ARTICLE

HNRNPR Variants that Impair Homeobox Gene Expression Drive Developmental Disorders in Humans

Floor A. Duijkers,[1] Andrew McDonald,[2,13] Georges E. Janssens,[2,13] Marco Lezzerini,[2,13] Aldo Jongejan,[3] Silvana van Koningsbruggen,[1] Wendela G. Leeuwenburgh-Pronk,[4] Marcin W. Wlodarski,[5] Sébastien Moutton,[6,7,8] Frédéric Tran-Mau-Them,[6,7] Christel Thauvin-Robinet,[6,7,8] Laurence Faivre,[6] Kristin G. Monaghan,[9] Thomas Smol,[10,11] Odile Boute-Benejean,[10,11] Roger L. Ladda,[12] Susan L. Sell,[12] Ange-Line Bruel,[6,7] Riekelt H. Houtkooper,[2] and Alyson W. MacInnes[2,*]

The heterogeneous nuclear ribonucleoprotein (*HNRNP*) genes code for a set of RNA-binding proteins that function primarily in the spliceosome C complex. Pathogenic variants in these genes can drive neurodegeneration, through a mechanism involving excessive stress-granule formation, or developmental defects, through mechanisms that are not known. Here, we report four unrelated individuals who have truncating or missense variants in the same C-terminal region of hnRNPR and who have multisystem developmental defects including abnormalities of the brain and skeleton, dysmorphic facies, brachydactyly, seizures, and hypoplastic external genitalia. We further identified in the literature a fifth individual with a truncating variant. RNA sequencing of primary fibroblasts reveals that these *HNRNPR* variants drive significant changes in the expression of several homeobox genes, as well as other transcription factors, such as *LHX9, TBX1,* and multiple *HOX* genes, that are considered fundamental regulators of embryonic and gonad development. Higher levels of retained intronic *HOX* sequences and lost splicing events in the *HOX* cluster are observed in cells carrying *HNRNPR* variants, suggesting that impaired splicing is at least partially driving *HOX* deregulation. At basal levels, stress-granule formation appears normal in primary and transfected cells expressing *HNRNPR* variants. However, these cells reveal profound recovery defects, where stress granules fail to disassemble properly, after exposure to oxidative stress. This study establishes an essential role for *HNRNPR* in human development and points to a mechanism that may unify other "spliceosomopathies" linked to variants that drive multi-system congenital defects and are found in hnRNPs.

Introduction

The heterogeneous ribonucleoprotein R (hnRNPR) protein is a part of a multi-megadalton complex, termed the spliceosome C, that functions in the nucleus to process and transport mRNA.[1] At least 20 abundant hnRNPs have been designated the "core" hnRNP proteins (hnRNPA1 through hnRNPU) on the basis of their association with nascent RNA transcripts.[2,3] These hnRNPs are increasingly viewed as multi-functional, their roles extending past splicing in the nucleus to operating at several levels of gene regulation.[4]

The localization of certain hnRNPs is not limited to the nucleus; some hnRNPs are observed shuttling in and out of the cytoplasm.[5,6] In the cytoplasm, some hnRNPs have been found localized in stress granules. These are dense aggregations of protein, RNAs, and stalled initiation complexes that appear when cells are exposed to acute oxidative stress.[7,8] Stress granules appear very rapidly upon cellular exposure to oxidative insult and serve to retain and protect mRNAs that are not immediately required for the stress response. When the insult is over, the stress granules dissipate, and the mRNAs are released and then free to associate with the translation machinery. Excessive formation of stress granules is a feature of cells carrying *HNRNPA1* and *HNRNPA2/B1* variants that drive degenerative diseases such as amyotrophic lateral sclerosis 1 (ALS1; MIM: 105400) and frontotemporal dementia (FTD; MIM: 600274).[9,10] These variants occur in the "prion-like domains" (PrLDs) of the hnRNPs. These domains are generally low in complexity, unstructured, and enriched in glycines and uncharged polar amino acids (especially arginine, glutamine, and tyrosine).[11] These PrLD-residing variants appear to aggravate the tendency of the hnRNPA1 and A2/B1 proteins to self-polymerize, leading to increases of pathologic cellular inclusions in the motor neurons of affected individuals.[9,10]

[1]Amsterdam University Medical Centers, University of Amsterdam, Department of Clinical Genetics, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands; [2]Amsterdam University Medical Centers, University of Amsterdam, Laboratory Genetic Metabolic Diseases, Amsterdam Gastroenterology and Metabolism, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands; [3]Amsterdam University Medical Centers, University of Amsterdam, Bioinformatics Laboratory, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands; [4]Amsterdam University Medical Centers, University of Amsterdam, Department of Pediatrics, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands; [5]Department of Pediatric Hematology and Oncology, University of Freiburg, D-79106 Freiburg, Germany; [6]Institut National de la Santé et de la Recherche Médicale UMR 1231 GAD, Génétique des Anomalies du Dévelopement, Université de Bourgogne-Franche Comté, F-21079 Dijon, France; [7]Fédération Hospitalo-Universitaire Médecine TRANSLationnelle et Anomalies du Développement, Centre Hospitalier Universitaire et Université de Bourgogne-Franche Comté, 21000 Dijon, France; [8]Centre de Génétique et Centre de Référence Anomalies du Développement et Syndromes Malformatifs de l'Inter-région Est, Centre Hospitalier Universitaire Dijon Bourgogne, F-21079 Dijon, France; [9]GeneDx, Gaithersburg, MD 20877, USA; [10]Université de Lille, EA 7364 - RADEME, F-59000 Lille, France; [11]Centre Hospitalier Universitaire Lille, Institut de Génétique Médicale, F-59000 Lille, France; [12]Department of Pediatrics, Penn State Children's Hospital, Hershey, PA 17033, USA
[13]These authors contributed equally to this work
*Correspondence: a.w.macinnes@amc.nl
https://doi.org/10.1016/j.ajhg.2019.03.024.

Beyond neurodegeneration, several variants in genes coding for hnRNPs have also been reported in humans with neurodevelopmental defects and syndromic features. Variants in *HNRNPK* are found in Au-Kline syndrome (AUKS; MIM: 616580); variants in *HNRNPH2* link to X-linked, syndromic, Bain type mental retardation (MRXSB; MIM: 300986); and variants in *HNRNPU* are linked to early infantile epileptic encephalopathy-54 (EIEE54; MIM: 617391). The clinical phenotypes described within these syndromes include intellectual disability (ID), seizures, facial dysmorphism, abnormalities of the skeleton, and in some cases, defects in the central nervous system, heart, and kidneys.[12–17] However, despite these indications that hnRNPs have a function in neurodevelopment, the pathophysiological role of the variants reported to be in the genes coding for these proteins is unknown.

Here, we report *de novo* truncating and missense variants in *HNRNPR* in four unrelated individuals presenting with overlapping neurodevelopmental phenotypes and dysmorphic features. Our functional analysis of cells carrying these variants reveals a deregulation of the expression of several homeobox and T-box genes with well-established roles in embryonic development, and it indicates that the full-length hnRNPR protein is required to maintain the tight regulation of these genes.

Material and Methods

Individuals Participating in this Study

Written informed consent was obtained from the affected individuals and/or their parents prior to inclusion in this study, which was performed in accordance with the ethical standards of the Declaration of Helsinki. All procedures were carried out according to the ethical standards of the Amsterdam Universitair Medische Centra (UMC), the Université de Bourgogne-Franche Comté, the Université de Lille, and Penn State Children's Hospital, as well as their respective national boards.

Whole-Exome Sequencing

Whole-exome sequencing (WES) was performed with a trio diagnostic approach (affected individual and both parents). Libraries were prepared with the Kapa HTP kit (Illumina), and in the Netherlands, capture was performed with the SeqCap EZ MedExome kit (Roche NimbleGen). In France and the USA, the Agilent CRE capture kit (Agilent Technologies) was used. Sequencing was done on an Illumina HiSeq2500 (USA) or HiSeq4000 (Netherlands and France) HTv4 with paired-end, 125 bp reads. The read alignment to GRCh37 (hg19) and the variant calling were done with a pipeline based on the Burrows-Wheeler Aligner BWA-MEM 0.7 and the Genome Analysis Toolkit (GATK) 3.3.0. The median coverage of the captured target region was at least 100×. Variant annotation and prioritizing were done with Cartagenia Bench Lab next-generation-sequencing (NGS) (Agilent Technologies). We excluded variants located outside the ± 6 nucleotide exons and intron/exon boundaries and variants with a minor allele frequency (MAF) of >1% in control databases, including dbSNP137; the 1000 Genomes Project (release of February, 2012); the Exome Variant Server (EVS), the National Heart, Lung, and Blood Institute (NHLBI) Exome Sequencing Project GO (ESP5400 release); and our in-house exome controls. Variants that fit with a *de novo* or recessive mode of inheritance were further analyzed.

Cell Lines and Cell Culture

Fibroblasts were obtained by a punch biopsy in the upper arm and maintained in DMEM (GIBCO) + 10% fetal calf serum (FCS) and 1% penicillin/streptomycin. Lymphoblastoid cell lines (LCLs) were derived from Epstein-Barr virus (EBV)-immortalization of peripheral mononuclear cells isolated from whole blood with Ficoll (GE Life Sciences) and grown in RPMI (GIBCO) containing 15% FCS, 1% L-glutamine, and 1% penicillin/streptomycin, as previously described.[18] HeLa cells were maintained in DMEM (GIBCO) + 10% FCS and 1% penicillin/streptomycin.

Plasmids

N-terminal, HA-tagged wild-type (GenBank: 001102398), and mutant hnRNPR genes were synthesized and cloned into pCDNA3.1(-) (Invitrogen) by Genscript Biotech, and the sequences were verified by Sanger sequencing.

HeLa Cell Transfection

HeLa cells were plated in a 6-well dish between 70%–80% confluency. The following day, HA-tagged hnRNPR plasmids from WT, c.1609dupG (p.Ala537Glyfs*10), or c.1652dupG (p.Pro552Serfs*34) were expressed by transfection with jetPRIME DNA and siRNA transfection reagent (Polyplus transfection, #114-15) according to the manufacturer's specifications. In brief, a transfection mix consisting of 2 μg DNA (unless otherwise specified), 4 μl jetPRIME reagent, and 200 μl jetPRIME buffer was used in each well. Cells were harvested after a 24 hr incubation with transfection mix (at 37° C) and processed accordingly.

Confocal Analysis

Fibroblasts or transfected HeLa cells were plated to 50% confluence on glass coverslips in a 24-well plate. The following day, cells were treated for 60 min with 1 mM sodium arsenite (Sigma #S7400) dissolved in DMEM + 10% FCS at 37°C. For recovery experiments, cells were washed 3× with 1× PBS before the medium was replaced with fresh DMEM +10% FCS. Cells were fixed for 10 min at room temperature (RT) with 2% paraformaldehyde (Sigma #P6148) in PBS and washed 2× with PBS. Cells were then permeabilized for 5 min at RT with 0.1% Triton X-100 (Sigma #T8787) in PBS and washed 3× with PBS. Cells were then blocked for 30 min at RT with 1% BSA (Sigma #A2153) in PBS, and then they were incubated 1:100 in blocking buffer at RT for 30 min with antibodies against G3BP1 (BD Transduction Laboratories #611126), TIA1 (Abcam #ab40693), hnRNPR (Atlas Antibodies #HPA026092), or HA (Santa Cruz Biotechnology #sc-57592). Cells were then incubated at RT for 30 min with anti-mouse or anti-rabbit immunoglobulin conjugated to biotin and diluted at 1:200 in blocking buffer (Dako #E0433 [mouse] or #E0432 [rabbit]), then they were washed2× with PBS and subjected to a final incubation in the dark for 30 min at RT with streptavidin conjugated to FITC (Thermo Fisher Scientific #11-4317-87), or Alexa Fluor 633 (Life Technologies #A21052 [mouse] or #A21070 [rabbit]) at a 1:200 dilution in blocking buffer. Cells were then washed 3× with PBS and mounted on slides with Prolong Gold Antifade Mountant (Thermo Fisher Scientific #P36935) and left to dry overnight in the dark at RT. Images were acquired with a Leica SP-8 confocal microscope. For quantifications, biological triplicates of each

condition were generated on different days and analyzed by generating 3×3 tiles at 40×objective three times from each slide.

Immunoblots

LCLs or trypsinized transfected HeLa cells were washed 1× with 1× PBS and resuspended in RIPA buffer containing cOmplete Mini EDTA-free Protease Inhibitor Cocktail (Sigma-Aldrich). Lysates were incubated in ice for 10 min and then centrifuged at 14,000 rpm for 10 min in a pre-chilled centrifuge to remove debris. Protein quantification of the supernatant was carried out with bicinchoninic acid (BCA) assay, and equal amounts of protein (20–30 μg) were loaded onto NuPAGE Novex 4%–12% Bis-Tris Gel 1.5 mm, 10 Well (Invitrogen). Resolved protein lysates were subsequently transferred to nitrocellulose membranes via iBlot 2 Transfer Stacks (Invitrogen) for 7 min at 14 V in iBlot 2 Dry Blotting System (Invitrogen). Membranes were blocked for 1 h in 5% dried milk in PBS-T at RT on a rocking platform before being probed with rabbit anti-hnRNPR polyclonal antibody at 1:1000 (Novusbio, NBP1-89676), HA tag antibody (4810) (Novus Biologicals, NBP2-43714), or mouse anti-βactin monoclonal antibody at 1:5000 (Sigma, A5441). Next, they were incubated with appropriate secondary antibodies, then subjected to chemifluorescence detection.

RNA Isolation

Total RNA was isolated with TRIzol reagent (Invitrogen) according to the manufacturer's instructions. RNA was next treated with RNase-Free DNase Set (QIAGEN) for DNA removal and then with RNeasy MinElute Cleanup Kit (QIAGEN) for RNA concentration and purification.

RNA Sequencing

Primary fibroblasts from two affected individuals (one male and one female) were compared with fibroblasts from two healthy controls (one male and one female). Samples were processed and sequenced at a depth of 50 million reads by Genome Scan for Illumina sequencing platform with the NEBNext Ultra Directional RNA Library Prep Kit (NEB #E7420) according to the manufacturer's instructions. In brief, rRNA was depleted from total RNA with the rRNA depletion kit (New England Biotech #E6310). After fragmentation of the rRNA-reduced RNA, a cDNA synthesis was performed in order to ligate the sequencing adapters and perform PCR amplification of the resulting product. The quality and yield after sample preparation was measured with the Fragment Analyzer. The size of the resulting products was consistent with the expected size distribution (a broad peak between 300–500 bp). Clustering and DNA sequencing with the Illumina cBot and HiSeq 4000 was performed according to the manufacturer's protocol with a concentration of 3.0 nM of DNA. HiSeq control software HCS v3.4.0 was used to run the HiSeq sequencing platform and perform image analysis, base calling, quality checks, and conversion of the data into fastq format with the Illumina data analysis pipeline RTA v2.7.7 and Bcl2fastq v2.17 tools.

The quality of the reads in the fastq files was confirmed with FastQC version 0.11.4. FastQC performs multiple quality tests and provides the user with warnings about possible problems with the raw data. Paired-end reads were mapped to the *H. sapiens* genome with HISAT2 version 2.1.0.[19] After successful mapping, the counts per gene were extracted from the .SAM files with HTSeq version 0.9.1.[20]

The statistical analysis was performed in R with DESeq2 version 1.16.1.[21] DESeq2 models the raw counts with the negative binomial distribution and normalizes between samples by calculating size factors using the median ratio method.[22] The statistical test is based on a generalized linear model (GLM) that uses the negative binomial model as an error distribution model. To more accurately model dispersion parameters, the trend of dispersion to abundance is taken into account via an empirical Bayes procedure. The significance of the GLM coefficients are determined by a Wald test.

Results

Five Unrelated Individuals Carry Variants in *HNRNPR*

We identified three unrelated individuals who have overlapping clinical characteristics, including developmental delay, microcephaly, seizures, facial dysmorphism, brachydactyly, and other congenital abnormalities, and a *de novo* truncating variant in *HNRNPR* (GenBank: NM_001102398.2). One individual (P1) harbors a c.1609dupG (p.Ala537Glyfs*10) variant, and two individuals (P2 and P3) have the same c.1652dupG (p.Pro552Serfs*34) variant. Additionally, we identified a fourth individual (P4) who has developmental delay, microcephaly, facial dysmorphism, brachydactyly, and other congenital abnormalities and harbors a c.1763G>A (p.Arg588His) missense variant in *HNRNPR*. We then received some clinical information regarding a previously reported individual (P5) who has epileptic encephalopathy and a truncating variant (c.1663C>T [p.Gln555Ter]) in the last exon of *HNRNPR*.[23] We were able to surmise that the neurodevelopmental phenotype and brain MR imaging of individual P5 revealed developmental delay, corpus callosum agenesis, and cerebellar vermis hypoplasia, all of which are quite comparable to P2 and P4 in our study (Ambry Genetics, personal communication). Although the clinical information we were able to gather on this individual is limited, the similar C-terminal truncating variant in *HNRNPR* coupled to the similar clinical features that we were able to obtain leads us to believe that the genotype-phenotype correlation of this individual matches the others in this study. All individuals were born after a normal pregnancy and have healthy, non-consanguineous parents. The clinical information of individuals P1–P5 is summarized in Table 1. A full description of the clinical phenotypes is located in the Supplemental Note. Contact between the clinics was established through an entry for *HNRNPR* in GeneMatcher.[24]

Three individuals described in this paper and one individual described in the literature[23] have a truncating variant in the last exon of *HNRNPR*, and one individual harbors a missense variant in the same exon. None of these variants have been previously reported in variant population databases such as ExAC or gnomAD. The p.Arg588His variant substitutes a conserved arginine residue and is predicted to be "probably damaging," "not tolerated," or "disease causing" by PolyPhen-2, MutPred2, SIFT, and Mutation Taster.[25–28] The CADD score for the p.Arg588His

