ZNF423 patient variants, truncations, and in-frame deletions in mice define an allele-dependent range of midline brain abnormalities.

Ojas Deshpande, Raquel Z. Lara, Oliver R. Zhang*, Dorothy Concepcion, and Bruce A. Hamilton

Department of Cellular and Molecular Medicine

Department of Medicine

Institute for Genomic Medicine

Rebecca and John Moores UCSD Cancer Center

University of California, San Diego School of Medicine

9500 Gilman Drive

La Jolla, CA 92093

* Current address: Rocky Vista University College of Osteopathic Medicine, 8401 S Chambers Rd, Parker, CO 80134

ORCIDs

OD: 0000-0002-6029-2370
RZL: 0000-0003-3769-9313
ORZ: 0000-0001-9710-9843
DC: 0000-0002-8506-1229
BAH: 0000-0001-5599-9139

Short title: *ZNF423* patient variants, truncations, and deletions in mice

ABSTRACT

Interpreting rare variants remains a challenge in personal genomics, especially for disorders with several causal genes and for genes that cause multiple disorders. *ZNF423* encodes a transcriptional regulatory protein that intersects several developmental pathways. *ZNF423* has been implicated in rare neurodevelopmental disorders, consistent with midline brain defects in *Zfp423*-mutant mice, but pathogenic potential of most patient variants remains uncertain. We engineered ~50 patient-derived and small deletion variants into the highly-conserved mouse ortholog and examined neuroanatomical measures for 791 littermate pairs. Three substitutions previously asserted pathogenic appeared benign, while a fourth was effectively null. Heterozygous premature termination codon (PTC) variants showed mild haploabnormality, consistent with loss-of-function intolerance inferred from human population data. In-frame deletions of specific zinc fingers showed mild to moderate abnormalities, as did low-expression variants. These results affirm the need for functional validation of rare variants in biological context and demonstrate cost-effective modeling of neuroanatomical abnormalities in mice.

(149 words)

AUTHOR SUMMARY

Gene identification in rare disorders is typically supported by finding different mutations of the same gene in multiple families with the same disorder. However, causal evidence for any specific mutation found in one or a few related individuals is weaker, especially if the disorder can be caused by any of several genes and the functional effect of the mutation is not certain. Experimental models can be helpful in testing causal effects, but only to the extent that the model is validated to recapitulate one or more aspects of the disorder. We used CRISPR/Cas9-based genome engineering to create a wide range of mutations in mouse *Zfp423*, whose human cognate is implicated in neurodevelopmental disorders, especially cerebellar vermis hypoplasia and Joubert syndrome. This large collection of animal models shows that both reduced *Zfp423* expression, including heterozygosity for loss-of-function mutations, and normally-expressed domain deletions, including specific zinc finger domains, produce measureable abnormalities in midline development. Despite this high level of validation, most patient-derived amino acid substitution variants tested do not produce measureable effects. The single exception is a substitution, H1277Y, that destroys a structural element in the last zinc finger domain and results in dramatic loss of steady-state Zfp423 protein level.

INTRODUCTION

Variant effect prediction remains a challenge in medical genomics [1, 2]. Progress from large reference databases such as ExAC [3], gnomAD [4], and UK Biobank [5] allows powerful statistical evidence against pathogenicity, based on allele frequency [6] for rare variants that had appeared unique to patients in smaller samples. Recessive phenotypes, low or context-dependent penetrance, effects on pre-term viability, and demography, however, may create exceptions often enough to be relevant to patients with rare disorders. Population frequencies also provide limited guidance for singleton and de novo variants. Predictive algorithms based on evolutionary constraint, physico-chemical similarity between residues, or average replacement effects in deep mutational scanning data continue to improve, but these constraints are neither necessary nor sufficient for disease association as loss of human-specific traits may present as disease while evolution selects on subtler variation than disease presentation. Attempts to model prediction accuracy can suffer where ground truth is not available and clinical variant databases in current use include assertions often based on limited evidence. The problem can be particularly acute for disorders where a substantial number of genes are mutable to overlapping phenotypes, including ciliopathies such as Joubert syndrome and related disorders (JSRD). For example, *ZNF423* mutations have been reported as pathogenic in JSRD patients [7] and other neurodevelopmental disorders [8], but most patient variants have uncertain significance and even those asserted pathogenic in public databases rely on very limited data. This is true for many rare disorders.

