	c.3160C>A	p.His1054Asn	no	*de novo*	damaging (0)	damaging (−6.1)	0.994	29.6	Variant of uncertain significance. PM2: Pathogenic Moderate (absent GnomAD despite good coverage) PP3: Pathogenic Supporting (8/8 pathogenic predictions)
2	GenBank: NC_000012.11: g.7048298C>T	c.3172C>T	p.His1058Tyr	no	*de novo*	damaging (0)	damaging (−5.26)	0.992	23	Variant of uncertain significance. PM2:Pathogenic Moderate (absent GnomAD despite good coverage) PP3: Pathogenic Supporting (6/8 pathogenic predictions)
3	GenBank: NC_000012.11: g.7048303-7048304insAACCTG	c.3177_3178insAACCTG	p.Ser1059_His1060insAsnLeu	no	*de novo*	NA	NA	NA	NA	Variant of uncertain significance. PM2: Pathogenic Moderate (absent GnomAD despite good coverage) PM4: Pathogenic Moderate (in frame variant in *ATN1*, and is not in a repeat region.) PP3: Pathogenic Supporting: 1 pathogenic prediction from GERP (versus no benign predictions).
4	GenBank: NC_000012.11: g.7048303-7048304insGACCTG	c.3177_3178insGACCTG	p.Ser1059_His1060insAspLeu	no	*de novo*	NA	NA	NA	NA	Variant of uncertain significance. PM2: Pathogenic Moderate (absent GnomAD despite good coverage) PM4: Pathogenic Moderate (in frame variant in *ATN1*, and is not in a repeat region.) PP3: Pathogenic Supporting: 1 pathogenic prediction from GERP (versus no benign predictions).
5	GenBank: NC_000012.11: g.7048304C>T	c.3178C>T	p.His1060Tyr	no	*de novo*	damaging (0)	damaging (−5.21)	0.998	28.7	Variant of uncertain significance. PM2: Pathogenic Moderate (absent GnomAD despite good coverage) PP3: Pathogenic Supporting (8/8 pathogenic predictions)
6	GenBank: NC_000012.11: g.7048310C>G	c.3184C>G	p.His1062Asp	no	*de novo*	damaging (0)	damaging (−7.86)	0.992	26.2	Variant of uncertain significance. PM2: Pathogenic Moderate (absent GnomAD despite good coverage) PP3: Pathogenic Supporting (8/9 pathogenic predictions)
7	GenBank: NC_000012.11: g.7048314T>G	c.3188T>G	p.Leu1063Arg	no	*de novo*	damaging (0)	damaging (−5.33)	0.997	29.6	Variant of uncertain significance. PM2: Pathogenic Moderate (absent GnomAD despite good coverage)
8	GenBank: NC_000012.11: g.7048311A>G	c.3185A>G	p.His1062Arg	no	*de novo*	damaging (0)	damaging (−6.89)	0.9938	24.1	Variant of uncertain significance. PM2: Pathogenic Moderate (absent GnomAD despite good coverage) PP3: Pathogenic Supporting (8/8 pathogenic predictions)

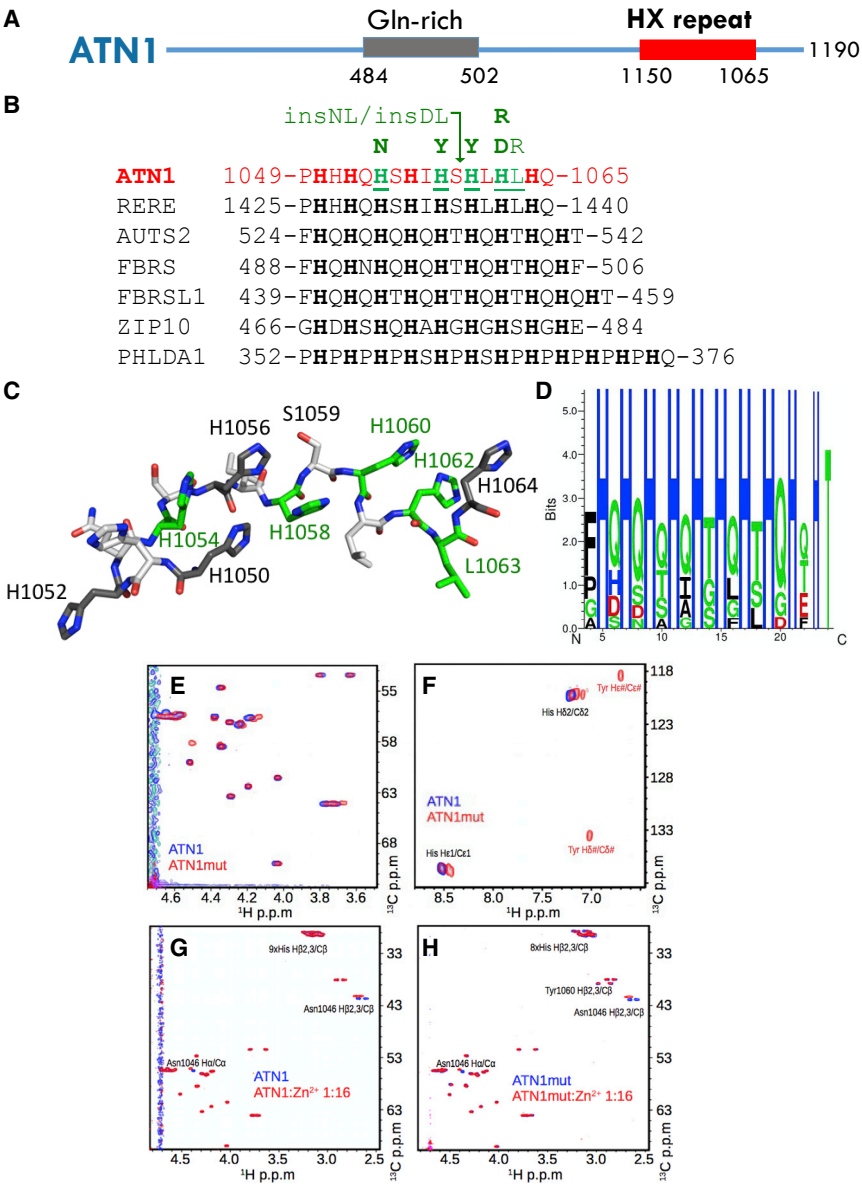


Figure 2. *De novo* Variants in *ATN1* Affect a Highly Invariant Motif

(A) Schematic overview of ATN1; amino acid positions of the Gln-repeat region expanded in DRPLA and the HX repeat motif are illustrated.

(B) Human proteins with their (HX)n repeat motifs, with n ≥ 8. Start and end residue numbers are given and histidines are highlighted. For ATN1, residues found mutated in this study are underlined. The variant substitutions present in affected individuals are indicated above the sequence alignment (green). The arrow head marks the position of the two amino acid insertions Asn-Leu (insNL) and Asp-Leu (insDL). Proteins and database accession numbers are: ATN1: atrophin-1, NP_001007027.1; RERE: arginine-glutamic acid dipeptide repeats protein isoform a, NP_001036146.1; AUTS2: autism susceptibility gene 2 protein, NP_056385.1; FBRS: probable fibrosin-1, NP_001098549.2; FBRSL1: fibrosin-1-like protein, NP_001136113.1; ZIP10: zinc transporter ZIP10 precursor, NP_001120729.1; PHLADA1: pleckstrin homology-like domain family A member 1, NP_031376.3.