Table 1. Clinical Features of Individuals Carrying *HNRNPR* Variants

	This Study				Previously Published Case[23]
	Individual 1 (P1)	**Individual 2 (P2)**	**Individual 3 (P3)**	**Individual 4 (P4)**	**Individual 5 (P5)**
hnRNPR variant	c.1609dupG (p.Ala537Glyfs*10)	c.1652dupG (p.Pro552Serfs*34)	c.1652dupG (p.Pro552Serfs*34)	c. 1763G>A (p.Arg588His)	c.1663C>T, (p.Gln555Ter)
Sex	f	m	f	m	NA
Age at last follow-up	12 years	3 years	10 years	8 years	infant
Height (percentile)	140 cm (90^{th})	83 cm ($<1^{st}$)	112.4 cm (3^{rd})	101 cm ($<1^{st}$)	NA
Weight (weight per height percentile)	43 kg (95^{th})	10.6 kg (10^{th})	23.8 kg (95^{th})	13.4 kg ($<1^{st}$)	NA
Head circumference (percentile)	49 cm ($<1^{st}$; 50^{th} at birth)	44 cm ($<1^{st}$; 30^{th} at birth)	44,5 cm (3^{rd}; unknown at birth)	48 cm ($<1^{st}$; 50^{th} at birth)	NA
Neurological Features					
Developmental delay	severe	moderate to severe	moderate (IQ 52)	severe	+
Seizures	+ (febrile)	+	+	−	+ (epileptic encephalopathy)
Generalized hypotonia	+	−	+	+	NA
Brain abnormalities; magnetic resonance imaging	short corpus callosum (conserved proportions with microcephaly)	hypoplastic corpus callosum, frontal atrophy, cerebellar vermis hypoplasia	microcephaly	short corpus callosum, cerebellar vermis hypoplasia, microcephaly	corpus callosum agenesis, cerebellar vermis hypoplasia
Behavior	Stereotypic movements, attention deficit disorder	Stereotypic movements, attention deficit disorder	Stereotypic movements, pervasive developmental disorder	−	NA
Cardiac defects	−	VSD	−	−	NA
Urogenital defects	−	cryptorchidism, micropenis	hypoplastic labia	−	NA
Facial dysmorphism	+	+	+	+	mild
Unusual head hair 1^{st} year of life	+	+	+	NA	NA
Brachycephaly	−	+	+	+	NA
Face shape	bitemporal narrowing	round, bitemporal narrowing	angular lower facies, round face	round face	NA
Ear abnormalities	+	+	+	+	NA
Upslanted palpebral fissures	+	+	+	−	NA
Telecanthus	−	+	−	−	NA
Short palpebral fissures	+	+	−	−	NA
Abnormal nasal bridge	wide	depressed	narrow	−	NA

(*Continued on next page*)

Table 1. ***Continued***

	This Study				Previously Published Case[23]
Low columella	–	–	+	+	NA
Tooth abnormalities	+	+	+	–	NA
Micrognathia and/or retrognathia	–	+	+	+	NA
Short and/or broad neck	–	+	+	–	NA
Thorax and/or Spine					
Rib abnormalities	–	11 rib pairs	Hypoplastic 12th rib pair; abnormal articulation of the costovertebral junction and ribs	–	NA
Abnormalities in vertebra and/or pelvis	coxa valga	hip dysplasia	Subtle narrowing interpediculate distance lumbosacrally, increased cartilage pelvis, widening symphysis pubis, scoliosis	–	NA
Extremities					
Short hands	–	+	+	+	NA
Short feet	+	+	+	–	NA
Pes valgus	+	–	–	+	NA
Brachydactyly	+	+	+	+	NA
Broad thumb and/or big toe	+	±	±	–	NA
Clinodactyly in digit 2 of the hands and feet	+	+	–	–	NA
Hypoplastic end phalanges of the fifth digit and/or toe and clinodactyly	+	+	+	+	NA
Eye Abnormalities					
Abnormal retina	–	Pale retina, pale papillae	–	Pale retina	NA
Oculomotor defects	–	–	+	–	NA
Nystagmus	–	+	–	–	NA
Hypermetropia	+	–	+	–	NA
Strabismus	+	+	+	–	NA

(*Continued on next page*)

Table 1. *Continued*

	This Study				Previously Published Case[23]
Others					
Hearing loss	−	−	−		NA
Laryngomalacia	−	+	±	−	NA
Feeding difficulty	+	++	++	+	NA
Precocious puberty	+	suspected	−	−	NA
Hirsutism	+	+	+	−	NA

A list of the molecular variants in *HNRNPR* and the clinical features described in four unrelated individuals (P1–P4) in this study, plus one previously reported individual (P5).[23] NA = not available and VSD = ventricular septal defect.

variant (chromosome 1, position 23637095, CADD GRCH37-v1.4) is 33, suggesting that it is highly pathogenic.[29] The degree of developmental delay in the four individuals included in this study is moderate to severe, and they all have severe feeding difficulties. They all have microcephaly and facial dysmorphism that is partly identical, e.g., upslanted palpebral fissures (Figures 1, S1, and S2). However, some facial characteristics are different, and some features are only present in two or three of the four individuals (see also Table 1 and Figure 1). Brachydactyly is present consistently and most outspoken in the first and fifth ray of the hand (Figures 2 and S3). Abnormalities of the toes are also observed in P1–P3 (Figure S4). Only 11 rib pairs are present in P2 and P3 (Figure S5). Also present are other differences, such as the presence of cardiac abnormalities (P2), structural brain abnormalities, apart from microcephaly, such as reduced corpus callosum and cerebellar vermis hypoplasia (P2, P4, and P5) (Figures 1D, 1E, and S6), genital abnormalities (P2 and P3), and the differing extent of skeletal defects. Taken together, it is clear that many similarities exist between these individuals; however, it is difficult to say at this moment in time if this is a recognizable phenotype.

HNRNPR Mutations Truncate a Predicted Prion-like Domain

Fibroblasts and immortalized lymphoblast cell lines (LCLs) were generated from two individuals with *HNRNPR* variants; one individual (P1) harbored the c.1609dupG (p.Ala537Glyfs*10) variant and one (P2) harbored the c.1652dupG (p.Pro552Serfs*34) variant. Sanger sequencing verified the presence of these variants (Figure 3A). Both of these variants lie in the hnRNPR arginine/glycine-rich domain (RGG domain), a motif that mediates interactions between proteins and nucleic acids (frequently via methylation of the arginine residues).[30] As such, these truncating variants in P1 and P2 are predicted to abolish the protein's glutamine/asparagine-rich domain (QN domain), a domain that is commonly found in prion proteins and that drives self-propagating amyloid fibrils in neurodegenerative diseases such as Huntington (Figure 3B).[31] Moreover, the p.Ala537Glyfs*10 variant is predicted to truncate two of the four arginine residues that are known to be methylated by PRMT1 in the RGG domain,[32] although the p.Pro552Serfs*34 and p.Gln555Ter variants should truncate only the last of the four sites (Figure 3B). The missense variant p.Arg588His occurs in the QN domain (Figure 3B). As mentioned above, none of these variants nor the p.Gln555Ter variant[23] have been reported in variant population databases. In fact, the gnomAD database reports far fewer observed numbers of variants of *HNRNPR* compared to expected numbers [expected/observed synonymous variants = 125.2/112 (Z = 0.93, o/e = 0.89); expected/observed missense variants = 364/174 (Z = 3.54, o/e = 0.48); expected/observed loss-of-function (LoF) = 30/2 (pLI = 1, o/e = 0.07)]. These expected/observed ratios and the pLI

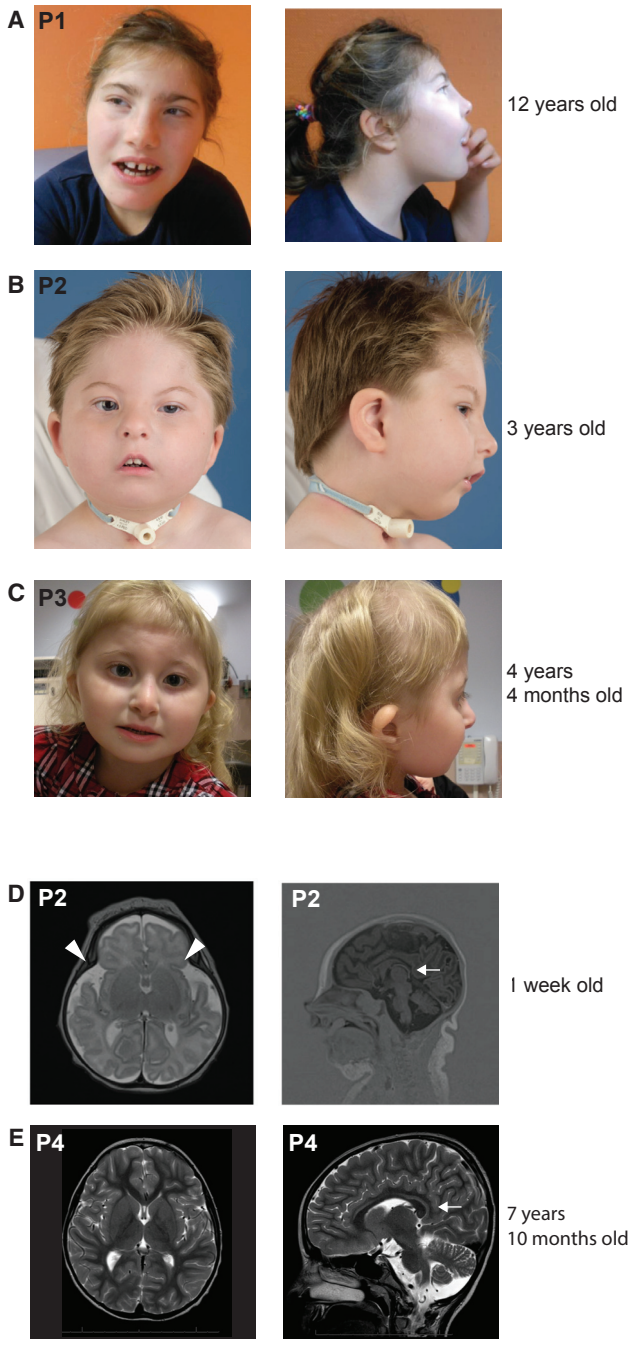


Figure 1. Facial Features and MRIs of Individuals Carrying Variants in *HNRNPR*

(A) Individual 1 (P1), who carries the heterozygous variant c.1609dupG (p.Ala537Glyfs*10). The photographs were taken when the subject was 12 years of age.

(B) Individual 2 (P2), who carries the heterozygous variant c.1652dupG (p.Pro552Serfs*34). The photographs were taken when the subject was 3 years of age.

(C) Individual 3 (P3), who carries the heterozygous variant c.1652dupG (p.Pro552Serfs*34). The photographs were taken when the subject was 4 $^{1}/_{4}$ years of age.

(D) Magnetic resonance images (MRIs) of P2 taken when the subject was 1 week of age. A small cerebellum can be observed. Arrowheads indicate the frontal atrophy, and an arrow indicates the reduced corpus callosum.

(E) MRIs of P4 taken when the subject was 7 years and 10 months of age. An arrow indicates the reduced corpus callosum.

score suggest that LoF variants are especially not tolerated well in *HNRNPR*.

Immunoblot analysis of LCLs done with antibodies against hnRNPR reveals the presence of the truncated proteins in addition to the full-length and short isoforms of hnRNPR (Figure 3C). It appears that the p.Ala537Glyfs*10 variant is co-migrating with the short isoform of hnRNPR because this band is more intense than that in the healthy control cells and because the long isoform appears diminished in both variant cell lines. This migration pattern is replicated by transfecting HeLa cells with hemagglutinin (HA)-tagged plasmids encoding either the full-length wild-type hnRNPR protein, p.Ala537Glyfs*10, or p.Pro552Serfs*34 and blotting the cells with antibodies against the HA tag (Figure 3D). Algorithms for calculating putative PrLDs on the basis of amino acid sequences[33] reveal a predicted PrLD in the last ~120 amino acids of the C-terminal region of hnRNPR (Figure 3E). This predicted PrLD in hnRNPR would be disrupted by either variant in *HNRNPR*.

The spliceosome C complex is localized in the nucleus of cells, and as such, most of the integral hnRNPs (including hnRNPR) contain a nuclear localization signal (NLS) (Figure 3B). Some hnRNPs (such as hnRNPA1) are also reported to shuttle in and out of the cytoplasm.[5] Previous proteomic analysis identified hnRNPR as a component present in cytoplasmic stress granules induced by sodium arsenite ($NaAsO_2$) treatment.[8] Because disease-linked variants in the predicted PrLD of hnRNPA1 result in an excess of stress-granule formation at basal levels,[9] we stained fibroblasts derived from affected individuals and healthy controls with antibodies against hnRNPR and the stress-granule marker Ras GTPase-activating protein-binding protein1 (G3BP1). hnRNPR localization in untreated cells was predominantly nuclear in all cell lines (Figure S7). Oxidative stress induction with sodium arsenite resulted in the formation of stress granules that co-localized with hnRNPR in healthy control cells, as well as cells derived from both affected individuals, confirming previous proteomics results suggesting hnRNPR is a component of stress granules (Figure S7).[8] However, no stress granules were observed in any cells in the absence of oxidative stress, suggesting that unlike the ALS-driving mutations in hnRNPA1 and A2/B1, the variant hnRNPR proteins reported here do not possess an intrinsic tendency to assemble into self-seeding fibrils (Figure S7).[9]

RNA Sequencing Reveals Differential Expression of Genes Important for Embryonic Development

To determine whether variants in *HNRNPR* affect RNA expression, unbiased next-generation RNA sequencing (RNA-seq) was performed on duplicate fibroblast samples from P1 (p.Ala537Glyfs*10), P2 (p.Pro552Serfs*34), and two unrelated, healthy controls (one male and one female). RNA was isolated and sequenced on the Illumina platform at a read depth of 50 million. The multidimensional scaling (MDS) plot shows the expected clustering

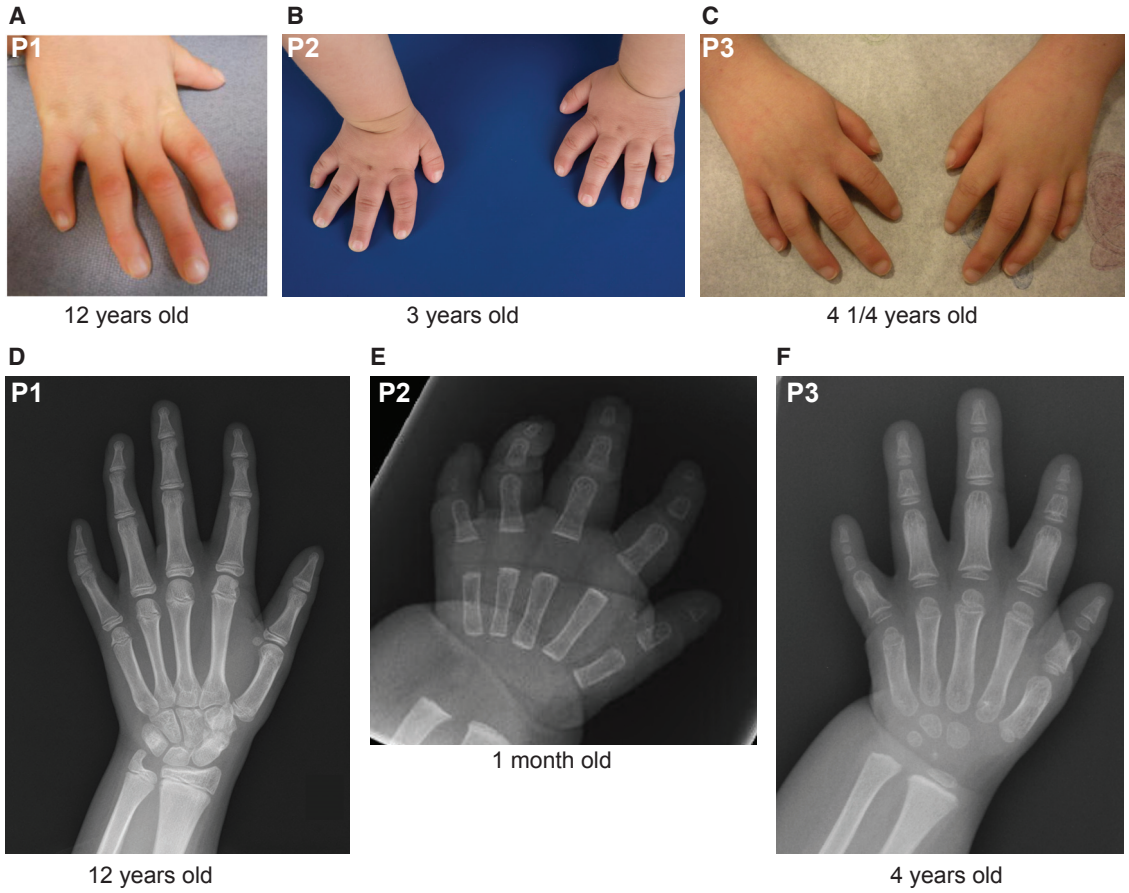


Figure 2. Hand and Foot Features of Individuals Carrying Variants in *HNRNPR*
(A–C) Photographs of the hands of P1 (A), P2 (B), and P3 (C), taken when the subjects were the ages of 12 years, 3 years, and 10 years, respectively, showing short first and fifth digits with clinodactyly.
(D) An X-ray of the hand of P1, taken when the subject was 12 years of age, showing brachymesophalangy of the fifth finger, extra sesame bones, advanced skeletal age, and a special scaphotrapezoid coalition.
(E) An X-ray of the hand of P2, taken when the subject was one month of age, showing mild shortening of the basal phalanx of digit 1, brachyacrophalangy of digits 1 and 2, brachyacro- and mesophalangy of digit 5, and clinodactyly of the distal phalangeal joint of digit 2. No ossification of the middle phalanx of the fourth digit and both phalanges of the fifth digit was yet noted.
(F) An X-ray of the hand of P3, taken when the subject was 4 years of age, showing abnormal epiphyses, a short first and second digit, brachymesophalangy of the second and fourth digit, and brachyacro- and mesophalangy of the fifth digit. Incomplete ossification of the thumb was noted.

of the duplicate cell samples from the same individuals (Figure 4A). In untreated cells, 40 individual RNAs were revealed to be significantly upregulated in the cells carrying *HNRNPR* variants compared to healthy control cells (Table 2), and 32 RNAs were significantly downregulated (Table 3).

Analysis of the gene ontology (GO) terms linked to the significantly up- and downregulated genes revealed that many of the differentially expressed RNAs play established roles in embryonic development (Figures 4B and 4C). The most significantly enriched GO terms associated with the upregulated RNAs in the cells carrying *HNRNPR* variants include functions such as embryonic skeletal system morphogenesis and development, anatomical structure morphogenesis, and embryonic organ morphogenesis (Figures 4B and S8). The most significantly enriched GO terms associated with the downregulated RNAs include functions such as female and male gonad development, development of primary female and male sexual characteristics, and sex differentiation (Figures 4C and S9). Further GO analysis of Interpro domains revealed that the most significant enrichment of differentially expressed RNAs (both up- and downregulated) are in those coding for homeobox proteins (Figure 4D).