Mice can be a useful model for *ZNF423* function. The ZNF423 orthology group is highly constrained across vertebrates [9]; after accounting for annotation differences in alternative 5' exons, mouse Zfp423 and human ZNF423 share >98% amino acid identity (99% in zinc finger domains). Most human variants will therefore be in sequence

contexts that are similarly constrained in mice. Null mutations in mouse *Zfp423* have defects in midline brain development similar to human JSRD, including anterior rotation of the cerebellar hemispheres and hypoplasia or agenesis of the cerebellar vermis with more modest effects on forebrain structures [10-12]. Roof plate defects in Zfp423 mutant mice also impact development of hindbrain choroid plexus [10, 13]. Other work in mice showed notable effects of Zfp423 on olfactory neurogenesis [14], neocortex development [15], adipogenesis [16-18], and wound healing [19]. Gene-trap alleles that reduced expression of an otherwise normal Zfp423 protein showed hypomorphic phenotypes, indicating a graded response to genetic function [10]. In cerebellum, loss of Zfp423 prevents or limits response to SHH by granule precursor cells ex vivo, consistent with a functional abnormality in the cilium [20]. ZNF423 homologs interact with a diverse set of lineage-determining or signal-dependent transcription factors in alternate and potentially competing complexes [7, 21-26] and mutational effect will likely depend on which if any of these contacts is altered [27]. Cell-based models may afford screening of all potential variants in a protein [28-30], but the limited context of cells ex vivo could miss key features of ZNF423 function as its expression is dynamic across communicating cell types during development, including both germinal zones in the cerebellum. Whether (or to what extent) all ZNF423 interactions and functions are required in any one cell type is not clear. With the ability to multiplex germline editing at high efficiency and specificity, mouse brain development might therefore be the simplest robust assay for impact of ZNF423 variants on human brain development, allowing quantitative assessment of variant effects on brain development at a scale commensurate with ascertainment of rare disease patient alleles while being agnostic to developmental stage or cell types in which specific functions are compromised.

Here we developed simple, quantitative measures with good statistical power to assess structural brain abnormalities in ~50 mouse strains with *Zfp423* mutations created by genome editing. This extensive set of comparisons allowed us to show differences in sensitivity among phenotypes, test pathogenic potential of patient-derived and other variants, and identify previously unreported haploabnormality in null allele heterozygotes. Among patient-derived substitution alleles, H1277Y, at a zinc-coordinating histidine in the last of 30 C2H2 zinc fingers, was effectively null. By contrast, three other patient variants asserted pathogenic or likely pathogenic based on single patients and algorithmic predictions (R89H, P913L, and E1124K) appeared benign. Premature truncation variants, including humanized alleles that encode protein tails to model patient frame-shift variants, were predominantly null with no evidence for dominant negative activity. An early frame-shift variant in exon 3 was an exception, evading nonsense-mediated decay to produce a partial protein at reduced abundance and a partial loss-of-function phenotype. Null allele heterozygotes showed slightly lower weight, smaller cerebellar vermis, and shorter stride length than control littermates, providing functional evidence for loss-of-function intolerance observed in human population data. In-frame deletions had a range of domain-specific effects. Deletions that remove zinc finger 1, zinc fingers that bind BMP-dependent SMAD proteins, or a non-motif region containing two CXXC sites, and deletions that reduced overall protein expression showed measureable effects, while deletion of zinc finger 12 did not.

RESULTS

Induction of patient-derived variants and collateral mutations.

We developed an editing pipeline using standard CRISPR/Cas9 tools. We prioritized 13 *ZNF423* amino acid substitution variants and two frame-shifting alleles from patients

(S1_Table). Four substitution alleles and one frameshift were published [7, 8] and reported in ClinVar [31] as pathogenic or likely pathogenic; a second frameshift allele was reported in MyGene2 [32]. Other substitution alleles were patient variants of uncertain significance communicated by Drs. Joseph Gleeson and Friedhelm Hildebrandt. Targets were selected to include a range of predicted effects in commonly used variant effect algorithms (Table 1) and to include a range of allele frequencies in databases depleted for close relatives and patients with Mendelian disorders [3, 4]. We edited FVB/NJ embryos because this strain background improved both postnatal viability of the *Zfp423*^{nur12} null mutation and heterozygote breeding performance relative to C57BL/6J, without compromising penetrance of severe ataxia among surviving homozygotes (Alcaraz et al., bioRxiv 2020/091629 in process). For most variants, we co-injected two or more alternative repair template oligonucleotides (S2_Table) to create silent substitution controls or edit adjacent sites in a single injection series.