(C) 3D structural visualization of the ATN1 HX repeat motif. The 3D structure is not derived experimentally, but only chosen to illustrate the localization of the histidines.

(D) Amino acid enrichment within the HX regions of the 51 human sequences most similar to the ATN1 HX repeat motif (excluding His-only or His-Pro motifs). Figure was produced with Seq2Logo 2.0.[20]

(E-F) 2D heteronuclear ^{1}H - ^{13}C HSQC correlation spectra. Peptides were recorded using natural abundance of ^{1}H and ^{13}C in synthesized peptides. ATN1: $ATN1_{1046\text{-}1067}$; ATN1-mut: $ATN1_{1046\text{-}1067}^{His1060Tyr}$. C-D: "1:16" indicates a peptide: Zn^{2+} ion ratio of 1:16. pD corresponds to the pH in D_2O.

Data). We found that only the histidines of the p.His1060Tyr mutant, but not of the wild-type peptide, bound Zn^{2+} at a pD of 5.5, as demonstrated by the histidine $^{1}H\beta2,3/^{13}C\beta$ chemical shift and signal intensity changes upon addition of Zn^{2+} (Figures 2G and 2H). This particular feature was lost upon deprotonation of the histidines at higher pD (Figures S2A and S2B). Thus, under the conditions used, the HX -repeat motif created specific and unusual pH-dependent zinc-binding properties for the histidine-rich sequence, and those properties were abolished by the variant.

Herein, we describe a recognizable constellation of severe neurocognitive impairment, distinctive facial features, and pleiotropic but overlapping congenital anomalies in eight affected individuals with *de novo* variants in the HX repeat motif encoded within exon 7 of *ATN1*. This static (non-progressive) syndromic phenotype is distinct from the neurodegenerative condition of DRPLA, which is caused by a triplet repeat expansion in exon 5 of *ATN1*. This repeat expansion is thought to result in a toxic gain of function. We propose the name CHEDDA (congenital hypotonia, epilepsy, developmental delay, digit abnormalities) to distinguish this previously unreported condition from *ATN1*-related DRPLA. Our clinical observations are consistent with a role for ATN1 as a key nuclear transcriptional regulator involved in the regulation of organ development, including development of the brain and heart,[9,10,34] and hint at a critical role of the ATN1 HX repeat motif in the control of human embryonic development. That the variants cause a simple haploinsufficiency of *ATN1* is unlikely given the presence of a number (albeit a very small number) of healthy individuals with heterozygous stop-gain, frameshift, and canonical splice variants in gnomAD and BRAVO and the clustering of the variants in CHEDDA within a specific restricted protein motif.

The HX repeat motif was only briefly described in 1991 in a purely bioinformatics-based "hypothesis" publication.[35] The authors reported the presence of $(HX)_n$ repeat

motifs (where n is the number of times the motif is repeated) in certain *Drosophila* transcription factors and suggested that this motif might be used for coordinating zinc binding. Interestingly, the wild-type ATN1 HX repeat sequence showed a strong pH dependency for zinc binding *in vitro*. This feature appears to be linked to the regular histidine spacing, which introduces a specific synchronization of the histidine side chains. The introduction of the p.His1060Tyr variant present in an affected individual in the cohort disrupted this synchronization and allowed zinc binding, as expected for poly-histidine motifs, thus endowing ATN1 with a novel property, albeit with unclear molecular consequences. On a molecular level, the ATN1 HX repeat motif appears therefore to give rise to specific features that distinguish it from other poly-histidine motifs. The *ATN1* HX repeat might serve as a specific pH-dependent interaction motif for ions and/or proteins or other biomolecules.

The hypothesis that disruption in the spacing of the histidines in HX motifs will affect critical functioning is supported by our observation that nine of the 19 reported individuals with neurodevelopmental disorder with or without anomalies of the brain, eye, and heart (NEDBEH, MIM: 616975) have *de novo* variants disrupting the HX motif of RERE and that those individuals with variants in the HX motif, as opposed to the rest of the protein, are more likely to have congenital anomalies, including septal cardiac, eye, and brain anomalies.[36] We also note that three rare *de novo* variants (SCV000571291.3, SCV000837721.1, and SCV000493076.1) in the HX motif of AUTS2 are listed in ClinVar as likely pathogenic and occurred in individuals with an intellectual-disability and congenital-anomaly phenotype.

A possible link between dysregulation of *ATN1* expression and another congenital syndromic neurocognitive condition, Pallister Killian syndrome (PKS, mosaic tetrasomy 12p, MIM: 601803), has also been previously postulated by Kaur et al.[37] *ATN1* lies within the PKS critical region on 12p13.31, and dysregulation of the expression of *ATN1*, among other genes, was demonstrated in fibroblasts from affected individuals with PKS. Kaur et al., postulated that *ATN1* overexpression could be a key driver of the phenotype of PKS via dysregulation of the key developmental HOX genes through the action of the master transcriptional regulator *CREBBP*, although direct evidence was lacking. This speculation is intriguing, given certain similarities in phenotype between individuals in our cohort and individuals affected by PKS; such similarities included severe cognitive impairment, hypotonia, distinctive facial features including high forehead and sparse fronto-temporal hair at birth (the latter representing a relatively rare clinical finding), and variable congenital anomalies, including high arched or cleft palate, polymicrogyria, limb and genitourinary anomalies, and congenital heart defects.[38] Indeed, PKS was considered as a differential diagnosis in individual 7 of our cohort.

The *de novo* variants in the HX motif of ATN1 reported here account for one out of 6,100 individuals who had a neurological phenotype and had exome sequencing performed through Baylor Genetics and for five out of 13,640 individuals who had neurodevelopmental delay and had exome sequencing performed through GeneDx; these latter individuals included another affected individual who had the variant ATN1 (c.3178C>T [p.His1060Tyr] GenBank: NM_001007026.1) (ClinVar: SCV000620232.1) and whose family did not give consent for the inclusion of clinical data. These data suggests a frequency of CHEDDA between 1.6×10^{-4} and 3.7×10^{-4} individuals with neurocognitive and/or neurological disorders. Despite this apparent rarity of CHEDDA, we postulate that more affected individuals might have already had a variant detected in this region through diagnostic ES and WGS, but the lack of the progressive neurological phenotype characteristic of DRPLA might have caused the variants to be unreported or classified as variants of uncertain clinical significance. This situation is similar to those of other clinically distinct conditions our groups have recently described,[2–4] and it highlights genotype-phenotype complexity and the importance of rigorous evaluation of the possible pathogenicity of unreported variants through international clinical and basic science collaborations.[39,40] To assist with the dissemination of accessible information about CHEDDA to clinicians and families of affected individuals, we have adopted *ATN1* on the Human Disease Gene Webseries.

Clarifying the biological role of the HX repeat motif in the ATN1 and AUTS2 protein families and understanding how this function might be altered in CHEDDA and potentially linked conditions such as PKS will require further work. Such work might lead to the development of targeted therapies. For example, Zhang et al. demonstrated that the clinical LSD1 inhibitor, tranylcypromine, could suppress *ATN1* expression, and they suggested that this agent might have potential therapeutic implications for conditions resultant from aberrant *ATN1* expression.[39,41] Further studies into the primary gene-regulatory functions of ATN1 and how this might be altered in CHEDDA and potentially linked conditions such as PKS are warranted.