Several Homeobox Genes and Developmental Regulators Are Differentially Expressed in Cells Carrying *HNRNPR* Variants

Figure 5A shows a volcano plot of the RNA-seq results, and the most significantly deregulated RNAs are labeled. A surprising number of these RNAs (boxed in Figure 5A) have been previously reported to have important roles in development. Indeed, the expression of each of these individual transcripts is markedly changed in cells with variants in

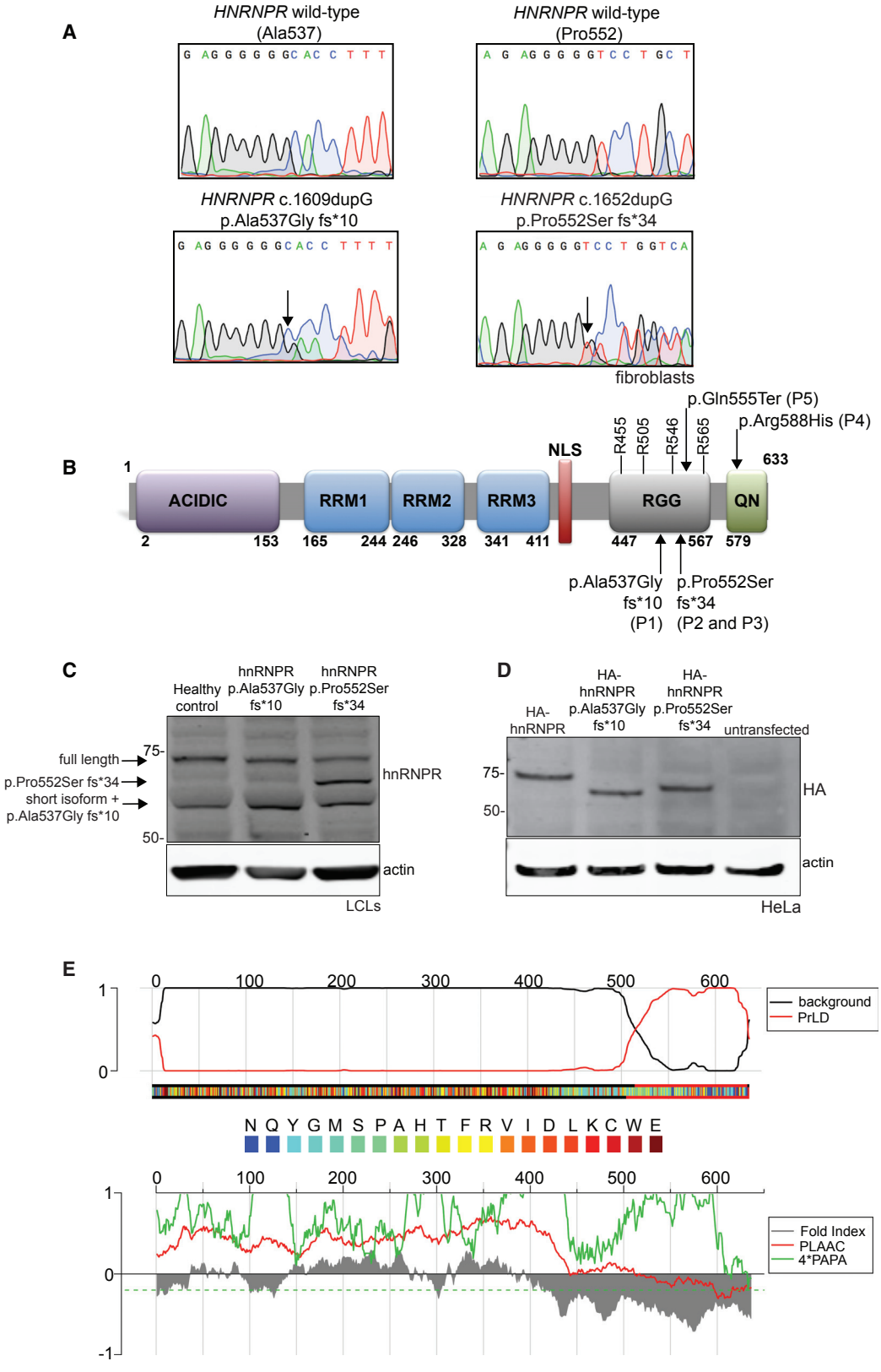


Figure 3. Frameshifting Variants in *HNRNPR* Truncate a Predicted PrLD in hnRNPR Protein

(A) Chromatograms of DNA sequencing of fibroblasts derived from healthy controls (top panels) and from P1 (p.Ala537Gly fs*10) and P2 (p.Pro552Ser fs*34) (bottom panels). Arrows in the bottom panels indicate the frameshift.

(B) Protein domains of hnRNPR. The acidic domain (in purple), RNA-binding domains (RRM1, RRM2, and RRM3; in blue), the nuclear localization signal (NLS; in red), the arginine/glycine domain (RRG; in gray), and the glutamine/asparagine domain (QN; in green) are

(legend continued on next page)

HNRNPR compared to controls (Figure 5B). *TBX1*, upregulated 15-fold in the cells carrying *HNRNPR* variants (Figures 5A 5and B), is an RNA coding for a T-box transcription factor that has a critical and well-established role in the development of all animal species.[34] Specifically, the overexpression of *Tbx1* in mice is linked to microcephaly, heart defects (ventricular septal defect [VSD]), and the impairment of trachea development.[35–37] In humans, *TBX1* variants can cause heart defects, and it is also the culprit gene suspected to drive the heart defects in individuals with deletions of the 22q11.2 chromosomal region in DiGeorge syndrome (DGS; MIM:188400).[38,39] Another noteworthy gene found upregulated 18-fold in the cells carrying the *HNRNPR* variants is *FOXG1*, a gene that has duplications and deletions and has been implicated in the congenital variant of Rett syndrome (MIM: 613454).[40,41] Also interesting in terms of embryonic development is the upregulation of pregnancy-specific glycoprotein 4 (*PSG4*). This gene has an unspecified function but is normally expressed exclusively in the placenta and has not been reported expressed ex utero (Figure 5A).[42] The most significantly upregulated gene in the *HNRNPR* mutant cells is glutathione S-transferase μ1 (*GSTM1*), which is increased in expression more than 400-fold over the cells from healthy controls (Figure 5A). This enzyme has a well-studied role in the metabolic detoxification of electrophilic compounds and xenobiotic toxins, including the detoxification of antiseizure medications.[43–45]

Most strikingly, several clustered homeobox *HOX* genes (*A5*, *B4*, and *B5*) are found highly expressed in the cells carrying *HNRNPR* variants (Figures 5A and 5B). These genes code for highly conserved DNA-binding transcription factors that regulate some of the most essential developmental decisions such as anterior-posterior axis formation and cell fate.[46,47] Sashimi plots of *HOXA5* suggest that in healthy control cells, *HOXA5* might be kept at low levels by a splicing event (boxed in red); this event has been revealed to remove a segment of RNA, transcribed from the genomic position chr7:27,140,186–27,141,774, that would include the 3′ UTR of *HOXA5* mRNA located between chr7:27,141,051–27,141,835 (Figure 6A, shaded in blue on the genomic map). This splicing event is not present in either of the cell lines carrying *HNRNPR* variants, and as such, expression of the *HOXA5* 3′ UTR is unimpeded, and expression of the 5′ end of the RNA is significantly higher (Figure 6A). In addition to the increased expression of *HOXA5*, the RNA-seq also identified a significant increase in the expression of *HOXB4*, *HOXB5*, and a sense intronic sequence of *HOXB3* in cells carrying the *HNRNPR* variants (Figures 5A and 5B). Sashimi plots covering a portion of the *HOXB* cluster reveal the increased expression of this intronic sequence located at chr17: 48,579,630–48,582,259 (boxed in blue) in addition to a general increase in expression of non-annotated intergenic sequences between the *HOXB* genes (Figure 6B). This intronic sequence includes the microRNA mir10A, which interestingly has been identified as one of the top upregulated microRNAs in murine embryonic stem cells induced with retinoic acid to differentiate toward a neuronal lineage.[48] Given that the tightly coordinated regulation of *HOX* genes is indispensable for myriad aspects of human development, including limb and digit formation,[49] it is likely that even subtle alteration or deviation of the *HOX* code can have profound downstream effects in tissue patterning.

Several genes that are downregulated in cells carrying *HNRNPR* variants are associated with GO terms describing male and female gonad development. The most substantially downregulated RNA in cells carrying the mutation is *LHX9*, which codes for the LIM homeobox 9 protein (Figures 5A, 5B, and 6C). The counts per million (CPM) reads in Figure 5B, as well as the Sashimi plots in Figure 6C, strongly suggest that transcription of *LHX9* is completely silenced in cells carrying the *HNRNPR* variants. Importantly, this gene has been characterized as a crucial regulator of gonadogenesis and forebrain development in mice.[50,51] Also downregulated in the cells carrying the *HNRNPR* variants (and contributing to the GO terms describing gonad development) are RNAs coding for the proto-oncogene c-Kit and the FOG family member zinc-finger protein ZFPM2, which play a role in male germ cell migration, proliferation, and survival, as well as in the development of the fetal testis, respectively (Figures 5A and 5B).[52,53] Another downregulated homeobox gene with a developmental function in the *HNRNPR* mutant cells is *IRX3* (Figures 5A and 5B), which has been shown in different animal models to play a critical role in dorsoventral patterning of the neural tube and the heart.[54,55]

Although it is not possible to determine which deregulated gene is potentially driving a specific clinical feature of the individuals reported here, the results of the RNA-seq taken all together reveal that *HNRNPR* variants drive a combinatorial effect on the expression of several genes

illustrated. Arginine residues reported to be methylated by PRMT1 in the RGG domain are indicated, as are the positions of the variants in individuals P1–P5.

(C) An immunoblot of lymphoblast cell lines derived from a healthy control, P1, and P2 by using antibodies against hnRNPR. The two known isoforms (full length and short form) of hnRNPR are indicated by the arrows, as are the truncated proteins (hnRNPR p.Ala537Gly fs*10 runs at the same height as the hnRNPR short form).

(D) An immunoblot of HeLa cells transfected with plasmids that coded for full-length HA-hnRNPR, HA-hnRNPR p.Ala537Gly fs*10, and HA-hnRNPR p.Pro552Ser fs*34 done by using antibodies against the HA tag.

(E) FoldIndex predicts an intrinsically unfolded region in the C terminus of hnRNPR (red curve reaching 1 on the top panel, gray curve below zero on bottom panel). The algorithms of Alberti et al.[62] (red curve below zero on the bottom panel) and Toombs et al.[63] (green curve below zero) also predict a PrLD at the C terminus of hnRNPR.

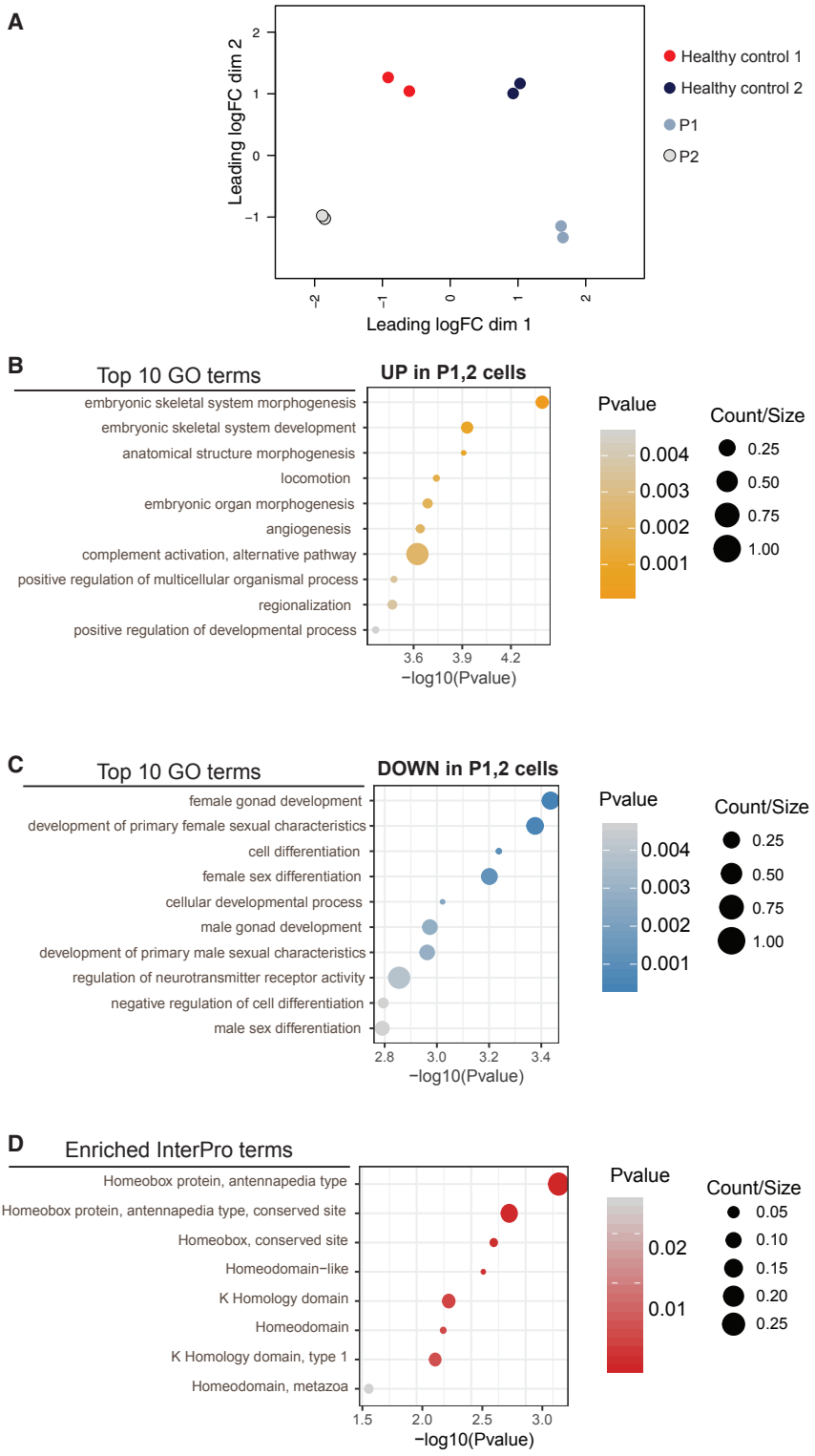


Figure 4. Gene Ontology Analysis of Deregulated RNAs in Cells Carrying *HNRNPR* Variants

(A) A multidimensional scaling (MDS) plot of RNA-seq samples showing the duplicate fibroblast samples in very close proximity.

(B) The top ten gene ontology (GO) terms enriched in RNAs that are upregulated in fibroblasts with *HNRNPR* variants compared to fibroblasts from healthy controls (using adjusted p values < 0.05).

(C) The top ten GO terms enriched in RNAs that are downregulated in fibroblasts with *HNRNPR* variants compared to fibroblasts from healthy controls (using adjusted p values < 0.05).

(D) All of the enriched InterPro terms from the analysis of both up- and downregulated RNAs in the fibroblasts carrying *HNRNPR* variants compared to fibroblasts from healthy controls.

that are well-established to have tightly regulated roles in human development.

Truncated hnRNPNR Proteins Excessively Localize to Stress Granules and Impair Their Disassembly

Given that it is known that the *HNRNPA1* and *A2/B1* variants driving ALS and FTD result in an increased tendency of the mutant hnRNPs to assemble into self-seeding fibrils that stain positively for stress-granule markers,[9] the localization of truncated hnRNPR proteins was examined. In order to determine whether the truncation of hnRNPR affects the protein's localization, HeLa cells were transfected with plasmids coding for HA-hnRNPR, HA-hnNRNPR p.Ala537Glyfs*10, or HA-hnRNPR p.Pro552Serfs*34. These cells were then fixed and stained with antibodies against the HA tag and the stress-granule marker TIA1. At basal levels, full-length and truncated hnRNPR proteins are localized predominantly in the nucleus, and there is no excessive formation of stress granules (Figure 7A), in line with what was shown in primary cells in Figure S7. However, cells exposed to sodium arsenite for 60 min revealed that despite the fact that a similar number of cells formed stress granules, a substantially larger fraction of cells expressing either of the mutant constructs displayed a strong co-localization between stress granules and the truncated HA-tagged variants compared to cells expressing the wild-type construct (Figures 7A, 7B, and S10). These results suggest that although the truncated hnRNPR proteins do not induce excessive formation of stress granules at basal levels, compared to full length hnRNPR, they do appear to have a higher affinity for stress granules once they have formed.

To determine whether this increased affinity of hnRNPR variants for stress granules affects their dynamics,

Table 2. Significantly (Adjusted P Value < 0.05) Upregulated RNAs Comparing Fibroblasts Carrying *HNRNPR* Variants (from P1 and P2) to Healthy Control Cells

Upregulated RNAs in P1 and P2's Cells Versus Healthy Control Cells					
HGNC Symbol	**Ensembl Gene ID**	**Entrez Gene**	**Description**	**Adjusted P Value**	**Log2 FC**
GSTM1	ENSG00000134184	2944	glutathione S-transferase mu 1	9.25×10^{-4}	8.84
PSG4	ENSG00000243137	5672	pregnancy specific beta-1-glycoprotein 4	1.33×10^{-2}	7.68
ADAMTS16	ENSG00000145536	170690	ADAM metallopeptidase with thrombospondin type 1 motif 16	1.17×10^{-2}	5.21
NA	ENSG00000234156	NA	Antisense of OR1B1	2.92×10^{-2}	4.72
MIR503	ENSG00000208005	574506	microRNA 503	2.24×10^{-2}	4.69
FOXG1	ENSG00000176165	2290	forkhead box G1	8.46×10^{-3}	4.26
NA	ENSG00000282556	NA	LincRNA	$8.68E \times 10^{-3}$	4.24
GMNC	ENSG00000205835	647309	geminin coiled-coil domain containing	4.09×10^{-2}	4.18
RGPD2	ENSG00000185304	729857	RANBP2-like and GRIP domain containing 2	5.25×10^{-3}	3.93
TBX1	ENSG00000184058	6899	T-box 1	3.53×10^{-2}	3.93
GNG4	ENSG00000168243	2786	G protein subunit gamma 4	3.61×10^{-2}	3.90
PLPPR5	ENSG00000117598	163404	phospholipid phosphatase related 5	3.32×10^{-2}	3.59
SNORD31B	ENSG00000201847	1E+08	small nucleolar RNA, C/D box 31B	4.90×10^{-2}	3.58
RPL34P19	ENSG00000230240	NA	ribosomal protein L34 pseudogene 19	3.96×10^{-2}	3.51
NA	ENSG00000257178	NA	Intronic sense RNA HOXB3	3.53×10^{-2}	3.50
DPYSL4	ENSG00000151640	10570	dihydropyrimidinase like 4	7.79×10^{-3}	2.68
C3	ENSG00000125730	718	complement C3	4.40×10^{-2}	2.51
TOR4A	ENSG00000198113	54863	torsin family 4 member A	4.44×10^{-2}	2.24
CFD	ENSG00000197766	1675	complement factor D	8.46×10^{-3}	2.20
HOXB5	ENSG00000120075	3215	homeobox B5	8.68×10^{-3}	2.15
SPON2	ENSG00000159674	10417	spondin 2	3.61×10^{-2}	2.10
PLAU	ENSG00000122861	5328	plasminogen activator, urokinase	1.22×10^{-3}	2.09
PI4KAP1	ENSG00000274602	728233	phosphatidylinositol 4-kinase alpha pseudogene 1	1.15×10^{-2}	2.01
HOXA5	ENSG00000106004	3202	homeobox A5	1.17×10^{-2}	1.92
SGIP1	ENSG00000118473	84251	SH3 domain GRB2 like endophilin interacting protein 1	4.54×10^{-3}	1.84
PCBP3	ENSG00000183570	54039	poly(rC) binding protein 3	2.36×10^{-2}	1.78
COL18A1	ENSG00000182871	80781	collagen type XVIII alpha 1 chain	3.96×10^{-2}	1.78
BMS1P2	ENSG00000251079	NA	BMS1, ribosome biogenesis factor pseudogene 2	4.40×10^{-2}	1.69
STEAP3	ENSG00000115107	55240	STEAP3 metalloreductase	3.18×10^{-4}	1.67
IGF2BP3	ENSG00000136231	10643	insulin like growth factor 2 mRNA binding protein 3	3.52×10^{-3}	1.60
B3GNT5	ENSG00000176597	84002	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 5	4.22×10^{-2}	1.57
MIR503HG	ENSG00000223749	84848	MIR503 host gene	1.17×10^{-2}	1.46
GFOD1	ENSG00000145990	54438	glucose-fructose oxidoreductase domain containing 1	7.79×10^{-3}	1.13
HOXB4	ENSG00000182742	3214	homeobox B4	3.48×10^{-2}	1.13
TBKBP1	ENSG00000198933	9755	TBK1 binding protein 1	3.49×10^{-2}	1.10