Accession Numbers

ATN1 c.3178C>T (p.His1060Tyr): GenBank: NM_001007026.1 and ClinVar: SCV000221666.1

ATN1 c.3177_3178insGACCTG (p.Ser1059_His1060insAspLeu): GenBank: NM_001007026.1 and ClinVar: SCV000619648.1

ATN1 c.3177_3178insAACCTG (p.Ser1059_His1060insAsnLeu): GenBank: NM_001007026.1 and ClinVar: SCV000571353.3

ATN1 c.3184C>G (p.His1062Asp): GenBank: NM_001007026.1 and ClinVar: SCV000528073.3

ATN1 c.3160C>A (p.His1054Asn): GenBank: NM_001007026.1 and ClinVar: SCV000678263.1

ATN1 c.3172C>T (p.His1058Tyr): GenBank: NM_001007026.1 and ClinVar: SCV000678264.1

ATN1: c.3188T>G (p.Leu1063Arg): GenBank: NM_001007026.1 and ClinVar: SCV000678265.1

ATN1 c.3185A>G (p.His1062Arg): GenBank: NM_001007026.1 and ClinVar: SCV000853264

Supplemental Data

Supplemental Data can be found with this article online at https://doi.org/10.1016/j.ajhg.2019.01.013.

Acknowledgments

The authors thank the affected individuals and their families for their participation in this study and also thank the clinical neurology, general pediatric, and clinical genetics teams involved with their clinical care and support. We thank Toby Baldwin, Toni Saville, and Michael Buckley at SEALS genetics laboratory for their technical assistance. The research was supported by the National Health and Medical Research Council (GNT11149630 to E.E.P.; GNT0512123 to T.R.), Office of Health and Medical Research, NSW Health (to E.E.P.), King Abdulaziz City for Science and Technology (13-BIO1113-20 and 15-BIO3688-20 to F.S.A.), King Salman Center for Disability Research (F.S.A.), Saudi Human Genome Program (F.S.A.), and the King Abdullah University of Science and Technology through baseline funds (S.T.A., S.H., J.M., H.A., F.A., L.J., M.J.) and the Award No. FCC1/1976-25 from the Office of Sponsored Research (S.T.A.). The Department of Molecular and Human Genetics at Baylor College of Medicine receives revenue from clinical genetic testing done at Baylor Genetics Laboratory.

Declaration of Interests

The authors declare no competing interests.

Received: January 8, 2018
Accepted: January 23, 2019
Published: February 28, 2019

Web Resources

BRAVO, https://bravo.sph.umich.edu/freeze5/hg38/
CADD, http://cadd.gs.washington.edu/
ClinVar, https://www.ncbi.nlm.nih.gov/clinvar/
dbSNP, https://www.ncbi.nlm.nih.gov/SNP/
DECIPHER, https://decipher.sanger.ac.uk/
Ensembl, https://www.ensembl.org/index.html
ExAC database, http://exac.broadinstitute.org
gnomAD database, http://gnomad.broadinstitute.org/
Human Disease Gene Webseries, http://humandiseasegenes.nl/
Mendelian Inheritance in Man, http://www.omim.org
PROVEAN, http://provean.jcvi.org
SPARKY, https://www.cgl.ucsf.edu/home/sparky
UCSC Genome Browser, http://genome.ucsc.edu

References

1. Boycott, K.M., Rath, A., Chong, J.X., Hartley, T., Alkuraya, F.S., Baynam, G., Brookes, A.J., Brudno, M., Carracedo, A., den Dunnen, J.T., et al. (2017). International cooperation to enable the diagnosis of all rare genetic diseases. Am. J. Hum. Genet. *100*, 695–705.
2. Palmer, E.E., Kumar, R., Gordon, C.T., Shaw, M., Hubert, L., Carroll, R., Rio, M., Murray, L., Leffler, M., Dudding-Byth, T., et al.; DDD Study (2017). A Recurrent de novo nonsense variant in ZSWIM6 results in severe intellectual disability without frontonasal or limb malformations. Am. J. Hum. Genet. *101*, 995–1005.
3. Patel, N., Faqeih, E., Anazi, S., Alfawareh, M., Wakil, S.M., Colak, D., and Alkuraya, F.S. (2015). A novel APC mutation defines a second locus for Cenani-Lenz syndrome. J. Med. Genet. *52*, 317–321.
4. Aldahmesh, M.A., Mohamed, J.Y., Alkuraya, H.S., Verma, I.C., Puri, R.D., Alaiya, A.A., Rizzo, W.B., and Alkuraya, F.S. (2011). Recessive mutations in ELOVL4 cause ichthyosis, intellectual disability, and spastic quadriplegia. Am. J. Hum. Genet. *89*, 745–750.
5. Alkuraya, F.S. (2016). Discovery of mutations for Mendelian disorders. Hum. Genet. *135*, 615–623.
6. Monies, D., Maddirevula, S., Kurdi, W., Alanazy, M.H., Alkhalidi, H., Al-Owain, M., Sulaiman, R.A., Faqeih, E., Goljan, E., Ibrahim, N., et al. (2017). Autozygosity reveals recessive mutations and novel mechanisms in dominant genes: implications in variant interpretation. Genet. Med. *19*, 1144–1150.
7. Wang, L., and Tsai, C.C. (2008). Atrophin proteins: An overview of a new class of nuclear receptor corepressors. Nucl. Recept. Signal. *6*, e009.
8. Shen, Y., Lee, G., Choe, Y., Zoltewicz, J.S., and Peterson, A.S. (2007). Functional architecture of atrophins. J. Biol. Chem. *282*, 5037–5044.
9. Zhang, F., Xu, D., Yuan, L., Sun, Y., and Xu, Z. (2014). Epigenetic regulation of Atrophin1 by lysine-specific demethylase 1 is required for cortical progenitor maintenance. Nat. Commun. *5*, 5815.
10. Wood, J.D., Nucifora, F.C., Jr., Duan, K., Zhang, C., Wang, J., Kim, Y., Schilling, G., Sacchi, N., Liu, J.M., and Ross, C.A. (2000). Atrophin-1, the dentato-rubral and pallido-luysian atrophy gene product, interacts with ETO/MTG8 in the nuclear matrix and represses transcription. J. Cell Biol. *150*, 939–948.
11. Onodera, O., Oyake, M., Takano, H., Ikeuchi, T., Igarashi, S., and Tsuji, S. (1995). Molecular cloning of a full-length cDNA for dentatorubral-pallidoluysian atrophy and regional expressions of the expanded alleles in the CNS. Am. J. Hum. Genet. *57*, 1050–1060.
12. Nagafuchi, S., Yanagisawa, H., Ohsaki, E., Shirayama, T., Tadokoro, K., Inoue, T., and Yamada, M. (1994). Structure and expression of the gene responsible for the triplet repeat disorder, dentatorubral and pallidoluysian atrophy (DRPLA). Nat. Genet. *8*, 177–182.
13. Koide, R., Ikeuchi, T., Onodera, O., Tanaka, H., Igarashi, S., Endo, K., Takahashi, H., Kondo, R., Ishikawa, A., Hayashi, T., et al. (1994). Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). Nat. Genet. *6*, 9–13.
14. Hasegawa, A., Ikeuchi, T., Koike, R., Matsubara, N., Tsuchiya, M., Nozaki, H., Homma, A., Idezuka, J., Nishizawa, M., and Onodera, O. (2010). Long-term disability and prognosis in dentatorubral-pallidoluysian atrophy: a correlation with CAG repeat length. Mov. Disord. *25*, 1694–1700.
15. Veneziano, L., and Frontali, M. (1993). Drpla. In GeneReviews, M.P. Adam, H.H. Ardinger, R.A. Pagon, S.E. Wallace, L.J.H.