(*Continued on next page*)

Table 2. ***Continued***

Upregulated RNAs in P1 and P2's Cells Versus Healthy Control Cells					
AGAP9	ENSG00000204172	642517	ArfGAP with GTPase domain, ankyrin repeat and PH domain 9	2.24×10^{-2}	1.05
SLFN12	ENSG00000172123	55106	schlafen family member 12	4.90×10^{-2}	0.81
PLXNB1	ENSG00000164050	5364	plexin B1	3.96×10^{-2}	0.80
SRGAP1	ENSG00000196935	57522	SLIT-ROBO Rho GTPase activating protein 1	3.48×10^{-2}	0.58
SASH1	ENSG00000111961	23328	SAM and SH3 domain containing 1	3.61×10^{-2}	0.44

fibroblasts from two unrelated, healthy controls and individuals carrying the p.Ala537Glyfs*10 or p.Pro552Serfs*34 variant of hnRNPR were exposed to sodium arsenite for 60 min. These cells were then either immediately fixed, or the sodium arsenite was removed, and the cells were allowed to recover for 30, 45, or 60 min before fixation and staining with antibodies against hnRNPR and G3BP1 (another marker of stress granules). The fibroblasts from the healthy controls revealed that most of the stress granules induced in the 60-min exposure to sodium arsenite had dissipated within the 60-min recovery period (Figures 7C, 7D, and S11). The fibroblasts carrying the truncated hnRNPR proteins did not reveal excessive formation of stress granules at basal levels and formed stress granules in response to sodium arsenite similarly to healthy control cells. However, by the end of the 60 min recovery period, a substantial number of cells carrying the truncated hnRNPR variants had persistent, visible stress granules (Figures 7C, 7D, and S11). Taken together, these results suggest that the presence of the truncated hnRNPR proteins in stress granules hinders their disassembly.

Discussion

This study reports four unrelated children who have overlapping neurodevelopmental and dysmorphic phenotypes and who all carry *de novo* truncating or missense variants in the gene coding for the spliceosome C complex protein hnRNPR. An in-depth and unbiased examination of the transcriptomes of primary cells from affected children carrying different variants of hnRNPR (p.Ala537Glyfs*10 and p.Pro552Serfs*34) reveals that the most significant changes in RNA expression occur in transcripts coding for homeobox and T-box genes that play fundamental roles in eukaryotic development. Although many mutations in homeobox and T-box genes have been linked to impaired human development, this report connects the deregulated expression of several of these genes collectively with a *de novo* variant affecting the spliceosome.

Given what has been previously published about specific homeobox and T-box genes, it seems reasonable to speculate on how some of the developmental phenotypes of the individuals reported here may be driven by changes in the expression of these genes. For example, mice lacking the LIM homeobox gene *Lhx9* fail to form discrete gonads but have no other major developmental phenotypes.[50] The silencing of *LHX9* is readily apparent in primary cells carrying *HNRNPR* variants and may help explain the hypoplastic defects in the external genitalia (micropenis and hypoplastic labia) of P2 and P3. The fact that one male and one female were present in both the affected individuals (P1 and P2) and healthy controls (1 and 2) suggests that our findings related to gonad development are shared across both genders, although we must point out that the singular data from either gender may have influenced this result. *Lhx9* is also expressed in the developing mouse forebrain,[51] an area of the brain that is substantially compromised in P2's development (see Figure 1D). The effect of *LHX9* silencing may well be compounded by the loss of c-Kit and *ZFPM2*, which also play roles in gonad cell development and differentiation.[52,53] The upregulated expression of *FOXG1* may contribute to the neurodevelopmental phenotype. The deregulated expression of *TBX1* and *IRX3* may contribute to the heart defects, given that mice engineered to overexpress *Tbx1* reveal ventricular septal defects (as did P2 at birth), and mice with deleted *Irx3* reveal heart abnormalities that are associated with arrhythmias (as observed in P3).[35,36,38,55,56] *IRX3* loss may also contribute to the brain phenotypes, given what is known about its fundamental role in the patterning of the neural tube.[54]

The deregulation of *HOX* gene expression in the cells from affected individuals may also contribute to the pathogenesis of these *HNRNPR* variants. Pathogenic variants in several *HOX* genes have been reported to be linked to an array of human developmental phenotypes, a number of which are reported in the individuals of this study; these phenotypes include brain and cognitive abnormalities, rib defects, ear defects, facial dysmorphism, cardiovascular defects, fifth finger clinodactyly, short hands and feet, and brachydactyly.[49] Although the abnormal expression of non-mutated *HOX* genes has not yet been reported in the context of human development, animal models have demonstrated that that overexpression of *HOX* genes has profound downstream effects on cell fate and tissue patterning.[57,58] It therefore seems likely that if the *HNRNPR* variants drive the combined deregulation of *TBX1, LHX9, FOXG1, IRX3,* and *HOX* expression, collectively this gives rise to many of the clinical phenotypes that we observe in the individuals of this study. Given

Table 3. Significantly (Adjusted P Value < 0.05) Downregulated RNAs Comparing Fibroblasts Carrying *HNRNPR* Variants (from P1 and P2) to Healthy Control Cells

Downregulated RNAs in P1 and P2's Cells Versus Healthy Control Cells					
HGNC Symbol	**Ensembl Gene ID**	**Entrez Gene**	**Description**	**Adjusted P Value**	**Log2 FC**
LHX9	ENSG00000143355	56956	LIM homeobox 9	1.00×10^{-2}	−7.51
RPSAP47	ENSG00000188856	NA	ribosomal protein SA pseudogene 47	3.52×10^{-3}	−6.88
DGKB	ENSG00000136267	1607	diacylglycerol kinase beta	8.68×10^{-3}	−5.78
LYPD6B	ENSG00000150556	130576	LY6/PLAUR domain containing 6B	1.02×10^{-2}	−5.45
NA	ENSG00000261573	NA	Antisense of PTPRC	2.49×10^{-2}	−5.32
CLEC2A	ENSG00000188393	387836	C-type lectin domain family 2 member A	2.24×10^{-2}	−5.24
ANKRD30B	ENSG00000180777	374860	ankyrin repeat domain 30B	3.85×10^{-2}	−4.37
KRTAP4-9	ENSG00000212722	100132386	keratin associated protein 4-9	3.85×10^{-2}	−3.99
PAX9	ENSG00000198807	5083	paired box 9	3.96×10^{-2}	−3.93
NA	ENSG00000268747	NA	Processed pseudogene	8.46×10^{-3}	−3.77
RPSAP15	ENSG00000237506	NA	ribosomal protein SA pseudogene 15	4.54×10^{-3}	−3.35
RCN1P2	ENSG00000214455	NA	reticulocalbin 1 pseudogene 2	8.63×10^{-3}	−3.24
PREX2	ENSG00000046889	80243	phosphatidylinositol-3,4,5-trisphosphate dependent Rac exchange factor 2	2.42×10^{-3}	−3.01
LYPD6	ENSG00000187123	130574	LY6/PLAUR domain containing 6	1.40×10^{-2}	−2.98
ZIC1	ENSG00000152977	7545	Zic family member 1	4.64×10^{-3}	−2.90
KIT	ENSG00000157404	3815	KIT proto-oncogene receptor tyrosine kinase	3.96×10^{-2}	−2.71
ZIC4	ENSG00000174963	84107	Zic family member 4	3.88×10^{-2}	−2.68
MEX3B	ENSG00000183496	84206	mex-3 RNA binding family member B	3.76×10^{-2}	−2.22
ZFPM2	ENSG00000169946	23414	zinc finger protein, FOG family member 2	5.75×10^{-3}	−2.07
CPS1	ENSG00000021826	1373	carbamoyl-phosphate synthase 1	1.17×10^{-2}	−1.70
MRPL23	ENSG00000214026	6150	mitochondrial ribosomal protein L23	3.96×10^{-2}	−1.61
PLK2	ENSG00000145632	10769	polo like kinase 2	3.85×10^{-2}	−1.55
SGCD	ENSG00000170624	6444	sarcoglycan delta	4.73×10^{-4}	−1.53
C5orf30	ENSG00000181751	90355	chromosome 5 open reading frame 30	1.33×10^{-2}	−1.38
SORT1	ENSG00000134243	6272	sortilin 1	2.92×10^{-2}	−1.19
ITGA4	ENSG00000115232	3676	integrin subunit alpha 4	3.85×10^{-2}	−1.12
RPSAP58	ENSG00000225178	NA	ribosomal protein SA pseudogene 58	3.96×10^{-2}	−1.06
IRX3	ENSG00000177508	79191	iroquois homeobox 3	3.85×10^{-2}	−1.05
CTSC	ENSG00000109861	1075	cathepsin C	3.53×10^{-2}	−1.01
FNDC10	ENSG00000228594	643988	fibronectin type III domain containing 10	3.96×10^{-2}	−1.00
FBLN7	ENSG00000144152	129804	fibulin 7	3.95×10^{-2}	−0.95
DIXDC1	ENSG00000150764	85458	DIX domain containing 1	2.24×10^{-2}	−0.65

that truncated hnRNPR proteins are detectable in two distinct cell types, i.e., lymphocytes and fibroblasts, derived from individuals carrying the variants, and given that *HNRNPR* is ubiquitously expressed in human tissue,[59] it is likely that the truncated hnRNPR proteins are also expressed during embryogenesis, when the levels of homeobox and T-box genes are under strict control to ensure proper development and tissue patterning.

Several other genes, including *HNRNPH2*, *HNRNPK*, and *HNRNPU*, that code for components of the spliceosome C complex have been described in individuals with developmental disorders.[12–17,60] Although the reports to date are limited to clinical data, and no pathophysiological data are available yet, there are many overlapping phenotypes between the individuals reported in these studies compared with those reported here. The neurodevelopmental

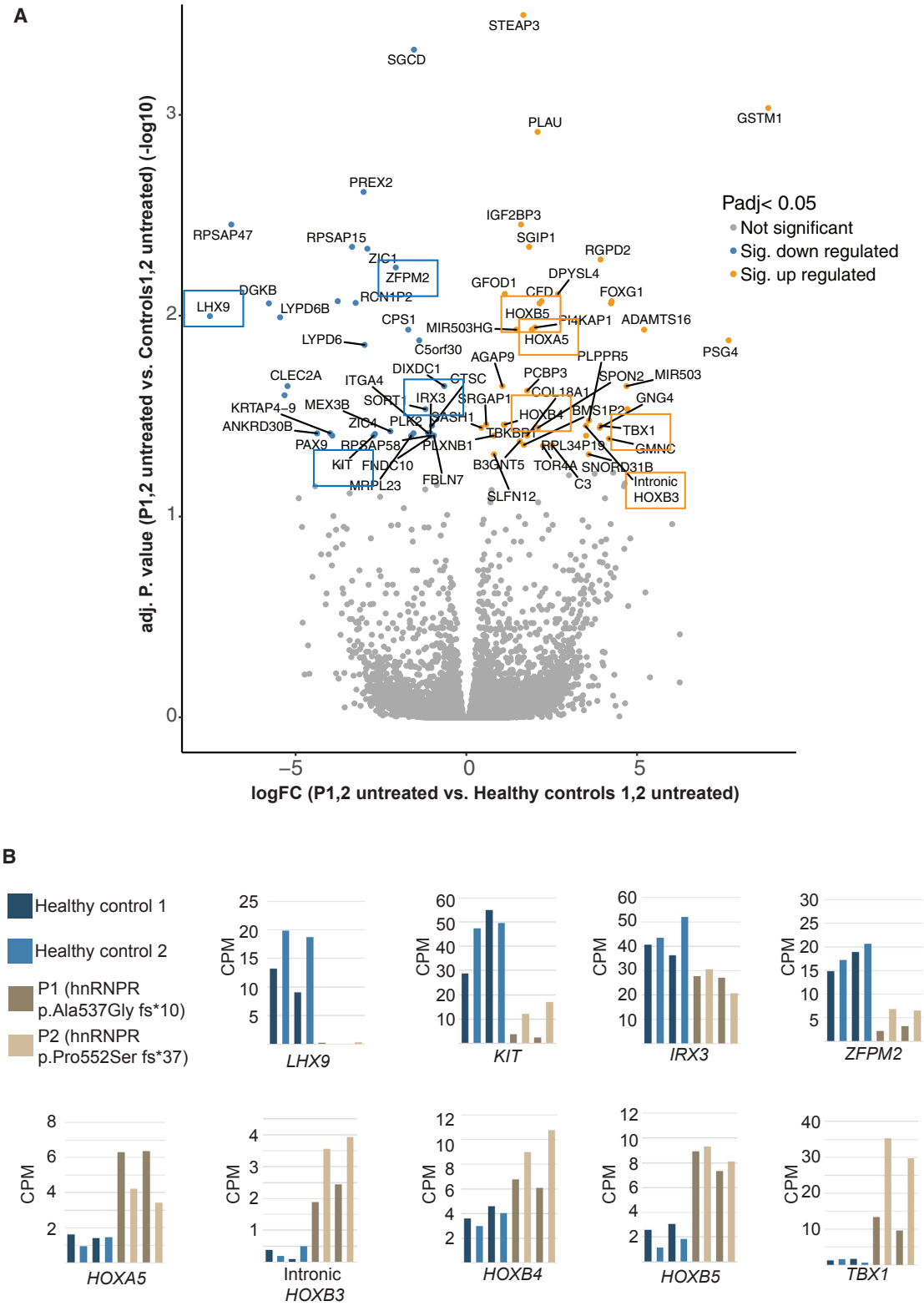


Figure 5. RNA-Seq Results Reveal the Deregulation of Many RNAs Coding for Homeobox Proteins in Cells Carrying *HNRNPR* Variants
(A) A volcano plot of RNA-seq results showing deregulated RNAs in fibroblasts carrying *HNRNPR* variants compared to fibroblasts from healthy controls, using a cut-off of an adjusted p value < 0.05. Downregulated RNAs are plotted in blue, and upregulated RNAs are plotted in yellow. Downregulated RNAs coding for homeobox proteins and those found in the GO enrichment term "gonad development" are boxed in blue, and upregulated RNAs coding for homeobox proteins and those found in the GO enrichment term "embryonic skeletal system morphogenesis" are boxed in yellow.
(B) Graphs depicting the RNA-seq counts per million (CPM) of each individual RNA boxed in (A). Blue bars represent the data from the healthy controls, and beige bars represent the data from the individuals carrying variants in *HNRNPR*.

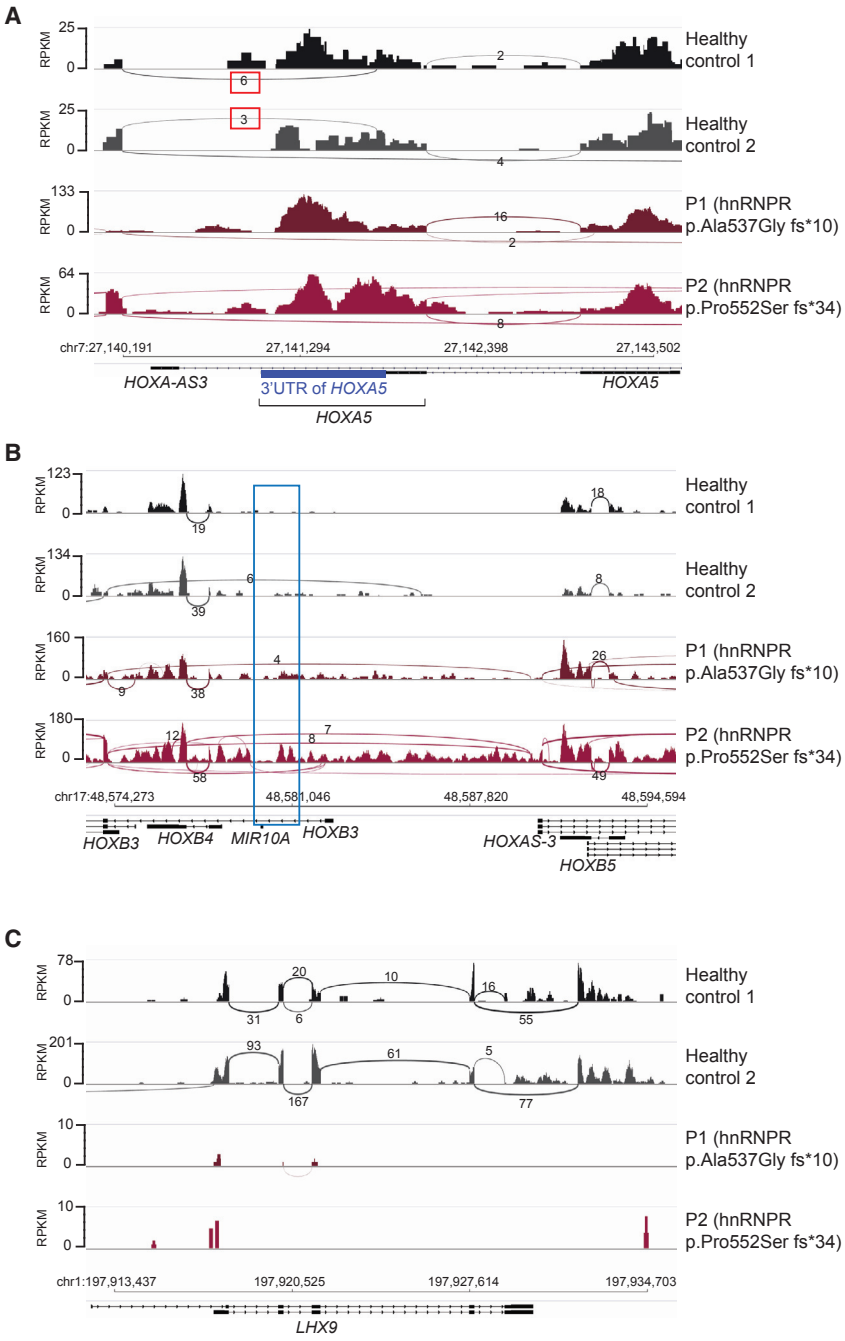


Figure 6. Sashimi Plots of RNA-Seq Data Reveal Deregulation of *HOX* Genes and *LHX9*
(A) A sashimi plot focused on *HOXA5* (3′ to 5′) and measuring RNA-seq reads by reads per kilobase million (RPKM) (note: not to scale). Red boxes indicate splice junctions found in both healthy control cell lines (top two panels, in black and gray) that are absent in the cells carrying *HNRNPR* variants (bottom panels, very dark red and fuchsia). The 3′ UTR of *HOXA5* is illustrated in blue.
(B) Sashimi plots focused on a portion of the *HOXB* cluster. The annotated intronic *HOXB3* sequence is boxed in blue.
(C) A sashimi plot focused on *LHX9* (5′ to 3′).

abnormalities, such as ID, developmental delay, and seizures, together with the facial dysmorphism and skeletal abnormalities, seem to be unifying phenotypes. Although it is currently beyond the scope of this study, it will be interesting to test whether primary cells from individuals affected by variants in other *HNRNP* genes also reveal deregulated expression levels of RNAs, such as homeobox and T-box RNAs, that are crucial for development when the RNAs are sequenced at a high read depth.