Bean, K. Stephens, and A. Amemiya, eds. (Seattle, WA: University of Washington).
16. Ross, C.A., and Poirier, M.A. (2004). Protein aggregation and neurodegenerative disease. Nat. Med. *10* (*Suppl*), S10–S17.
17. Nucifora, F.C., Jr., Ellerby, L.M., Wellington, C.L., Wood, J.D., Herring, W.J., Sawa, A., Hayden, M.R., Dawson, V.L., Dawson, T.M., and Ross, C.A. (2003). Nuclear localization of a non-caspase truncation product of atrophin-1, with an expanded polyglutamine repeat, increases cellular toxicity. J. Biol. Chem. *278*, 13047–13055.
18. Okamura-Oho, Y., Miyashita, T., Nagao, K., Shima, S., Ogata, Y., Katada, T., Nishina, H., and Yamada, M. (2003). Dentatorubral-pallidoluysian atrophy protein is phosphorylated by c-Jun NH2-terminal kinase. Hum. Mol. Genet. *12*, 1535–1542.
19. Fischbeck, K.H. (1997). Kennedy disease. J. Inherit. Metab. Dis. *20*, 152–158.
20. Thomsen, M.C., and Nielsen, M. (2012). Seq2Logo: a method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. Nucleic Acids Res. *40*, W281-7.
21. Anazi, S., Maddirevula, S., Faqeih, E., Alsedairy, H., Alzahrani, F., Shamseldin, H.E., Patel, N., Hashem, M., Ibrahim, N., Abdulwahab, F., et al. (2017). Clinical genomics expands the morbid genome of intellectual disability and offers a high diagnostic yield. Mol. Psychiatry *22*, 615–624.
22. Mosca, A.L., Laurent, N., Guibaud, L., Callier, P., Thauvin-Robinet, C., Mugneret, F., Huet, F., Grimaldi, M., Gouyon, J.B., Sandre, D., and Faivre, L. (2007). Polymicrogyria, cerebellar vermis hypoplasia, severe facial dysmorphism and cleft palate: a new syndrome? Eur. J. Med. Genet. *50*, 48–53.
23. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature *536*, 285–291.
24. Ng, P.C., and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res. *31*, 3812–3814.
25. Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., and Chan, A.P. (2012). Predicting the functional effect of amino acid substitutions and indels. PLoS ONE *7*, e46688.
26. Quang, D., Chen, Y., and Xie, X. (2015). DANN: A deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics *31*, 761–763.
27. Rentzsch, P., Witten, D., Cooper, G.M., Shendure, J., and Kircher, M. (2019). CADD: Predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. *47* (D1), D886–D894.
28. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al.; ACMG Laboratory Quality Assurance Committee (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. *17*, 405–424.
29. Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kübler, J., Lozajic, M., Gabler, F., Söding, J., Lupas, A.N., and Alva, V. (2018). A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J. Mol. Biol. *430*, 2237–2243.
30. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR *6*, 277–293.
31. Wüthrich, K. (1986). NMR with proteins and nucleic acids. Europhys. News *17*, 11–13.
32. Witkowska, D., Rowinska-Zyrek, M., Valensin, G., and Kozlowski, H. (2012). Specific poly-histidyl and poly-cysteil protein sites involved in Ni^{2+} homeostasis in Helicobacter pylori. Impact of Bi^{3+} ions on Ni^{2+} binding to proteins. Structural and thermodynamic aspects. Coord. Chem. Rev. *256*, 133–148.
33. Remelli, M., Brasili, D., Guerrini, R., Pontecchiani, F., Potocki, S., Rowinska-Zyrek, M., Watly, J., and Kozlowski, H. (2017). Zn(II) and Ni(II) complexes with poly-histidyl peptides derived from a snake venom. Inorg. Chim. Acta *472*, 149–156.
34. Zhang, C.L., Zou, Y., Yu, R.T., Gage, F.H., and Evans, R.M. (2006). Nuclear receptor TLX prevents retinal dystrophy and recruits the corepressor atrophin1. Genes Dev. *20*, 1308–1320.
35. Janknecht, R., Sander, C., and Pongs, O. (1991). (HX)n repeats: a pH-controlled protein-protein interaction motif of eukaryotic transcription factors? FEBS Lett. *295*, 1–2.
36. Fregeau, B., Kim, B.J., Hernández-García, A., Jordan, V.K., Cho, M.T., Schnur, R.E., Monaghan, K.G., Juusola, J., Rosenfeld, J.A., Bhoj, E., et al. (2016). De Novo mutations of RERE cause a genetic syndrome with features that overlap those associated with proximal 1p36 deletions. Am. J. Hum. Genet. *98*, 963–970.
37. Kaur, M., Izumi, K., Wilkens, A.B., Chatfield, K.C., Spinner, N.B., Conlin, L.K., Zhang, Z., and Krantz, I.D. (2014). Genome-wide expression analysis in fibroblast cell lines from probands with Pallister Killian syndrome. PLoS ONE *9*, e108853.
38. Izumi, K., and Krantz, I.D. (2014). Pallister-Killian syndrome. Am. J. Med. Genet. C. Semin. Med. Genet. *166C*, 406–413.
39. Shi, Z., Fujii, K., Kovary, K.M., Genuth, N.R., Röst, H.L., Teruel, M.N., and Barna, M. (2017). Heterogeneous ribosomes preferentially translate distinct subpools of mRNAs genome-wide. Mol. Cell *67*, 71–83.e7.
40. Galan, A., Lozano, G., Piñeiro, D., and Martinez-Salas, E. (2017). G3BP1 interacts directly with the FMDV IRES and negatively regulates translation. FEBS J. *284*, 3202–3217.
41. Kondrashov, N., Pusic, A., Stumpf, C.R., Shimizu, K., Hsieh, A.C., Ishijima, J., Shiroishi, T., and Barna, M. (2011). Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. Cell *145*, 383–397.