It is intriguing that all the pathogenic variants now reported for *HNRNPR* occur in the last exon of the gene and cluster in the same C-terminal region (amino acids 537–588), and that far fewer missense and LoF variants are observed in the gnomAD variant population database than are expected. The few *HNRNPR* LoF variants that have been reported in gnomAD all affect residues that are at least 75 amino acids further upstream of Ala537. The fact that truncated hnRNPR proteins are observed in primary cells carrying the variants (Figure 3C) supports an argument that mRNAs carrying pathogenic truncating variants resulting in premature stop codons escape nonsense-mediated decay and are expressed, whereas this may not be the case for the variants that introduce more N-terminal premature stop codons.[61] Taken together, these observations raise the possibility that the pathogenic *HNRNPR* variants are more than LoF and might likely be acting as dominant-negative or gain-of-function variants that require partial expression of the RGG domain (which is predicted to be abolished in the LoF variants reported in gnomAD). This will be an interesting area of future work when a clear readout of hnRNPR protein function is available.

In contrast to the self-seeding stress-granule assembly that is driven by the missense variants that are driving ALS and FTD and are reported in *HNRNPA1* and *HNRNPA2/B1*,[9] the truncated hnRNPR proteins remain localized in the nucleus and do not drive the formation of cytoplasmic stress granules at basal levels. However, cells that carry truncated hnRNPR and that are exposed to oxidative stress in the form of sodium arsenite do reveal impairments of stress-granule dynamics. A far greater proportion of truncated hnRNPR proteins was revealed to co-localize with stress-granule markers, suggesting that these proteins have a higher affinity for binding stress-granule components, and this finding would be in line with a gain-of-function variant. This is in line with the finding that stress granules in the fibroblasts of affected individuals do not

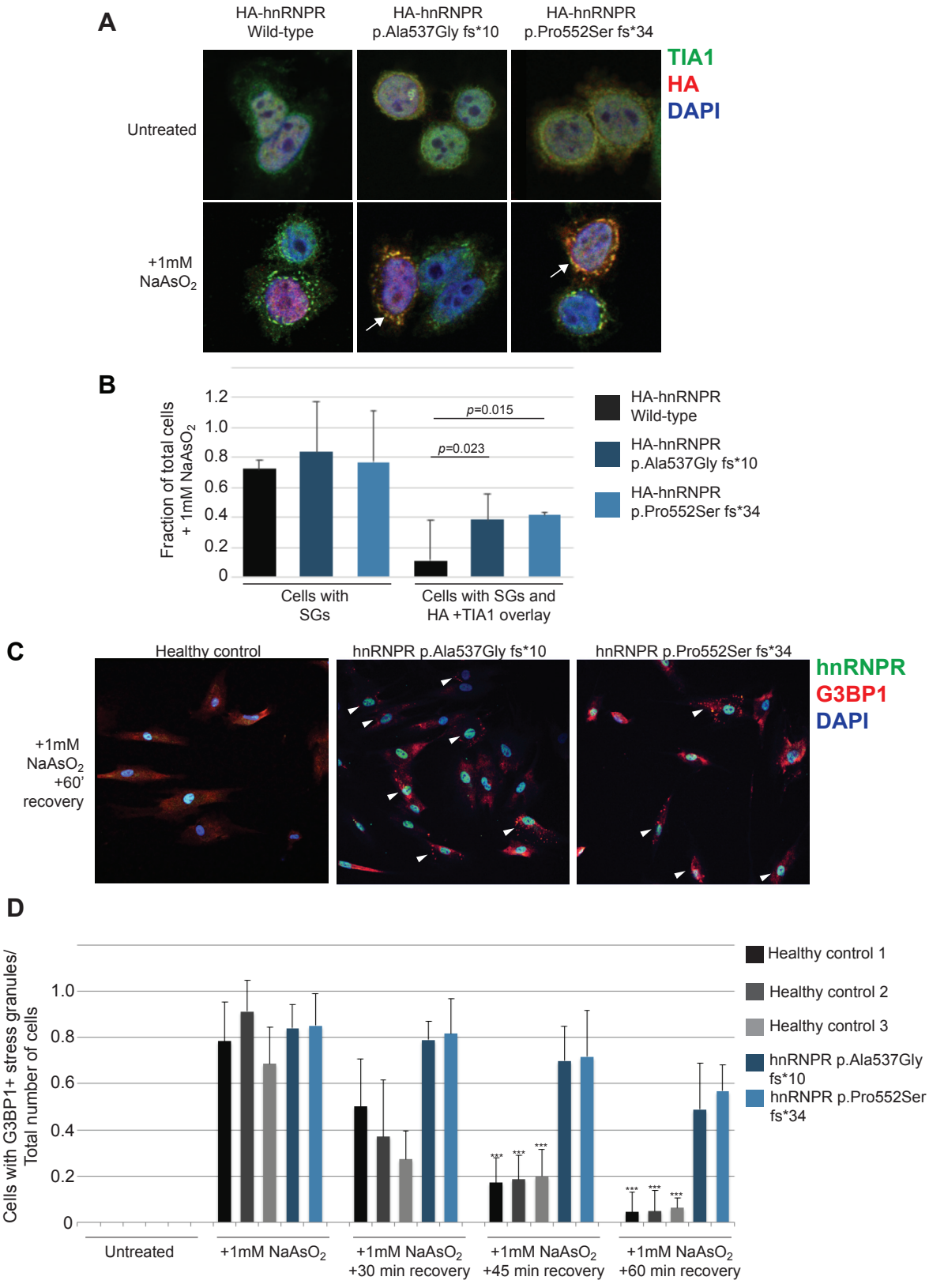


Figure 7. Truncated hnRNPR Proteins Affect Stress-Granule Dynamics

(A) A confocal microscopy analysis of HeLa cells that were transfected with plasmids coding for HA-tagged hnNRPR full-length and truncated proteins. The antibodies that were used for staining are against the stress-granule marker TIA1 (green) and the HA tag (red). DAPI (blue) is staining the nucleus. Stress granules form upon exposure to 1 mM $NaAsO_2$, and more cells are revealed with an overlay between TIA1 and HA-tagged truncated proteins (in yellow, indicated by arrows).

(legend continued on next page)

dissipate at the same rate as stress granules in healthy control cells once the sodium arsenite is removed and the cells are allowed to recover. Thus, it appears that although the truncated proteins do not drive the formation of stress granules per se, once they are associated with them, they associate very tightly. Although this cellular phenotype is readily detectable, it is far from clear how this might contribute to the developmental phenotypes of the affected individuals. At present, there is no evidence of neurodegeneration in any of the individuals in this study or those who have other hnRNP-related developmental syndromes. However, it cannot be excluded that these individuals may be at an increased risk of ALS, FTD, or other degenerative diseases in the future, given all the recent evidence linking RNA-binding-protein variants, stress-granule formation, and degeneration. Although the missense variant p.Arg588His was not tested directly, it is interesting to note that this variant occurs in the predicted PrLD of hnRNPR, similarly to the missense variants reported in hnRNPA1 and hnRNPA2/B1.[9]

This study identifies *de novo* loss-of-function variants in *HNRNPR*, which has not been previously linked to human disease, and these variants drive a developmental disorder in several individuals with overlapping clinical phenotypes. Importantly, the RNA-seq results suggest that different variants in the same gene drive similar changes in the expression of a collection of genes that play fundamental roles in the extraordinarily complex process of human embryonic development. Given the number of other individuals who have similar developmental phenotypes and who carry variants in the family of *HNRNP* genes, this report establishes the groundwork for future studies that may reveal that these disorders are driven by a similar deregulation of other developmentally important genes.

Supplemental Data

Supplemental Data can be found online at https://doi.org/10.1016/j.ajhg.2019.03.024.

Acknowledgments

Our most special thanks go to the individuals and their families who consented to be a part of this study. We thank Daisy Picavet in the Amsterdam Universitair Medische Centra (UMC) Cell Imaging Facility for assistance with the microscopy, Dirk Lebrecht at the University of Freiburg for the generation of the LCLs, and Aurore Garde at the University of Burgundy for assistance with collecting clinical data and writing the clinical history of P1. Also, thanks to Fred Vaz in the Genetische Metabole Ziekten (GMZ) Department of the Amsterdam UMC for initially sharing the variant information. Moreover, we thank Rick van Rijn (radiologist, Amsterdam UMC) for evaluating the X-rays. We would also like to thank GeneMatcher.org for assisting with the contact made between the clinics. A.W.M., M.L., A.M., and M.W.W. are supported under the frame of E-Rare-2, ERA-Net for Research on Rare Diseases (ZonMw #40-44000-98-1008 in the Netherlands, #BMBF 01GM1609 in Germany). G.E.J. is supported by a 2018 Federation of European Biochemical Society (FEBS) long-term fellowship. R.H.H. is supported by a European Research Council (ERC) Starting Grant (no. 638290) and a Vidi grant from ZonMw (no. 91715305).

Declaration of Interests

The authors declare no competing interests.

Received: October 17, 2018
Accepted: March 25, 2019
Published: May 9, 2019

Web Resources

CADD: Combined Annotation Dependent Depletion, https://cadd.gs.washington.edu
dbSNP137, https://www.ncbi.nlm.nih.gov/projects/SNP
Exome Aggregation Database, http://exac.broadinstitute.org
FastQ, http://www.bioinformatics.babraham.ac.uk/projects/fastqc
GeneMatcher, https://www.genematcher.org
Genome Aggregation Database, https://gnomad.broadinstitute.org
Mutation Taster, http://www.mutationtaster.org
MutPred2, http://mutpred.mutdb.org
National Heart, Lung, and Blood Institute (NHLBI) GO Exome Sequencing Project (ESP5400 release), http://evs.gs.washington.edu/EVS/
The Online Mendelian Inheritance in Man, http://www.omim.org
PLAAC: Prion-Like Amino Acid Composition, http://plaac.wi.mit.edu
PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/
SIFT: Sorting Intolerant from Tolerant, https://sift.bii.a-star.edu.sg

References

1. Choi, Y.D., Grabowski, P.J., Sharp, P.A., and Dreyfuss, G. (1986). Heterogeneous nuclear ribonucleoproteins: Role in RNA splicing. Science *231*, 1534–1539.

(B) A quantification of the cells in (A) (N = 3 independent transfections) that formed stress granules, as well as the cells that formed an overlay between TIA1 and HA-tagged hnRNPR proteins.
(C) A confocal microscopy analysis of fibroblasts derived from a healthy control or individuals carrying variants in *HNRNPR*; the fibroblasts were exposed to 1 mM $NaAsO_2$ for 60 min and then allowed to recover for 60 min. The antibodies used for staining are against hnRNPR (green) and the stress-granule marker G3BP1 (red). DAPI (blue) is staining the nucleus. Arrowheads indicate cells revealing stress granules.
(D) A quantification of fibroblasts that formed stress granules in untreated cells, cells exposed to 1 mM $NaAsO_2$ for 60 min and immediately fixed, and a recovery time course (N = between 3–5 independent experiments). Three unrelated cell lines from healthy controls were used (black and gray bars) for a comparison to cells carrying *HNRNPR* variants (blue bars). A Student's t test was used for statistical analysis, ***$p < 0.001$.
All error bars were derived using the standard deviation.

2. Beyer, A.L., Christensen, M.E., Walker, B.W., and LeStourgeon, W.M. (1977). Identification and characterization of the packaging proteins of core 40S hnRNP particles. Cell *11*, 127–138.
3. Chaudhury, A., Chander, P., and Howe, P.H. (2010). Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. RNA *16*, 1449–1462.
4. Dreyfuss, G., Matunis, M.J., Piñol-Roma, S., and Burd, C.G. (1993). hnRNP proteins and the biogenesis of mRNA. Annu. Rev. Biochem. *62*, 289–321.
5. Shen, E.C., Henry, M.F., Weiss, V.H., Valentini, S.R., Silver, P.A., and Lee, M.S. (1998). Arginine methylation facilitates the nuclear export of hnRNP proteins. Genes Dev. *12*, 679–691.
6. Piñol-Roma, S., and Dreyfuss, G. (1993). hnRNP proteins: Localization and transport between the nucleus and the cytoplasm. Trends Cell Biol. *3*, 151–155.
7. Protter, D.S.W., and Parker, R. (2016). Principles and properties of stress granules. Trends Cell Biol. *26*, 668–679.
8. Jain, S., Wheeler, J.R., Walters, R.W., Agrawal, A., Barsic, A., and Parker, R. (2016). ATPase-modulated stress granules contain a diverse proteome and substructure. Cell *164*, 487–498.
9. Kim, H.J., Kim, N.C., Wang, Y.D., Scarborough, E.A., Moore, J., Diaz, Z., MacLea, K.S., Freibaum, B., Li, S., Molliex, A., et al. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature *495*, 467–473.
10. King, O.D., Gitler, A.D., and Shorter, J. (2012). The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease. Brain Res. *1462*, 61–80.
11. Toombs, J.A., McCarty, B.R., and Ross, E.D. (2010). Compositional determinants of prion formation in yeast. Mol. Cell. Biol. *30*, 319–332.
12. Au, P.Y.B., You, J., Caluseriu, O., Schwartzentruber, J., Majewski, J., Bernier, F.P., Ferguson, M., Valle, D., Parboosingh, J.S., Sobreira, N., et al.; Care for Rare Canada Consortium (2015). GeneMatcher aids in the identification of a new malformation syndrome with intellectual disability, unique facial dysmorphisms, and skeletal and connective tissue abnormalities caused by de novo variants in HNRNPK. Hum. Mutat. *36*, 1009–1014.
13. Dentici, M.L., Barresi, S., Niceta, M., Pantaleoni, F., Pizzi, S., Dallapiccola, B., Tartaglia, M., and Digilio, M.C. (2018). Clinical spectrum of Kabuki-like syndrome caused by HNRNPK haploinsufficiency. Clin. Genet. *93*, 401–407.
14. Lange, L., Pagnamenta, A.T., Lise, S., Clasper, S., Stewart, H., Akha, E.S., Quaghebeur, G., Knight, S.J., Keays, D.A., Taylor, J.C., and Kini, U. (2016). A de novo frameshift in HNRNPK causing a Kabuki-like syndrome with nodular heterotopia. Clin. Genet. *90*, 258–262.
15. Thierry, G., Bénéteau, C., Pichon, O., Flori, E., Isidor, B., Popelard, F., Delrue, M.A., Duboscq-Bidot, L., Thuresson, A.C., van Bon, B.W., et al. (2012). Molecular characterization of 1q44 microdeletion in 11 patients reveals three candidate genes for intellectual disability and seizures. Am. J. Med. Genet. A. *158A*, 1633–1640.
16. Bramswig, N.C., Lüdecke, H.J., Hamdan, F.F., Altmüller, J., Beleggia, F., Elcioglu, N.H., Freyer, C., Gerkes, E.H., Demirkol, Y.K., Knupp, K.G., et al. (2017). Heterozygous HNRNPU variants cause early onset epilepsy and severe intellectual disability. Hum. Genet. *136*, 821–834.
17. Leduc, M.S., Chao, H.T., Qu, C., Walkiewicz, M., Xiao, R., Magoulas, P., Pan, S., Beuten, J., He, W., Bernstein, J.A., et al. (2017). Clinical and molecular characterization of de novo loss of function variants in HNRNPU. Am. J. Med. Genet. A. *173*, 2680–2689.
18. Hui-Yuen, J., McAllister, S., Koganti, S., Hill, E., and Bhaduri-McIntosh, S. (2011). Establishment of Epstein-Barr virus growth-transformed lymphoblastoid cell lines. J. Vis. Exp. *57*, 3321.
19. Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: A fast spliced aligner with low memory requirements. Nat. Methods *12*, 357–360.
20. Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.
21. Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.
22. Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.
23. Helbig, K.L., Farwell Hagman, K.D., Shinde, D.N., Mroske, C., Powis, Z., Li, S., Tang, S., and Helbig, I. (2016). Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. Genet. Med. *18*, 898–905.
24. Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: A matching tool for connecting investigators with an interest in the same gene. Hum. Mutat. *36*, 928–930.
25. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat. Methods *7*, 248–249.
26. Pejaver, V., Urresti, J., Lugo-Martinez, J., Pagel, K.A., Lin, G., Nam, H., Mort, M., Cooper, D.N., Sebat, J., Iakoucheva, L.M., et al. (2017). MutPred2: Inferring the molecular and phenotypic impact of amino acid variants. bioRxiv.
27. Vaser, R., Adusumalli, S., Leng, S.N., Sikic, M., and Ng, P.C. (2016). SIFT missense predictions for genomes. Nat. Protoc. *11*, 1–9.
28. Schwarz, J.M., Cooper, D.N., Schuelke, M., and Seelow, D. (2014). MutationTaster2: Mutation prediction for the deep-sequencing age. Nat. Methods *11*, 361–362.
29. 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.
30. Thandapani, P., O'Connor, T.R., Bailey, T.L., and Richard, S. (2013). Defining the RGG/RG motif. Mol. Cell *50*, 613–623.
31. Michelitsch, M.D., and Weissman, J.S. (2000). A census of glutamine/asparagine-rich regions: Implications for their conserved function and the prediction of novel prions. Proc. Natl. Acad. Sci. USA *97*, 11910–11915.
32. Wada, K., Inoue, K., and Hagiwara, M. (2002). Identification of methylated proteins by protein arginine N-methyltransferase 1, PRMT1, with a new expression cloning strategy. Biochim. Biophys. Acta *1591*, 1–10.
33. Lancaster, A.K., Nutter-Upham, A., Lindquist, S., and King, O.D. (2014). PLAAC: A web and command-line application to identify proteins with prion-like amino acid composition. Bioinformatics *30*, 2501–2502.

34. Papaioannou, V.E., and Silver, L.M. (1998). The T-box gene family. BioEssays *20*, 9–19.
35. Freyer, L., Nowotschin, S., Pirity, M.K., Baldini, A., and Morrow, B.E. (2013). Conditional and constitutive expression of a Tbx1-GFP fusion protein in mice. BMC Dev. Biol. *13*, 33.
36. Vitelli, F., Huynh, T., and Baldini, A. (2009). Gain of function of Tbx1 affects pharyngeal and heart development in the mouse. Genesis *47*, 188–195.
37. Hasten, E., McDonald-McGinn, D.M., Crowley, T.B., Zackai, E., Emanuel, B.S., Morrow, B.E., and Racedo, S.E. (2018). Dysregulation of TBX1 dosage in the anterior heart field results in congenital heart disease resembling the 22q11.2 duplication syndrome. Hum. Mol. Genet. *27*, 1847–1857.
38. Yagi, H., Furutani, Y., Hamada, H., Sasaki, T., Asakawa, S., Minoshima, S., Ichida, F., Joo, K., Kimura, M., Imamura, S., et al. (2003). Role of TBX1 in human del22q11.2 syndrome. Lancet *362*, 1366–1373.
39. Pan, Y., Wang, Z.G., Liu, X.Y., Zhao, H., Zhou, N., Zheng, G.F., Qiu, X.B., Li, R.G., Yuan, F., Shi, H.Y., et al. (2015). A novel TBX1 loss-of-function mutation associated with congenital heart disease. Pediatr. Cardiol. *36*, 1400–1410.
40. Ariani, F., Hayek, G., Rondinella, D., Artuso, R., Mencarelli, M.A., Spanhol-Rosseto, A., Pollazzon, M., Buoni, S., Spiga, O., Ricciardi, S., et al. (2008). FOXG1 is responsible for the congenital variant of Rett syndrome. Am. J. Hum. Genet. *83*, 89–93.
41. Brunetti-Pierri, N., Paciorkowski, A.R., Ciccone, R., Della Mina, E., Bonaglia, M.C., Borgatti, R., Schaaf, C.P., Sutton, V.R., Xia, Z., Jelluma, N., et al. (2011). Duplications of FOXG1 in 14q12 are associated with developmental epilepsy, mental retardation, and severe speech impairment. Eur. J. Hum. Genet. *19*, 102–107.
42. Teglund, S., Olsen, A., Khan, W.N., Frängsmyr, L., and Hammarström, S. (1994). The pregnancy-specific glycoprotein (PSG) gene cluster on human chromosome 19: Fine structure of the 11 PSG genes and identification of 6 new genes forming a third subgroup within the carcinoembryonic antigen (CEA) family. Genomics *23*, 669–684.
43. Tang, W., Borel, A.G., and Abbott, F.S. (1996). Conjugation of glutathione with a toxic metabolite of valproic acid, (E)-2-propyl-2,4-pentadienoic acid, catalyzed by rat hepatic glutathione-S-transferases. Drug Metab. Dispos. *24*, 436–446.
44. Bu, H.Z., Kang, P., Deese, A.J., Zhao, P., and Pool, W.F. (2005). Human in vitro glutathionyl and protein adducts of carbamazepine-10,11-epoxide, a stable and pharmacologically active metabolite of carbamazepine. Drug Metab. Dispos. *33*, 1920–1924.
45. Bu, H.Z., Zhao, P., Dalvie, D.K., and Pool, W.F. (2007). Identification of primary and sequential bioactivation pathways of carbamazepine in human liver microsomes using liquid chromatography/tandem mass spectrometry. Rapid Commun. Mass Spectrom. *21*, 3317–3322.
46. Pearson, J.C., Lemons, D., and McGinnis, W. (2005). Modulating Hox gene functions during animal body patterning. Nat. Rev. Genet. *6*, 893–904.
47. Krumlauf, R. (1994). Hox genes in vertebrate development. Cell *78*, 191–201.
48. Terranova, C., Narla, S.T., Lee, Y.W., Bard, J., Parikh, A., Stachowiak, E.K., Tzanakakis, E.S., Buck, M.J., Birkaya, B., and Stachowiak, M.K. (2015). Global developmental gene programing involves a nuclear form of fibroblast growth factor receptor-1 (FGFR1). PLoS ONE *10*, e0123380.
49. Quinonez, S.C., and Innis, J.W. (2014). Human HOX gene disorders. Mol. Genet. Metab. *111*, 4–15.
50. Birk, O.S., Casiano, D.E., Wassif, C.A., Cogliati, T., Zhao, L., Zhao, Y., Grinberg, A., Huang, S., Kreidberg, J.A., Parker, K.L., et al. (2000). The LIM homeobox gene Lhx9 is essential for mouse gonad formation. Nature *403*, 909–913.
51. Rétaux, S., Rogard, M., Bach, I., Failli, V., and Besson, M.J. (1999). Lhx9: A novel LIM-homeodomain gene expressed in the developing forebrain. J. Neurosci. *19*, 783–793.
52. Mauduit, C., Hamamah, S., and Benahmed, M. (1999). Stem cell factor/c-kit system in spermatogenesis. Hum. Reprod. Update *5*, 535–545.
53. Bouma, G.J., Washburn, L.L., Albrecht, K.H., and Eicher, E.M. (2007). Correct dosage of Fog2 and Gata4 transcription factors is critical for fetal testis development in mice. Proc. Natl. Acad. Sci. USA *104*, 14994–14999.
54. Briscoe, J., Pierani, A., Jessell, T.M., and Ericson, J. (2000). A homeodomain protein code specifies progenitor cell identity and neuronal fate in the ventral neural tube. Cell *101*, 435–445.
55. Kim, K.H., Rosen, A., Hussein, S.M., Puviindran, V., Korogyi, A.S., Chiarello, C., Nagy, A., Hui, C.C., and Backx, P.H. (2016). Irx3 is required for postnatal maturation of the mouse ventricular conduction system. Sci. Rep. *6*, 19197.
56. Zhang, S.S., Kim, K.H., Rosen, A., Smyth, J.W., Sakuma, R., Delgado-Olguín, P., Davis, M., Chi, N.C., Puviindran, V., Gaborit, N., et al. (2011). Iroquois homeobox gene 3 establishes fast conduction in the cardiac His-Purkinje network. Proc. Natl. Acad. Sci. USA *108*, 13576–13581.
57. Wolgemuth, D.J., Behringer, R.R., Mostoller, M.P., Brinster, R.L., and Palmiter, R.D. (1989). Transgenic mice overexpressing the mouse homoeobox-containing gene Hox-1.4 exhibit abnormal gut development. Nature *337*, 464–467.
58. Jegalian, B.G., and De Robertis, E.M. (1992). Homeotic transformations in the mouse induced by overexpression of a human Hox3.3 transgene. Cell *71*, 901–910.
59. Fagerberg, L., Hallström, B.M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., Habuka, M., Tahmasebpoor, S., Danielsson, A., Edlund, K., et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. Mol. Cell. Proteomics *13*, 397–406.
60. Bain, J.M., Cho, M.T., Telegrafi, A., Wilson, A., Brooks, S., Botti, C., Gowans, G., Autullo, L.A., Krishnamurthy, V., Willing, M.C., et al. (2016). Variants in HNRNPH2 on the X chromosome are associated with a neurodevelopmental disorder in females. Am. J. Hum. Genet. *99*, 728–734.
61. Hug, N., Longman, D., and Cáceres, J.F. (2016). Mechanism and regulation of the nonsense-mediated decay pathway. Nucleic Acids Res. *44*, 1483–1495.
62. Alberti, S., Halfmann, R., King, O., Kapila, A., and Lindquist, S. (2009). A systematic survey identifies prions and illuminates sequence features of prionogenic proteins. Cell *137*, 146–158.
63. Toombs, J.A., McCarty, B.R., and Ross, E.D. (2010). Compositional determinants of prion formation in yeast. Mol. Cell. Biol. *30*, 319–332.

The American Journal of Human Genetics, Volume *104*

Supplemental Data

HNRNPR Variants that Impair Homeobox Gene Expression Drive Developmental Disorders in Humans

Floor A. Duijkers, Andrew McDonald, Georges E. Janssens, Marco Lezzerini, Aldo Jongejan, Silvana van Koningsbruggen, Wendela G. Leeuwenburgh-Pronk, Marcin W. Wlodarski, Sébastien Moutton, Frédéric Tran-Mau-Them, Christel Thauvin-Robinet, Laurence Faivre, Kristin G. Monaghan, Thomas Smol, Odile Boute-Benejean, Roger L. Ladda, Susan L. Sell, Ange-Line Bruel, Riekelt H. Houtkooper, and Alyson W. MacInnes

Supplemental Note: Case reports

Individual 1 (P1) is a 12-year-old female (Figure 1A). Birth measurements at 39 weeks of gestation were normal with a weight of 3610 gram (75th percentile), a length of 50 cm (50th percentile), and a head circumference of 34 cm (50th percentile). During the neonatal period, she presented with poor weight gain and feeding difficulties, first thought to be a severe gastroesophageal reflux. At 3 months of age the individual was referred to the French clinic because of hypotonia, acquired microcephaly (below 1st percentile at 3 months while at 50th percentile at birth), and significant strabismus. At the age of 3 months rough hair was also observed and at the age of 18 months she experienced febrile seizure and esophagitis. At age 9 years and 6 months the patient developed precocious puberty and still had feeding problems, mainly bulimia, and regurgitation. She has a severe global developmental delay. She could sit unsupported at age 11 months and gait was acquired at 4 years of age. She had progressive valgus and flat feet that impaired walking. She has not developed any spoken language but can communicate with hand movements, and comprehension skills for simple tasks are correct. She is able to undress and eat alone but requires assistance for other daily tasks. Behavioral disturbances are characterized by hand stereotypies, attention disorder, and tantrums. Upon clinical examination she showed a narrow forehead, up-slanted palpebral fissures, convergent strabismus, ear lobule hypoplasia, wide nasal bridge, short and broad thumbs and halluces, and clinodactyly of the fifth fingers and toes (Figures 1A, 2A, and S4A). She had some degree of limb stiffness. Her primary teeth had to be extracted because of loss failure and dental crowding with permanent tooth growth.

Additional investigations were performed and hand X-rays show hypodense lesions on distal phalanges and the first metacarpal, as well as brachymesophalangy of the fifth finger (Figure 2D). MRI of the brain shows microcephaly and a short corpus callosum (but with conserved proportions with microcephaly). Ultrasound investigation of the kidneys, vertebral x-rays and audiometry are normal. Genetic DNA-tests such as karyotyping, array-CHG and DNA-investigations for Angelman syndrome were normal; as was the metabolic screening. Trio

whole exome sequencing (WES) identified *HNRNPR* as a candidate when a *de novo* frameshift variant chr1: g.23637248_23637249insC, NM_001102398.1: c.1609dupG, p.(Ala537Glyfs*10) was found.

Individual 2 (P2) is a 3-year old boy that presented at birth with multiple congenital abnormalities, breathing problems, and dysmorphic features (Figures 1B and S1). He was born with a severe laryngomalacia, for which wedge excision of the adenoids and a tracheostomy was performed. He also had a few small, muscular ventral septum defects (VSDs) that closed spontaneously. Furthermore, he had cryptorchidism, a micropenis, and is suspected to have a precocious puberty. He was diagnosed with nightly hypoventilation at the age of 9 months for which nocturnal respiratory support was started. He has had a few epileptic seizures, but this is not one of his main clinical problems. He was surgically operated on for severe hip dysplasia. Physical examination revealed a boy with characteristic facial features including bitemporal narrowing, short and narrow palpebral features, almond-shaped eyes, telecanthus, low set ears with thickened helix, and some degree of retrognathia (Figures 1B and S1). He also has a short and thick neck with excess skin folds. The hands appear short, with a combination of short middle phalanges and ulnar deviation of the distal phalanx of the second digit on both hands (Figures 2B, S3, and S4B). The distal bending folds of these fingers are absent. Clinodactyly of the fifth digit was also apparent and the feet show a chaotic implant of the toes (Figures 2B and S3). During his first year he had remarkable hair (Figure S1B) and he was diagnosed with hirsutism at 21 months of age. The boy was very cheerful in the first two years of life. After this his behavioral problems have become increasingly clear with attention deficit disorder, head banging, repeated removal of his cannulas and stereotypic behavior. Developmental delay was present from the beginning; however, the degree of developmental delay has also become more apparent in the last year. At the age of 3 years and 5 months his motor development is estimated at approximately 11 months but this may underestimate his developmental potential because of his severe hip dysplasia. He can now stand upright with

support but is unable to walk. He cannot speak, but this is probably due to his tracheostomy. He also does not understand sign language or pictograms yet. The only signal he uses is for his own name. This signal is named after his favorite position laying on this back with both hands behind the head and frog-like leg position.

Additional investigations that have been performed are an MRI-scan of the brain showing underdevelopment and atrophy of the frontal lobes, a hypoplastic corpus callosum, and small cerebellar hemispheres and vermis (Figure 1D). Thoracic x-rays show 11 rib pairs (Figure S5A). X-rays of the hands and feet show a remarkable pattern of brachydactyly, delayed ossification, and clinodactyly (Figure 2E. DNA diagnostics showed a small maternally inherited CNV by SNP-array (arr[hg19] 1p34.1(44,680,009-44,686,395)x1 mat), which is probably innocent. Trio-based exome sequencing revealed a heterozygous variant c.1652dupG p.(Pro552Serfs*34) in *HNRNPR*.

Individual 3 (P3) is a 10-year old female (Figures 1C and S2). At 2½ months of age she presented as a symmetrically small infant with a height and weight in the 3rd percentile and head circumference of at the 50th percentile. She has generalized hypotonia and swallowing problems for which she still requires tube feeding. At the first clinical examination she showed brachycephaly with prominent frontal bossing. Later a short, sloping forehead, with mild facial asymmetry, arched eyebrows, a narrow nasal bridge, depressions in the nasal tip, a low placed columella, short philtrum, angular lower facies, and mild micrognathia became apparent (Figure 1C). Her eyes are almond-shaped and her ears are small and cup-shaped (Figure 1C). Her left ear is smaller and shows an abnormal lobule (Figure 1C). The incisors are widely spaced, while the rest of the teeth are crowded. She has sparse, fine scalp hair and excess skin folds in the neck. More recently she still shows sparse scalp hair, but generalized hirsutism and a low posterior hairline. She has symmetrically small, puffy hands and feet (Figure 2C). There was bilateral clinodactyly of the 5th digits, proximal placed short thumbs and brachydactyly (more specific marked hypoplasia of the middle and distal phalanges;

Figure 2C). The digits are tubular in shape (Figure 2C). There are bilateral unusual palmar creases and hyperkeratosis pilaris. Her toes reveal a cutaneous syndactyly from the second to the fourth digit and short, curved 5th toes (Figures 2D and S4C). She has hypoplastic labia majora and a normal clitoris. Developmentally at 8 months, poor eye contact was noted. She was rolling, sitting, reaching, transferring objects, and babbling by this age. She crawled at 13 months and stood with support by 14 months. She walked with assistance by 20 months and walked independently by 33 months. A Denver II developmental assessment performed at 24 months revealed age appropriate personal-social skills. Fine motor-adaptive skills were delayed to about 13-17 months. Spoken language was at the approximate 16-month level but the patient knew body parts and could point to 4 pictures age-appropriately. Gross motor skills were delayed as she was not able to jump or climb stairs. A formal developmental evaluation was performed at an outside center at the age of 9 8/12 years. This revealed speech delay with the ability to use single words, use of some signs, and ability to use a tablet. Fine motor development was delayed; she was ambidextrous, could dress herself slowly, but not use zippers or buttons. She was working on utensils. She could write the first letter of her name and pick her name out from distractors. Gross motor development was delayed as she continued to need help with stairs. She was not completely toilet trained. By our last evaluation at 10 9/12 years, she carried a diagnosis of moderate developmental delay (IQ tested 52), pervasive developmental disorder and exhibited stereotypic behavior. She often rubs her knuckles into her bones, wants to touch her elbows or the elbows of others and chews on her hands or necklace. She was still not fully toilet trained.

Simple, focal seizures are present. Her visual maturation was delayed and she has myopic astigmatism, intermittent esotropia, and anisometropia. A heart exam revealed variable systolic murmur along the left sternal midline without abnormalities on ultrasound.

Other additional investigations were performed as well and brain MRI shows microcephaly, but no other abnormalities. Radiographs show 11 rib pairs (Figure S5B), abnormal articulation at the costovertebral junction of the ribs, subluxation of the radial head, subtle narrowing of the

interpediculate distance at the lumbosacral region. X-rays of the hands demonstrate hypoplastic phalanges with abnormal epiphyses and delayed ossification of the thumb (Figure 2F). The big toe of the foot shows shortening and broadening of both phalanges (Figure S4C). Trio exome sequencing revealed a de novo heterozygous variant in *HNRNPR* c.1652dupG p.(Pro552Serfs*34). This is the same variant identified in P2.

Individual 4 (P4) is an 8-year old male. Birth measurement at 39 weeks of gestation were normal with a weight of 3340g (50th percentile), a length of 50 cm (50-75th percentile) a head circumference of 34 cm (50th percentile). During the first year he presented with hypotonia, severe gastroesophageal reflux, feeding difficulties and constipation. He had Nissen fundoplication and was fed via percutaneous gastrostomy. Later a partial achalasia of the esophagus was diagnosed.

He was referred to clinical genetic at 6 years for severe global developmental delay, acquired microcephaly and growth delay. He could sit unsupported at two years and walked with ataxia at 4 years. At six years, the language was at a 9-month-old level. He could use three signs. He had no behavioral disturbances.

Clinical examination showed a high forehead, micrognathia, wide palpebral fissures, low columella, low-set and anteverted ears. He had short hands and brachydactyly, clinodactyly of fifth fingers, fetal pads and syndactyly of the fingers. He had a sacrococcygeal fistula.

Additional investigations performed included a brain MRI showing cerebellar hypoplasia, mega cisterna magna and a short corpus callosum (Figure 1E and S6). Medullary MRI showed normal medullary signs and arachnoidal cyst. Ultrasound of the kidneys and heart, vertebral X-Ray, and audiometry tests were normal. Ophthalmological examination showed normal vision and albinoid retina.

Karyotype, CGH-array, DNA investigations for Fragile X syndrome (FXS; MIM:300624) and Angelman syndrome (AS; MIM: 105830) were normal. Trio exome sequencing revealed a heterozygous missense variant in *HNRNPR* c.1763G>A p.(Arg588His).