The American Journal of Human Genetics, Volume *104*

Supplemental Data

De Novo Variants Disrupting the HX Repeat Motif of ATN1 Cause a Recognizable Non-Progressive Neurocognitive Syndrome

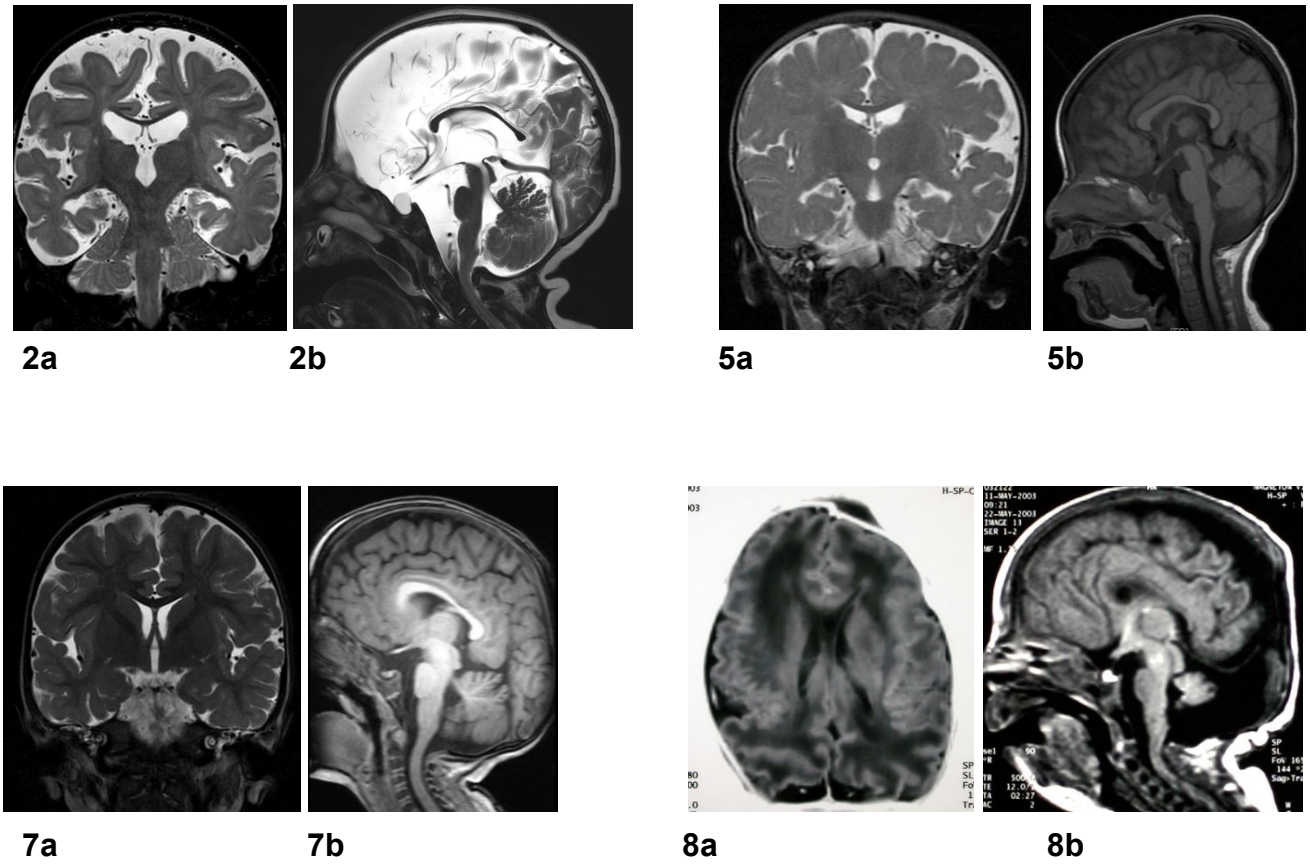
Elizabeth E. Palmer, Seungbeom Hong, Fatema Al Zahrani, Mais O. Hashem, Fajr A. Aleisa, Heba M. Jalal Ahmed, Tejaswi Kandula, Rebecca Macintosh, Andre E. Minoche, Clare Puttick, Velimir Gayevskiy, Alexander P. Drew, Mark J. Cowley, Marcel Dinger, Jill A. Rosenfeld, Rui Xiao, Megan T. Cho, Suliat F. Yakubu, Lindsay B. Henderson, Maria J. Guillen Sacoto, Amber Begtrup, Muddathir Hamad, Marwan Shinawi, Marisa V. Andrews, Marilyn C. Jones, Kristin Lindstrom, Ruth E. Bristol, Saima Kayani, Molly Snyder, María Mercedes Villanueva, Angeles Schteinschnaider, Laurence Faivre, Christel Thauvin, Antonio Vitobello, Tony Roscioli, Edwin P. Kirk, Ann Bye, Jasmeen Merzaban, Łukas Jaremko, Mariusz Jaremko, Rani K. Sachdev, Fowzan S. Alkuraya, and Stefan T. Arold

Supplemental Case Reports

Detailed phenotypic information on all affected individuals is provided in Table S1**.**

Supplemental Figures and Legends

Figure S1:



Legend: Neuroimaging of affected individuals with *de novo ATN1* variants.

Common features between affected individual include thinning of the corpus callosum with low hanging medial parietal lobule which results in anomalous course of the Vein of Galen, polymicrogyria of the sylvian fissure and steep clival angle. Affected individuals exhibit varying degrees of parenchymal atrophy, which is more pronounced in the Sylvian fissures. Of the four individuals where the neuroimaging was available two had clear absence of the falx cerebri and the other two had partial absence.

2a: Coronal T2 MRI of individual 2 illustrating absence of falx cerebri and peri-sylvian polymicrogyria. This patient also has significant parenchymal atrophy with widening of sulci; 2b: Sagittal T2 MRI of individual 2 showing steep clival angle, thin corpus callosum, abnormal course of Vein of Galen and parenchymal atrophy.

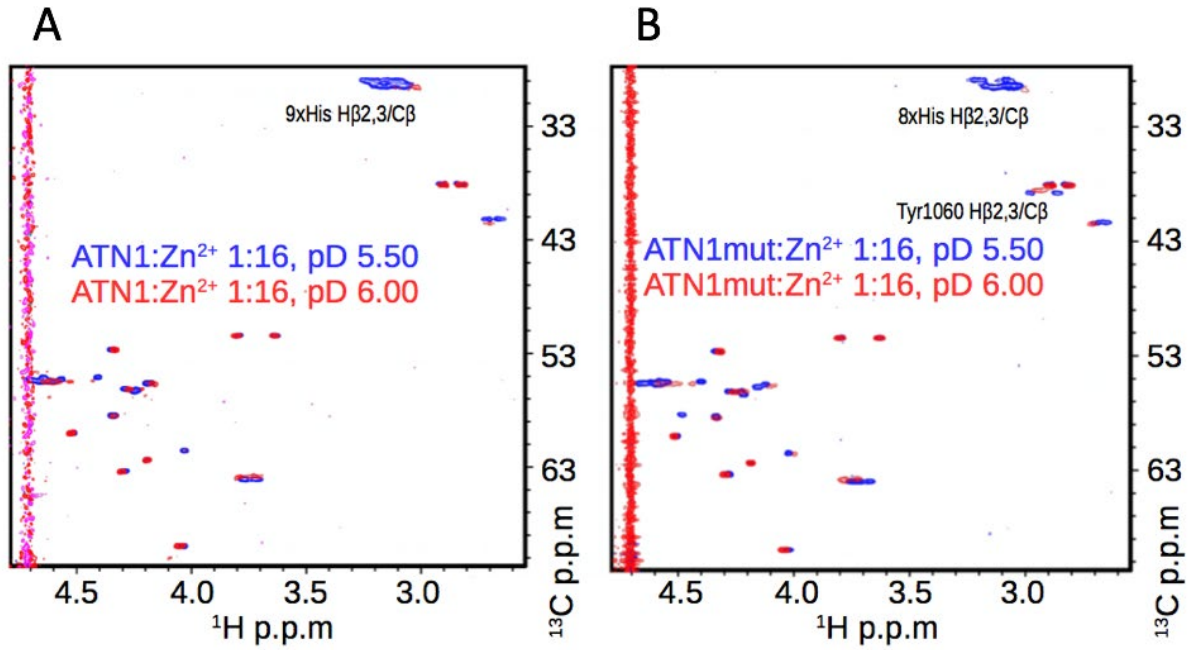
5a: Coronal T2 MRI of individual 5 showing peri-sylvian polymicrogyria, parenchymal atrophy as illustrated by widened Sylvian fissure and sulci and partial absence of falx cerebri;

5b: Sagittal T1 MRI illustrating thinned corpus callosum and abnormal course of Vein of Galen in Individual 5.