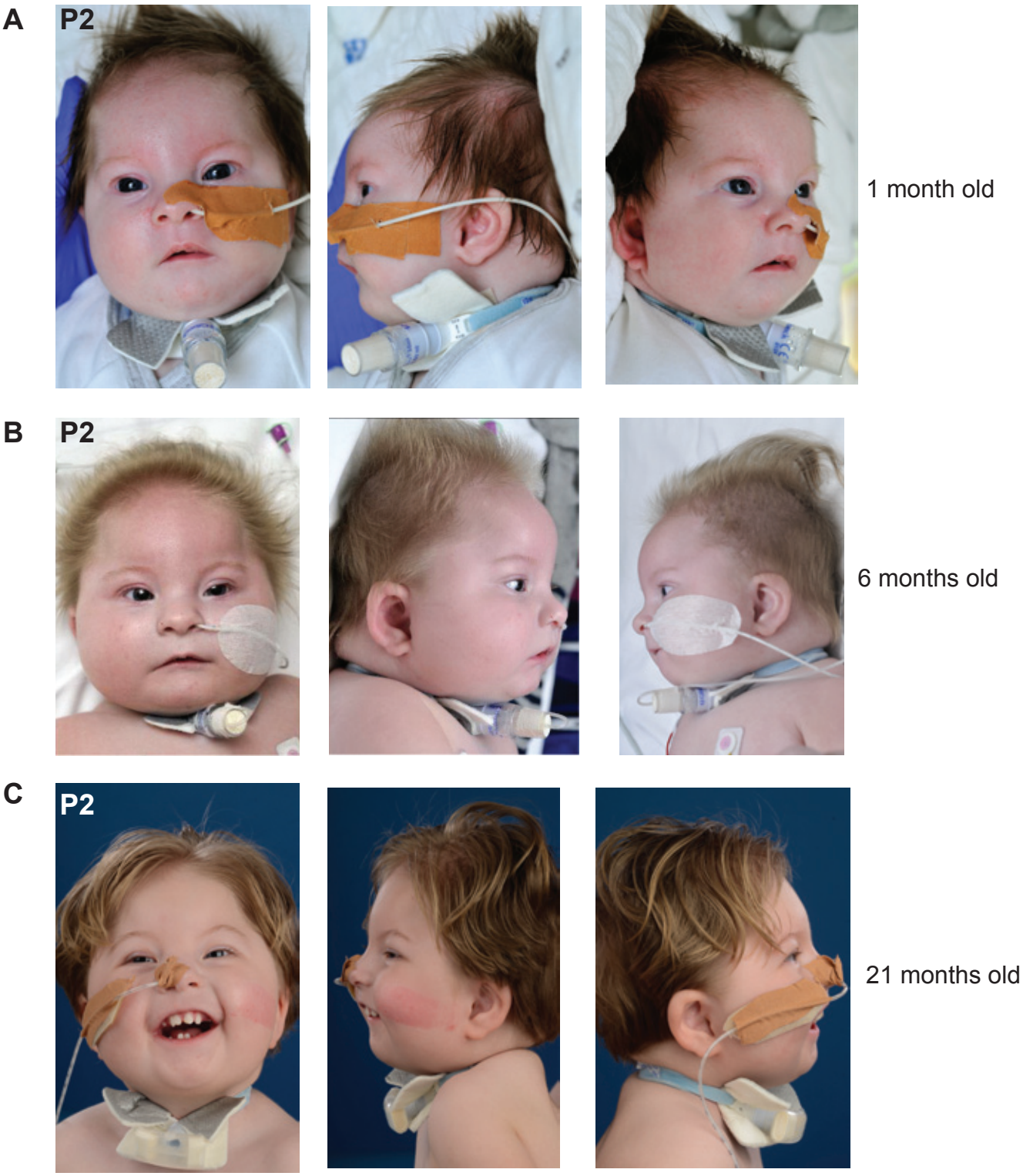


Figure S1. Photographs of P2 at different ages. Head and side portrait photographs of P2 at **A)** One month **B)** Six months and **C)** 21 months of age.

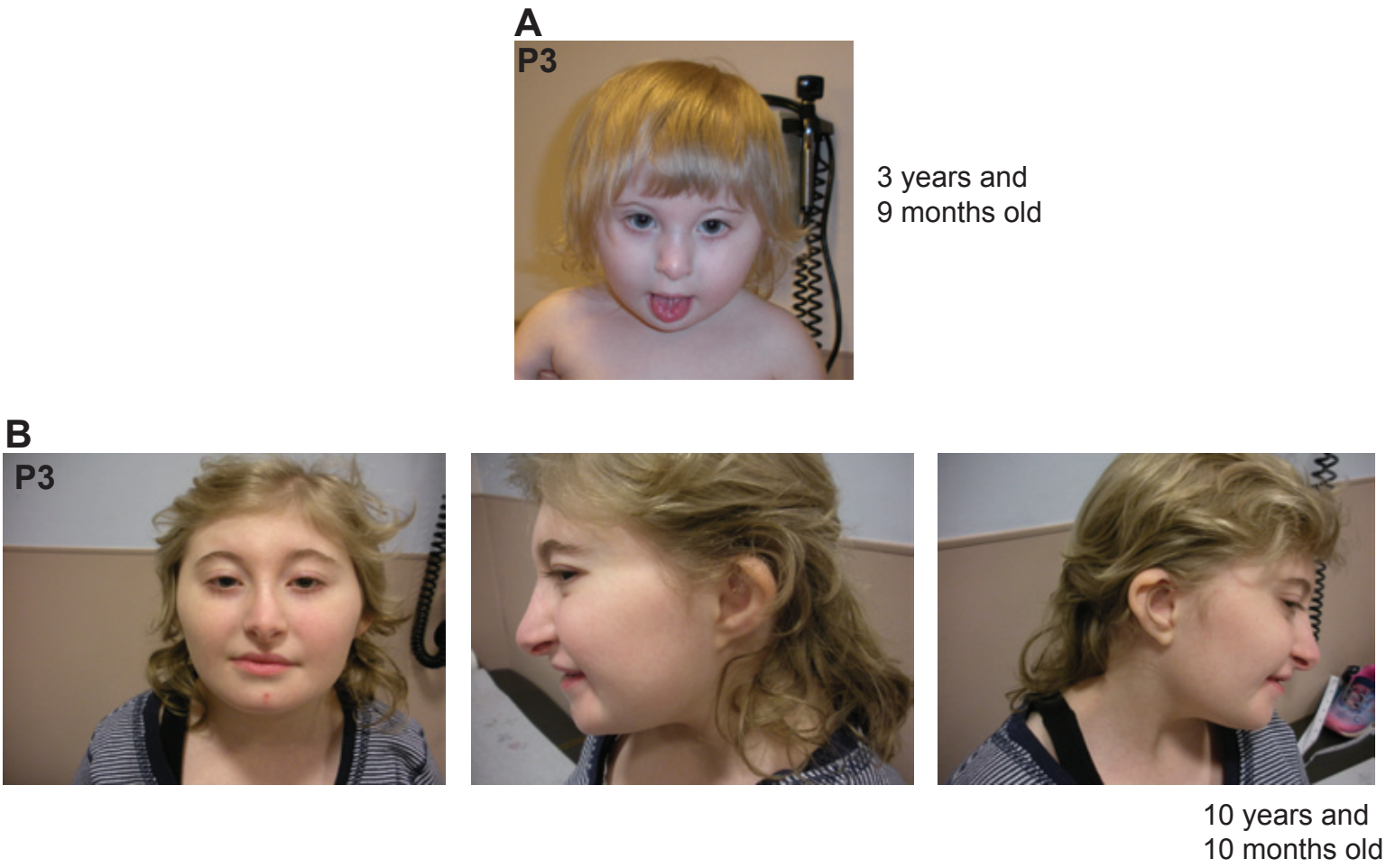
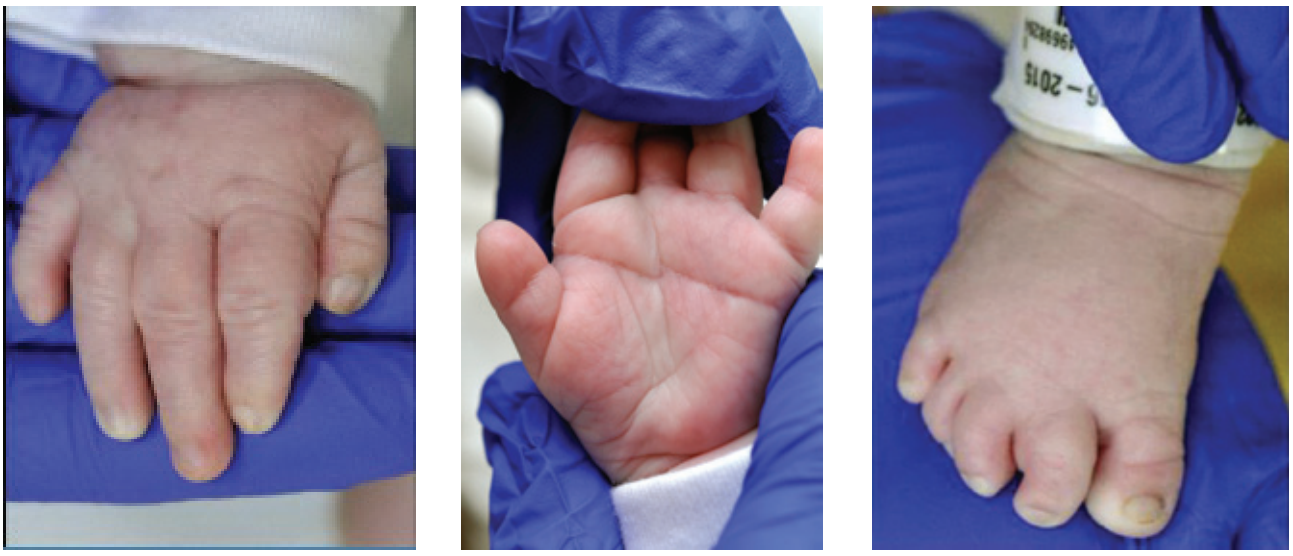


Figure S2. Photographs of P3 at different ages. Head and side portrait photographs of P3 at **A)** Three years and 9 months of age and **B)** 10 years and 10 months of age.

A



1 month old

B

6 months old

C

21 months old

Figure S3. Photographs of P2 at different ages. Photographs of the hands and feet of P2 at **A)** One month **B)** Six months and **C)** 21 months of age.

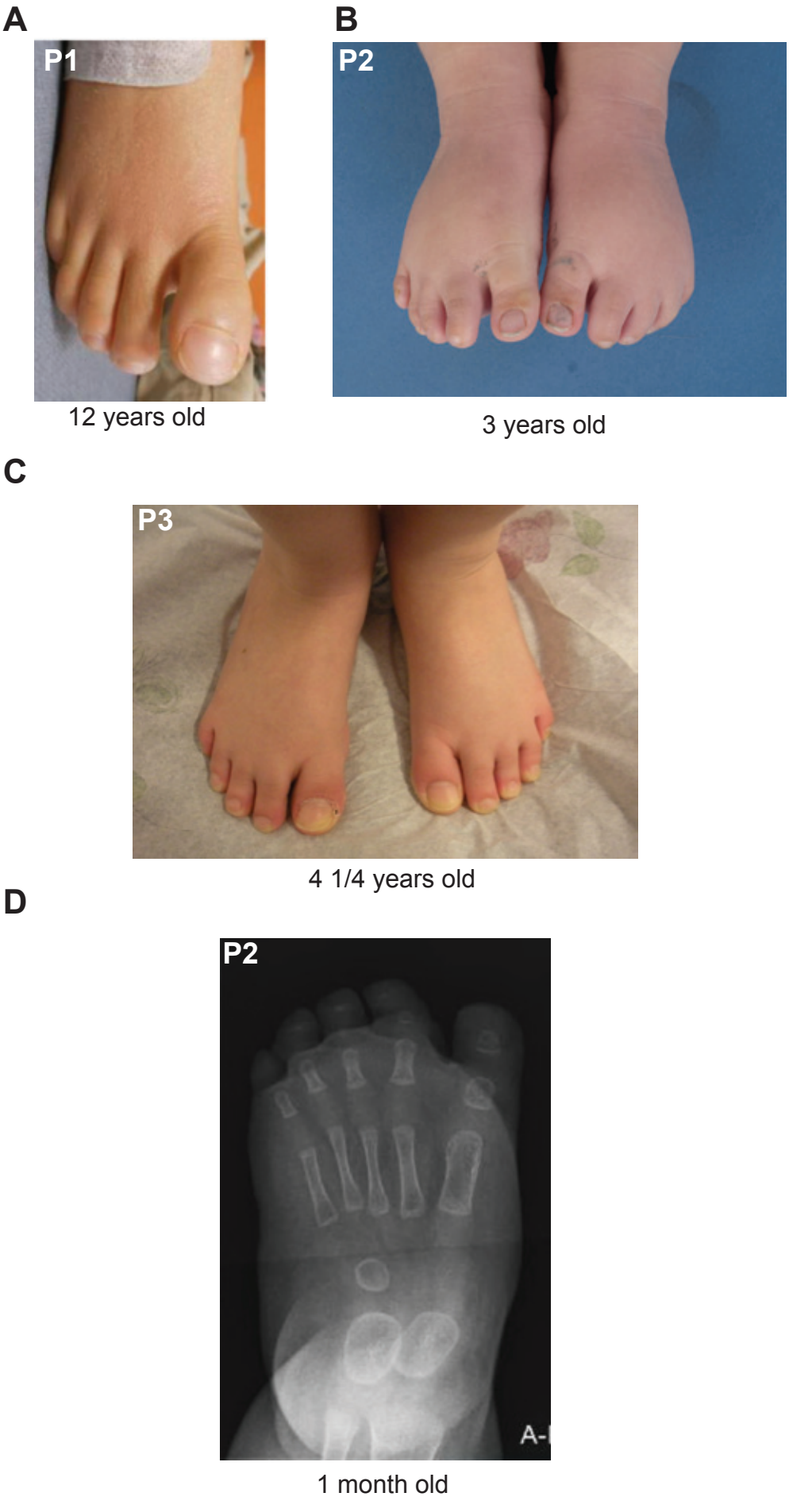


Figure S4. Foot features of individuals carrying variants in *HNRNPR.* Photographs of the feet of P1 **(A)**, P2 **(B)**, and P3 **(C)** taken at the ages of 12 years, 3 years, and 10 years, respectively showing broad 1st digit in P1 and P3 and small 5th digits with clinodactyly. **D)** X-ray of the foot of P2 taken at one month of age showing hypoplastic midphalanges of digits 2-3, hypoplastic endphalanges of digits 2-4 and absence of ossification of the mid phalanx of digits 4-5 and end phalanx of digit 5.

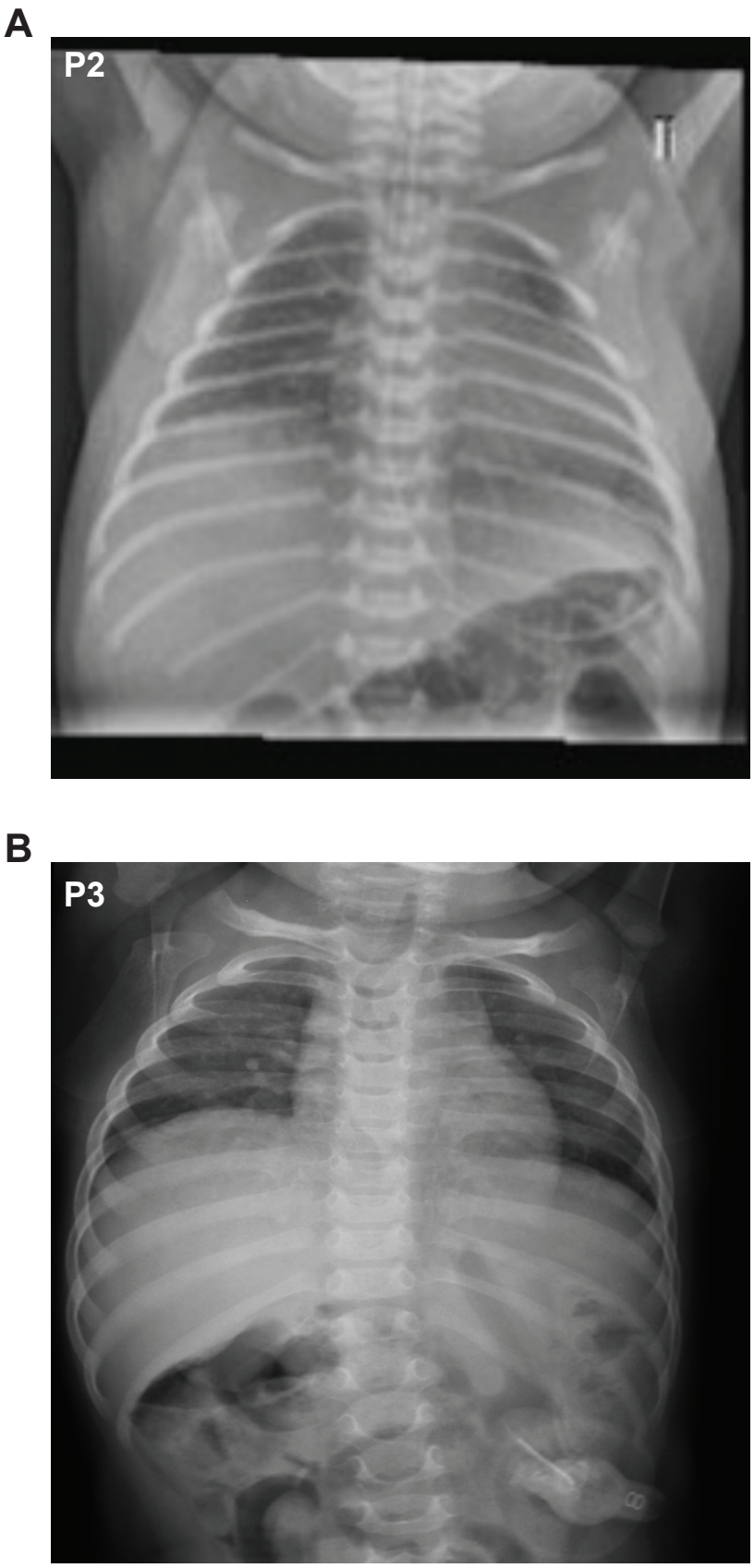


Figure S5. Skeletal features of individuals carrying variants in *HNRNPR.* **A)** X-ray of P2 at one week of age showing only 11 rib pairs. **B)** X-ray of P3 at 20.5 months of age showing only 11 rib pairs.

Figure S6. MRI of P4. MRI taken at 7 years and 10 months of age in an individual carrying the missense variant p.Arg588His in hnRNPR revealing cerebellar vermis hypoplasia (arrows).