7a: Coronal T2 MRI illustrating absence of falx cerebri and peri-sylvian polymicrogyria in individual 7; 7b: Sagittal T1 MRI of individual 7 showing steep clival angle, thin corpus callosum, and abnormal course of Vein of Galen.

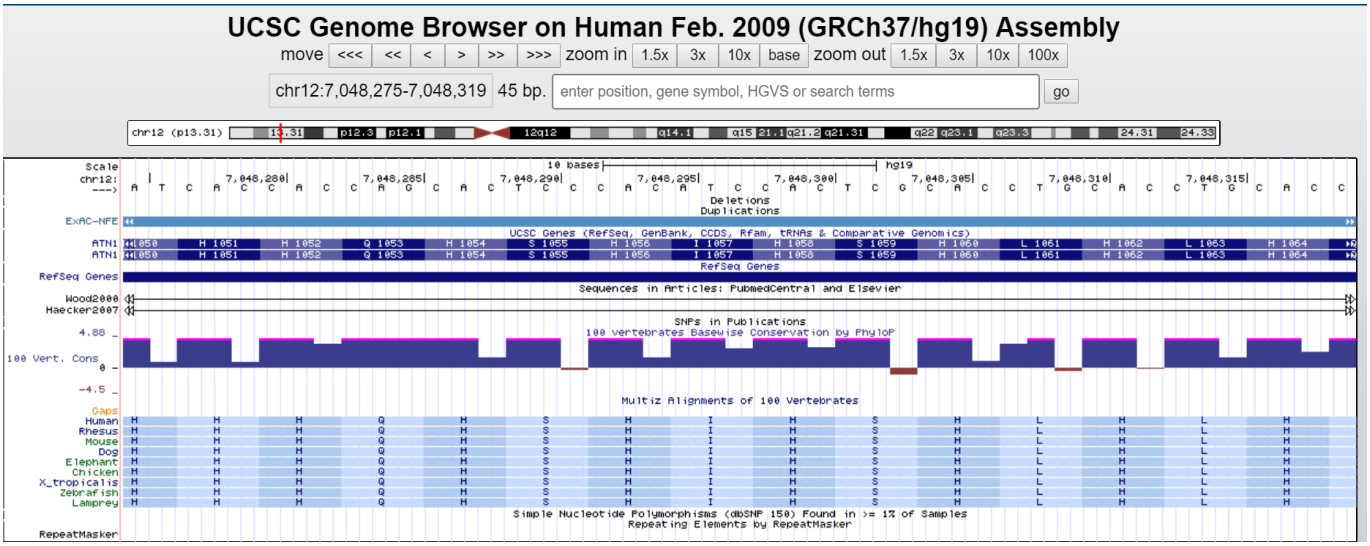
8a: Axial T1 MRI of individual 8 demonstrates polymicrogyria of the right Sylvian fissure; 8b Sagittal T1 MRI of individual 8 demonstrates thin corpus callosum and low hanging medial parietal gyrus. It is difficult to identify the vein of Galen. There is a steep clival angle. This individual exhibits vermian atrophy of the cerebellum.

Figure S2:



Legend: 2D heteronuclear ^{1}H - ^{13}C HSQC correlation spectra. Peptides were recorded using natural abundance of ^{1}H and ^{13}C in synthesized peptides. ATN1: 1046-NVTPHHHQHSHIHSHLHLHQQD-1067 ($ATN1_{1046-1067}$); ATN1mut: 1046-NVTPHHHQHSHIHSYLHLHQQD-1067 ($ATN1_{1046-1067}^{His1060Tyr}$). pD corresponds to the pH in D_2O. A-B: Comparison of Zn^{2+} interaction at pD of 5.5 and 6.0, at a peptide: Zn^{2+} ion ratio of 1:16

Figure S3:



Legend: Evolutionary conservation of the HX domain. Amino acids 150-164 in ATN1 are highly evolutionary conserved across species. Screenshot UCSC Genome Browser.

Supplemental Methods

Next generation sequencing and variant analysis

Individuals 1, 5 and 8 were screened by ES using previously described methodology[1-3].

Sequences of individuals 2, 3, 4, and 6 were captured using the SureSelect Human All Exon V4 (50 Mb), the Clinical Research Exome kit (Agilent Technologies, Santa Clara, CA) or the IDT xGen Exome Research Panel v1.0. Massively parallel (NextGen) sequencing was done on an Illumina system with 100bp or greater paired-end reads. Additional sequencing technology and variant interpretation protocols have been previously described[4].

Individual 7 was screened for sequence and structural variants by WGS using previously described methodology[5; 6] as well for variants in the mitochondrial genome using the in-house pipeline mity: mity has been designed to detect low heteroplasty SNPs and INDELs in the mitochondrial genome from a blood sample.

In all cases, the variants were confirmed and segregation analysis performed by Sanger sequencing using standard methodology**.**

Pattern Search for $(HX)_8$ motifs in the human proteome

Data were obtained using the H-x-H-x-H-x-H-x-H-x-H-x-H-x-H pattern in the PatternSearch program[7]. The initial list of 71 proteins was hand-curated to retain only HX sequences reminiscent to ATN1. Only the HX repeat of each protein sequence is shown.

```
>NP_001007027.1 atrophin-1
----PHHHQHSHIHSHLHLHQ------
>NP_001036146.1 arginine-glutamic acid dipeptide repeats protein isoform a
----PHHHQHSHIHSHLHLHQ------
>NP_001036147.1 arginine-glutamic acid dipeptide repeats protein isoform b
----PHHHQHSHIHSHLHLHQ------
>NP_001098549.2 probable fibrosin-1 *
----FHQHNHQHQHTHQHTHQHF----
>NP_001120703.1 autism susceptibility gene 2 protein isoform 2 *
----FHQHQHQHQHTHQHTHQHT----
>NP_001120729.1 zinc transporter ZIP10 precursor  *
----GHDHSHQHAHGHGHSHGHE------
>NP_001136113.1 fibrosin-1-like protein
----FHQHQHTHQHTHQHTHQHQHT--
>NP_001138356.1 zinc transporter 7
----AHSHDHAHGHGHFHSH-------
>NP_001931.2 atrophin-1
----PHHHQHSHIHSHLHLHQ------
>NP_036234.3 arginine-glutamic acid dipeptide repeats protein isoform a
----PHHHQHSHIHSHLHLHQ------
>NP_056385.1 autism susceptibility gene 2 protein isoform 1
----FHQHQHQHQHTHQHTHQHT----
>NP_065075.1 zinc transporter ZIP10 precursor
----GHDHSHQHAHGHGHSHGHE------
>NP_598003.2 zinc transporter 7
----AHSHDHAHGHGHFHSHD------
>XP_005246746.2 PREDICTED: zinc transporter ZIP10 isoform X1
----GHDHSHQHAHGHGHSHGHE------
>XP_005250314.1 PREDICTED: autism susceptibility gene 2 protein isoform X8
----FHQHQHQHQHTHQHTHQHT----
>XP_005263521.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X1
----PHHHQHSHIHSHLHLHQ------
>XP_005263523.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X3
----PHHHQHSHIHSHLHLHQ------
>XP_005266228.1 PREDICTED: fibrosin-1-like protein isoform X3
```

----FHQHQHTHQHTHQHTHQHQHT--
>XP_005266230.1 PREDICTED: fibrosin-1-like protein isoform X4
----FHQHQHTHQHTHQHTHQHQHT--
>XP_005266232.1 PREDICTED: fibrosin-1-like protein isoform X9
----FHQHQHTHQHTHQHTHQHQHT--
>XP_005266234.1 PREDICTED: fibrosin-1-like protein isoform X11
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011509806.1 PREDICTED: zinc transporter ZIP10 isoform X1
----GHDHSHQHAHGHGHSHGHE------
>XP_011509807.1 PREDICTED: zinc transporter ZIP10 isoform X1
----GHDHSHQHAHGHGHSHGHE------
>XP_011509808.1 PREDICTED: zinc transporter ZIP10 isoform X1
----GHDHSHQHAHGHGHSHGHE------
>XP_011509809.1 PREDICTED: zinc transporter ZIP10 isoform X2
----GHDHSHQHAHGHGHSHGHE------
>XP_011514312.1 PREDICTED: autism susceptibility gene 2 protein isoform X1
----FHQHQHQHQHTHQHTHQHT----
>XP_011514313.1 PREDICTED: autism susceptibility gene 2 protein isoform X2
----FHQHQHQHQHTHQHTHQHT----
>XP_011514314.1 PREDICTED: autism susceptibility gene 2 protein isoform X3
----FHQHQHQHQHTHQHTHQHT----
>XP_011514315.1 PREDICTED: autism susceptibility gene 2 protein isoform X4
----FHQHQHQHQHTHQHTHQHT----
>XP_011514316.1 PREDICTED: autism susceptibility gene 2 protein isoform X5
----FHQHQHQHQHTHQHTHQHT----
>XP_011514319.1 PREDICTED: autism susceptibility gene 2 protein isoform X6
----FHQHQHQHQHTHQHTHQHT----
>XP_011514320.1 PREDICTED: autism susceptibility gene 2 protein isoform X7
----FHQHQHQHQHTHQHTHQHT----
>XP_011533105.1 PREDICTED: fibrosin-1-like protein isoform X1
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533106.1 PREDICTED: fibrosin-1-like protein isoform X2
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533107.1 PREDICTED: fibrosin-1-like protein isoform X5
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533108.1 PREDICTED: fibrosin-1-like protein isoform X6
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533110.1 PREDICTED: fibrosin-1-like protein isoform X8
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533112.1 PREDICTED: fibrosin-1-like protein isoform X13
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533114.1 PREDICTED: fibrosin-1-like protein isoform X14
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011533118.1 PREDICTED: fibrosin-1-like protein isoform X15
----FHQHQHTHQHTHQHTHQHQHT--
>XP_011539081.1 PREDICTED: zinc transporter 7 isoform X2
----AHSHDHAHGHGHFHSHD------
>XP_011539812.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X2 [Homo
sapiens
----PHHHQHSHIHSHLHLHQ------
>XP_011544218.1 PREDICTED: probable fibrosin-1 isoform X1
----FHQHNHQHQHTHQHTHQHF----
>XP_011544219.1 PREDICTED: probable fibrosin-1 isoform X2
----FHQHNHQHQHTHQHTHQHF----
>XP_011544221.1 PREDICTED: probable fibrosin-1 isoform X3
----FHQHNHQHQHTHQHTHQHF----
>XP_016855889.1 PREDICTED: zinc transporter 7 isoform X1
----AHSHDHAHGHGHFHSHD------
>XP_016855890.1 PREDICTED: zinc transporter 7 isoform X3
----AHSHDHAHGHGHFHSHD------
>XP_016856847.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X1
----PHHHQHSHIHSHLHLHQ------
>XP_016856848.1 PREDICTED: arginine-glutamic acid dipeptide repeats protein isoform X1
----PHHHQHSHIHSHLHLHQ------
>XP_016860011.1 PREDICTED: zinc transporter ZIP10 isoform X2
----GHDHSHQHAHGHGHSHGHE------
>XP_016867440.1 PREDICTED: autism susceptibility gene 2 protein isoform X9
----FHQHQHQHQHTHQHTHQHT----

NMR Analysis, extended description

Free ATN1 and ATN1mut in solution

To probe the impact of the patient variants on the molecular behavior of the ATN1 protein we studied two polypeptides containing the $(HX)_n$ motif (n=8) in solution by NMR techniques - one of the native sequence 1046-NVTPHHHQHSHIHS**H**LHLHQQD-1067 (ATN1) and one bearing the pathological variant - with the 1060th histidine exchanged to tyrosine - 1046-NVTPHHHQHSHIHS**Y**LHLHQQD-1067 (ATN1mut). The peptides were synthesized using the solid-phase peptide synthesis (SPPS). Peptides were dissolved in 500 µL of 100% D_2O up to the concentration of 2.2 mg/mL. The pH electrode read-out was manually adjusted to 5.10 (what corresponds to pD 5.51, according to Covington *et al*.,[8]) with the stock solution of 0.1 M NaOD in 100% D_2O. All NMR experiments were performed on the 700 MHz spectrometer at 25ºC. The NMR data was processed by NMRPipe[9] and analyzed with Sparky software (https://www.cgl.ucsf.edu/home/sparky/). Analysis of the ^1H and ^{13}C resonances for the two peptides was done by application of a standard procedure[10] based on the inspection of the 2D homonuclear 2D ^1H-^1H TOCSY (with mixing times 10 and 80 ms) and 2D ^1H-^1H ROESY as well as 2D ^1H-^1H NOESY (with mixing times 300 and 500 ms) supported by the natural abundance 2D ^1H-^{13}C HSQC (separately tuned for the aliphatic and aromatic regions) experiments.