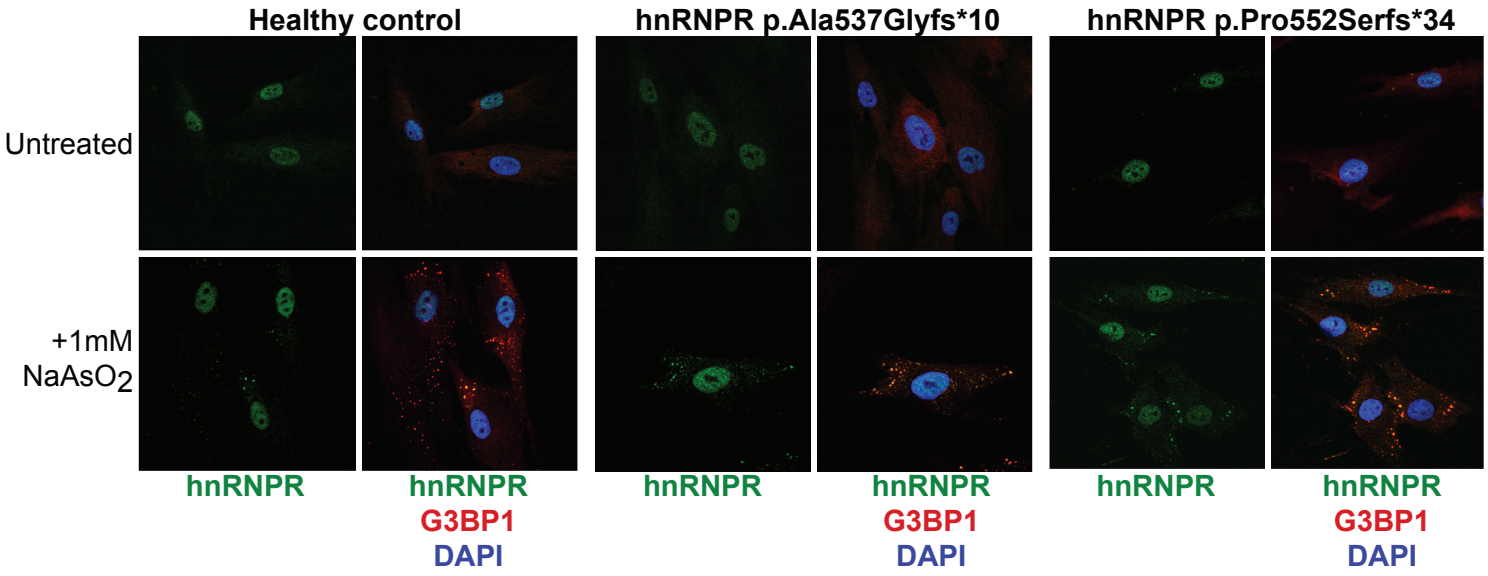


Figure S7. Truncated hnRNPR proteins do not drive stress granule formation at basal levels. Immunofluorescence of primary fibroblasts from healthy control individuals or those carrying variants truncating hnRNPR (P1 and P2). Cells were untreated (top) or treated with 1mM $NaAsO_2$ for 60' then stained with antibodies against hnRNPR (green), the stress granule marker G3BP1 (red), and DAPI (blue).

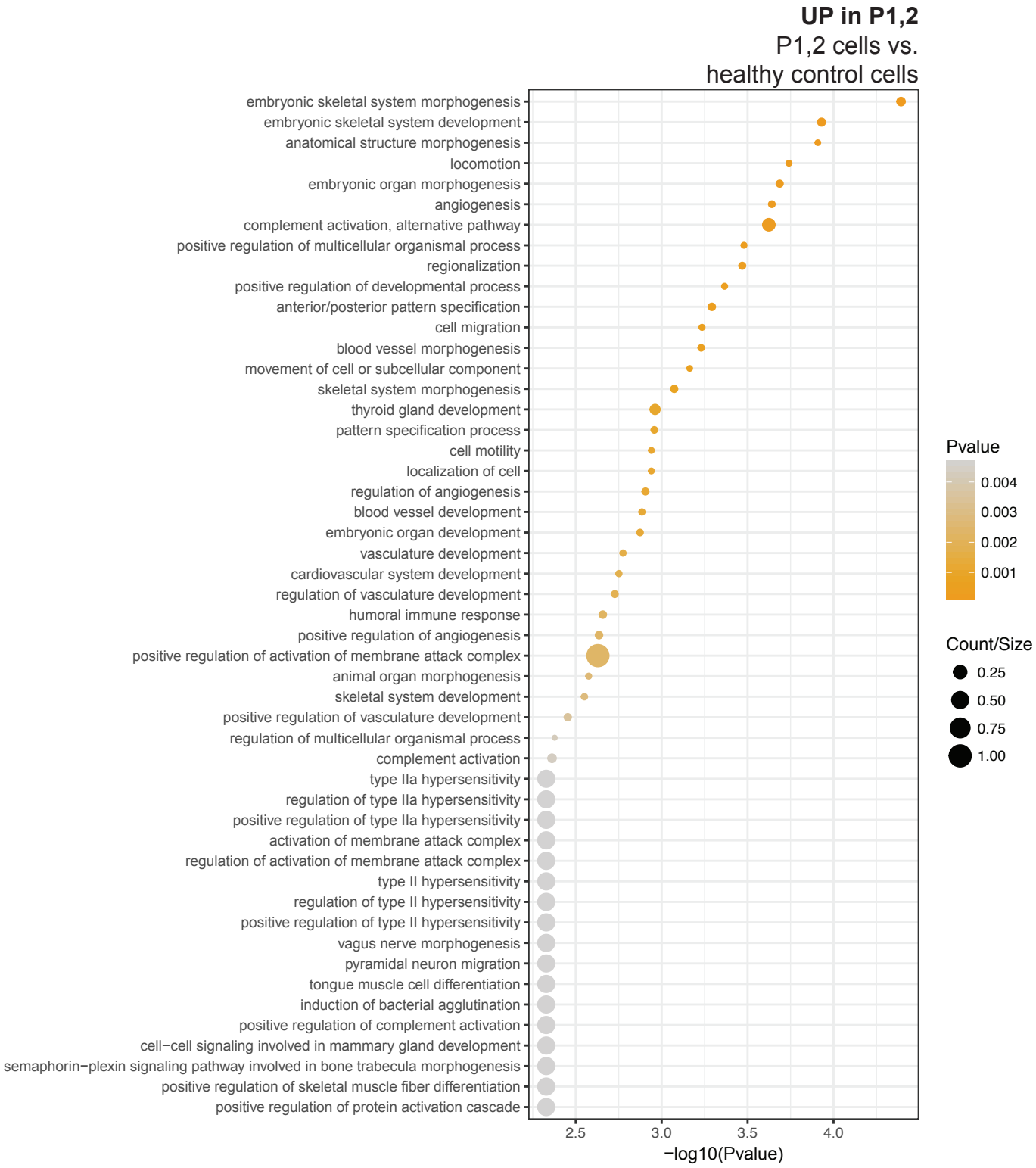


Figure S8. GO terms associated with upregulated RNAs in *HNRNPR* mutant cells. GO analysis of the differentially expressed RNAs (adjusted *p* value <0.05) deteremined by RNA-seq of untreated fibroblasts derived from individuals carrying *HNRNPR* mutations (P1 and P2) and healthy controls.

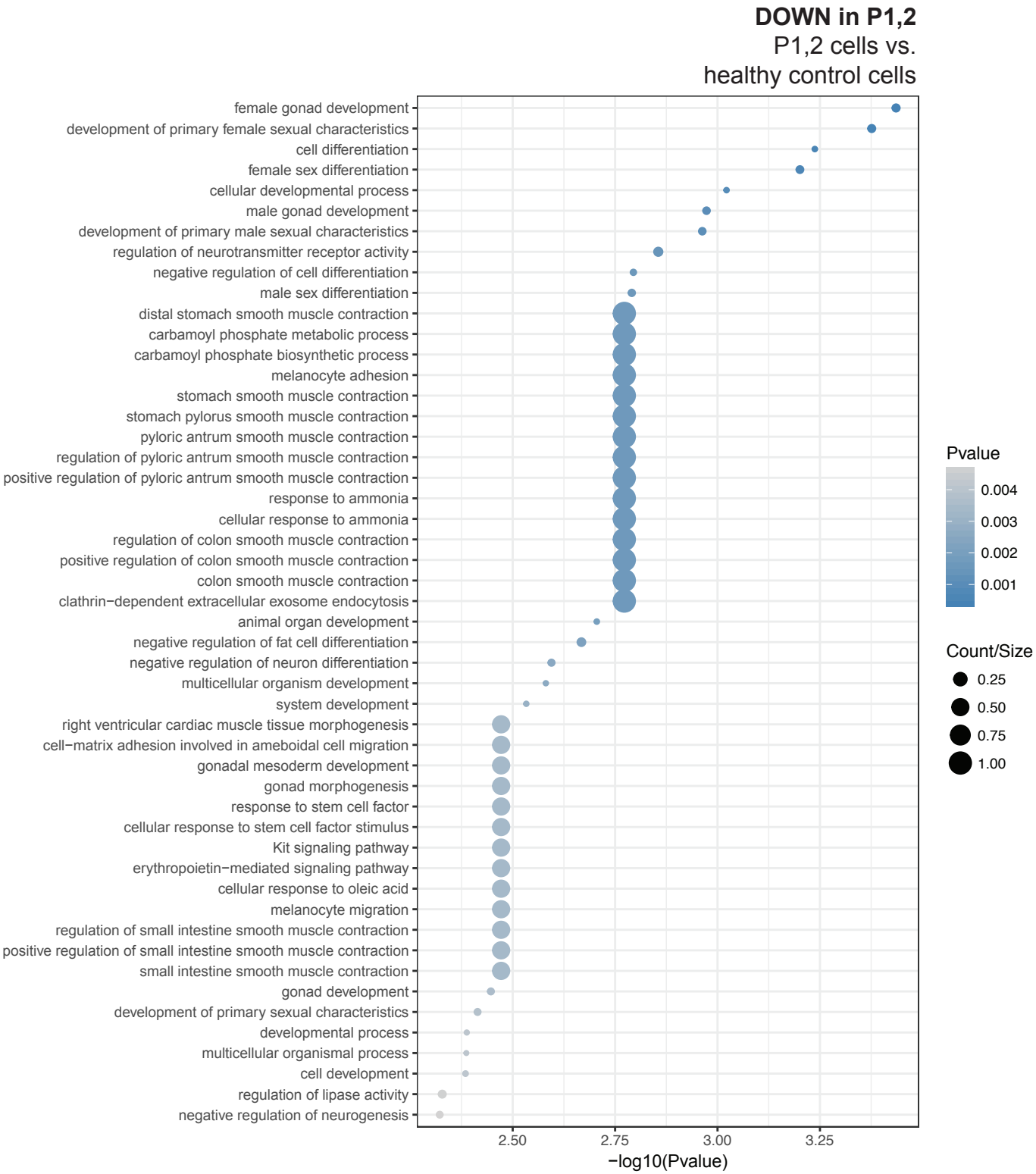


Figure S9. GO terms associated with downregulated RNAs in *HNRNPR* mutant cells. GO analysis of the differentially expressed RNAs (adjusted *p* value <0.05) deteremined by RNA-seq of untreated fibroblasts derived from individuals carrying *HNRNPR* variants (P1 and P2) and healthy controls.

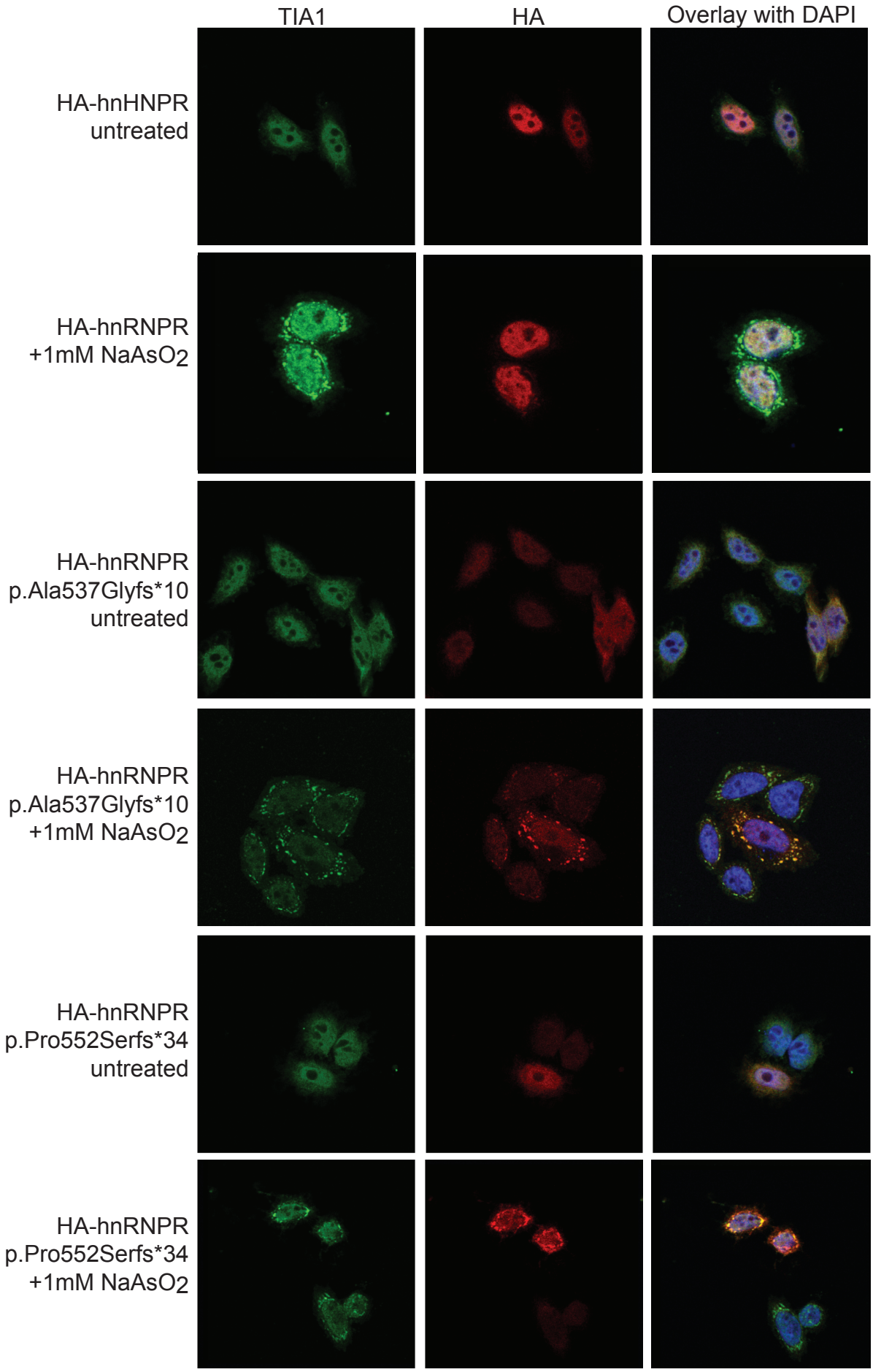


Figure S10. Truncated HA-hnRNPR variants co-localize more excessively with stress granules. HeLa cells transfected with HA-tagged hnRNPR contructs either untreated or exposed to 1mM $NaAsO_2$ for 30' and stained with antibodies against TIA1 (green), HA (red). DAPI (blue) is shown in the overlay.

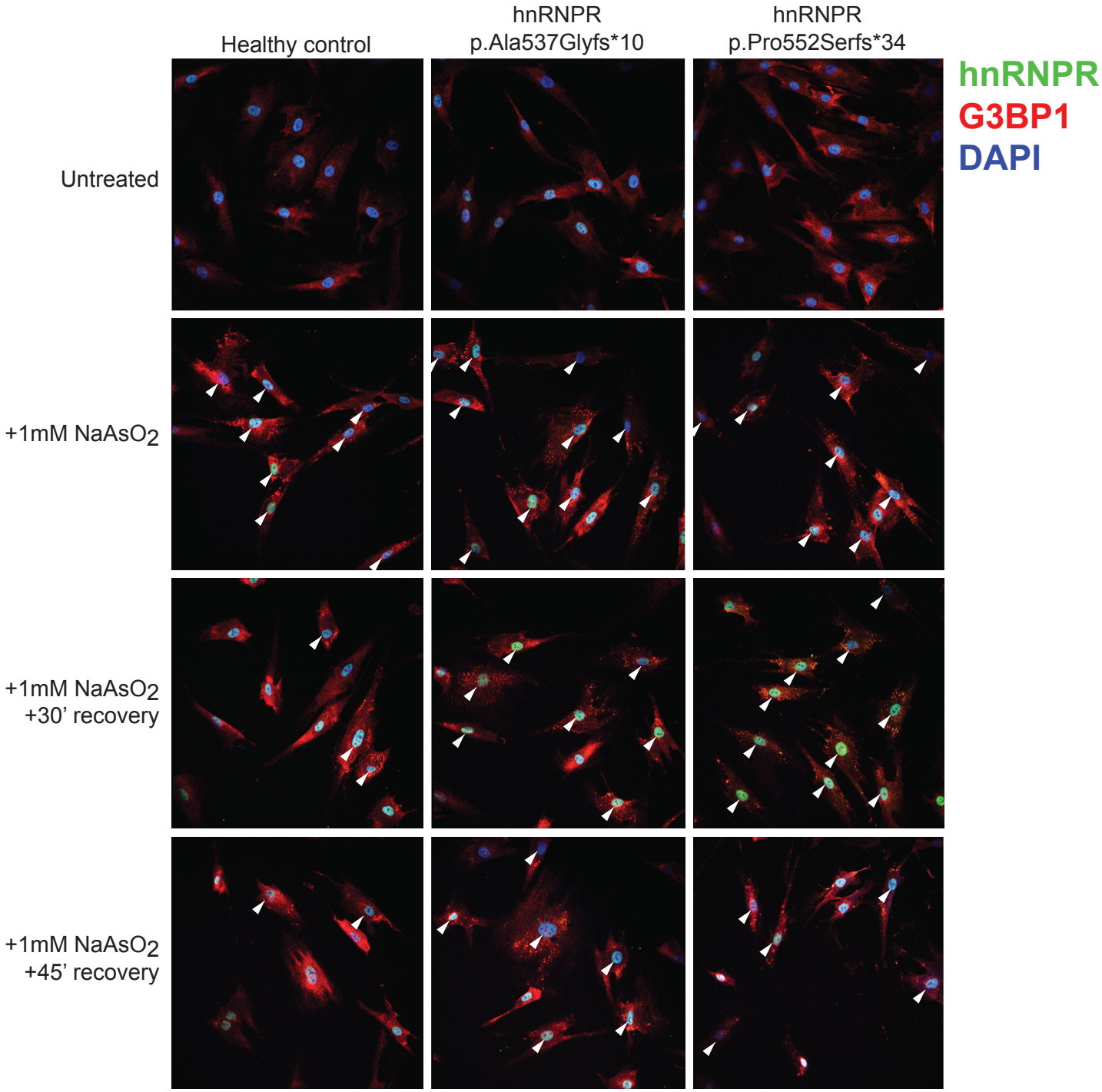


Figure S11. Stress granule disassembly is impaired in cells carrying truncated hnRNPR. Confocal microscopy of fibroblasts derived from a healthy control, P1 p.Ala537Glyfs*10, or P2 p.Pro552Serfs*34. Cells are untreated or treated with 1mM NaAsO2 for 30' and either immediately fixed or allowed to recover for 30' or 45'. Cells stained with DAPI (blue) and antibodies against hnRNPR (green) and the stress granule marker G3BP1 (red).Cells with stress granules are indicated with arrowheads.