Despite the fact that the primary sequences of peptides contain several HX repetitions the 2D heteronuclear ^1H-^{13}C HSQC correlation spectra of the aliphatic regions show decent signal dispersion **(Fig. 2E)**. This made the residue specific resonance assignment possible either for ATN1 as well as ATN1mut peptide. The peptide main chain $^1H\alpha$, $^{13}C\alpha$ chemical shifts showed that the free peptides adopt highly similar disordered conformation in solution. On the 2D ^1H-^{13}C HSQC spectra of the aromatic region coming from the wild-type ATN1 the cross-peaks corresponding to the imidazole Hδ2/Cδ2 and Hε1/Cε1 atoms of all histidines are clustered around one position - indicating similar averaged local environment as well as dynamical properties for each particular histidine side-chain moiety **(Fig. 2F)**. In contrast to the wild-type ATN1 the Hδ2/Cδ2 and Hε1/Cε1 atoms from the imidazole moieties of histidines within the mutated peptide ATN1mut contribute to the larger dispersion of the cross-peaks of the side-chain aromatic region **(Fig. 2F)**. This indicates that the side-chains from these residues are not "synchronized" thus experience their different local environments and mobility than in case of the wild-type ATN1. To further understand the origins and nature of the observed histidine side-chains "synchronization" phenomenon the 2D homonuclear ^1H-^1H ROESY and 2D ^1H-^1H NOESY (with mixing times 100, 300 and 500 ms) spectra, that provide information

about close contacts (up to 5.50 Å) between the protons through space, were recorded. In line with the low signal dispersion in the aromatic region for the wild-type ANT1 the trivial (|j-i| = 0) and short range (|j-i| < 2) NOE cross-peaks are not present or of substantially lower intensity than those observed for the analogical residues within ATN1mut. This observation indicates an increased side-chain mobility for the residues of ATN1 wild-type in comparison to the ATN1mut. Both peptides precipitate from the solution once pD exceeds the value of 6.00 due to deprotonation of the imidazole rings of histidines.

Titrations of ATN1 and ATN1mut with Zn^{2+} ions

The side-chain imidazole rings of histidines are known to coordinate metal ions [11; 12]. To decipher the impact of the pathological variant on the coordination of the metal ions we monitored the Zn^{2+} ions binding to the ATN1 and ANT1mut peptides by NMR. Measurements with Zn^{2+} ions were performed by the stepwise addition of $ZnCl_2$ stock solution (500 mM in 100% D_2O) to the solution of peptides reaching the peptide:Zn^{2+} molar ratios of 1:0.5, 1:1, 1:4, 1:8, 1:16, 1:32 and 1:48 respectively. For each peptide:Zn^{2+} ratio the pD was checked and corrected with the 0.1 M stock solution of NaOD (if needed) and the same set of 1D and 2D NMR spectra as for the free peptides was recorded.

At pD of 5.50 both peptides ATN1 and ATN1mut bind Zn^{2+} ions via their N-termini as indicated by the chemical shift changes of the N-terminal asparagine $^{1}H\alpha/^{13}C\alpha$ **(Fig. 2G,H)** and $^{1}H\beta2,3/^{13}C\beta$ signals. In line with the "synchronized" behavior of the histidine imidazoles the wild type ATN1 histidines did not bind Zn^{2+} but the ones from ATN1mut do bind as demonstrated by histidine $^{1}H\beta2,3/^{13}C\beta$ chemical shift and signal intensity changes upon addition of Zn^{2+} ions **(Fig. S2D)**. Increasing the pD up to 6.00 results in progressive deprotonation of histidine imidazoles, thus both peptides bind the zinc ions **(Fig. S2A,B)**. Moving pD towards 6.40 results in precipitation of both peptides (no signal left with 1D ^{1}H NMR spectra) - most likely due to forming larger zinc-coordinated aggregates, oligomers or by simultaneous precipitation of the zinc hydroxide $Zn(OH)_2$, accompanied by precipitation of peptides containing numerous deprotonated hydrophobic imidazole moieties. Interestingly the ATN1mut stays soluble in solution up to pD ca. 0.30 unit higher than the wild-type ATN1.

The "synchronization" effect of the histidine side-chains within the ATN1 can only be explained by the crowding effect of the $(HX)_8$ motif, that "synchronizes" all histidine side-chains to behave similarly in the pH-dependent manner. This effect increases the imidazole side chain dynamics significantly decreasing the ability to bind metal ions, as demonstrated by overlapped NMR signal positions of the imidazole $H\delta2/C\delta2$ and $H\varepsilon1/C\varepsilon1$ correlations **(Fig.**

2F) and zinc titration **(Fig. 2G)**, respectively. Once one of the central histidines from the $(HX)_8$ motif is mutated to tyrosine, all remaining histidines display the behavior more typical to the isolated, not crowded/synchronized histidines, thus the ATN1mut peptide is able to bind zinc ions in an expected mode(s) and pH range. This makes the entire ATN1mut sequence more soluble in the presence of metal ions and at higher pD values closer to the physiological ones.

Supplemental References

1. Yang, Y., Muzny, D.M., Reid, J.G., Bainbridge, M.N., Willis, A., Ward, P.A., Braxton, A., Beuten, J., Xia, F., Niu, Z., et al. (2013). Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med 369, 1502-1511.
2. Alazami, A.M., Patel, N., Shamseldin, H.E., Anazi, S., Al-Dosari, M.S., Alzahrani, F., Hijazi, H., Alshammari, M., Aldahmesh, M.a., Salih, M.a., et al. (2014). Accelerating Novel Candidate Gene Discovery in Neurogenetic Disorders via Whole-Exome Sequencing of Prescreened Multiplex Consanguineous Families. Cell reports, 1-14.
3. Moutton, S., Bruel, A.L., Assoum, M., Chevarin, M., Sarrazin, E., Goizet, C., Guerrot, A.M., Charollais, A., Charles, P., Heron, D., et al. (2018). Truncating variants of the DLG4 gene are responsible for intellectual disability with marfanoid features. Clin Genet 93, 1172-1178.
4. Tanaka, A.J., Cho, M.T., Millan, F., Juusola, J., Retterer, K., Joshi, C., Niyazov, D., Garnica, A., Gratz, E., Deardorff, M., et al. (2015). Mutations in SPATA5 Are Associated with Microcephaly, Intellectual Disability, Seizures, and Hearing Loss. Am J Hum Genet 97, 457-464.
5. Palmer, E.E., Kumar, R., Gordon, C.T., Shaw, M., Hubert, L., Carroll, R., Rio, M., Murray, L., Leffler, M., Dudding-Byth, T., et al. (2017). A Recurrent De Novo Nonsense Variant in ZSWIM6 Results in Severe Intellectual Disability without Frontonasal or Limb Malformations. Am J Hum Genet 101, 995-1005.
6. Gayevskiy, V., Roscioli, T., Dinger, M.E., and Cowley, M.J. (2018). Seave: a comprehensive web platform for storing and interrogating human genomic variation. Bioinformatics, bty540-bty540.
7. Zimmermann, L., Stephens, A., Nam, S.Z., Rau, D., Kubler, J., Lozajic, M., Gabler, F., Soding, J., Lupas, A.N., and Alva, V. (2018). A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol 430, 2237-2243.
8. Covington, A.K., Paabo, M., Robinson, R.A., and Bates, R.G. (2002). Use of the glass electrode in deuterium oxide and the relation between the standardized pD (paD) scale and the operational pH in heavy water.
9. Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J., and Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293.
10. Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids.(New York: Wiley).
11. Remelli, M., Brasili, D., Guerrini, R., Pontecchiani, F., Potocki, S., Rowinska-Zyrek, M., Watly, J., and Kozlowski, H. (2018). Zn(II) and Ni(II) complexes with poly-histidyl peptides derived from a snake venom. Inorganica Chimica Acta 472, 149-156.
12. Witkowska, D., Rowinska-Zyrek, M., Valensin, G., and Kozlowski, H. (2012). Specific poly-histidyl and poly-cysteil protein sites involved in $Ni2^+$ homeostasis in H*elicobacter pylori.* Impact of $Bi3^+$ ions on $Ni2^+$ binding to proteins. Structural and thermodynamic aspects. Coordination Chemistry Reviews 256, 133-148.