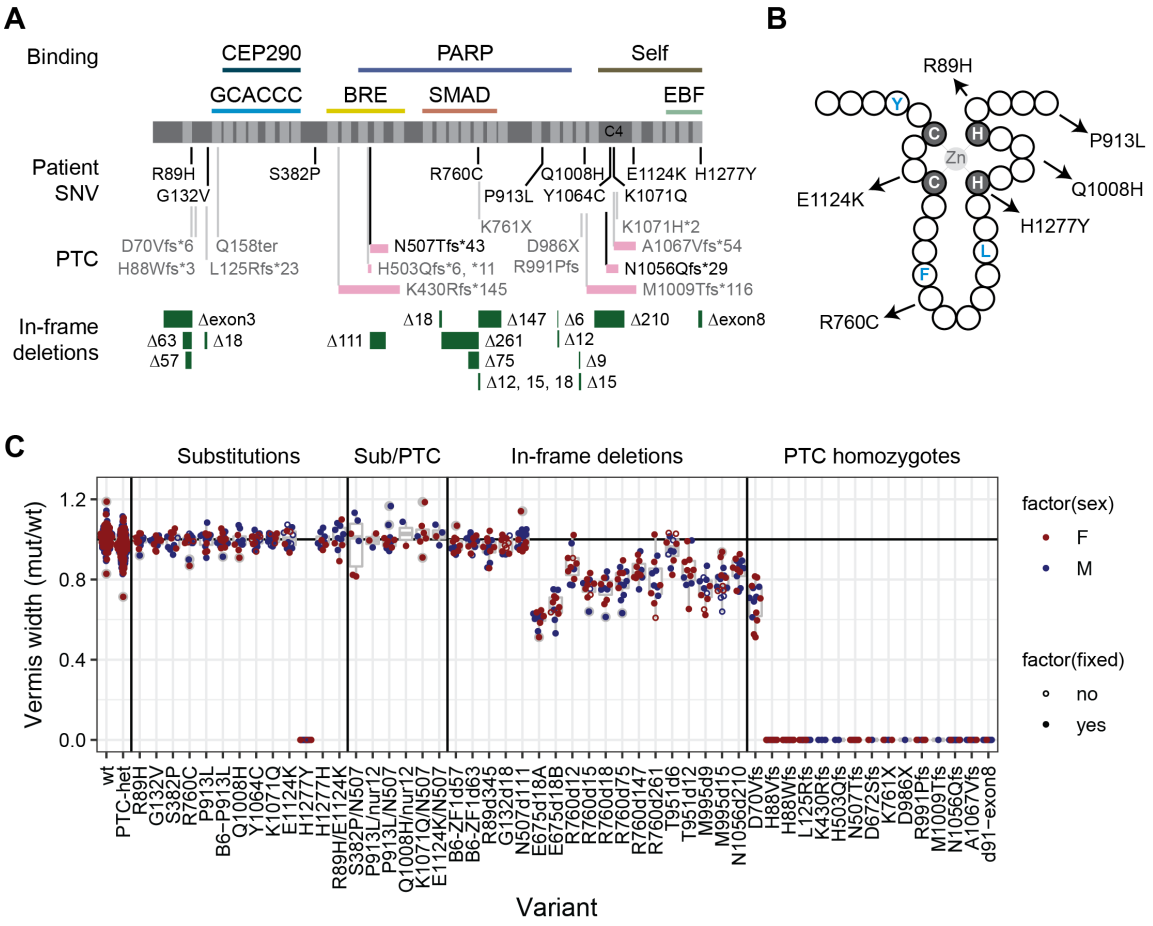
Table 1. Effect predictions for *ZNF423* substitution variants modeled in this study. Variant (domain) shows position relative to NP_ 055884.2, single-letter amino acid codes, and position relative to C2H2 zinc fingers (ZF). Column hg38 shows nucleotide position in the hg38 reference assembly with reference/variant nucleotide. For the gnomAD database (v2), allele count and minor allele frequency (MAF) are given. ClinVar assertions were current at time of writing. Categorical calls and scores are shown for 8 variant effect predictors.

Variant (domain)	hg38	gnomAD (MAF)	ClinVar 2019	PolyPhen 2 (HVAR)	SIFT (Score)	PROVEAN	Mutation Taster (rankscore)	Mutation Assessor (rankscore)	VEST3 Rank score	CADD PHRED	Envision
R89H (ZF1)	16:49730782 C/T	18 (6.4 e-5)	Likely pathogenic	Benign (0.001)	Tolerated (0.282)	Neutral (0.19)	Polymorphism (0.261)	Neutral (0.016)	0.126	19.16	0.99
G132V (ZF1-2 linker)	16:49638757 C/A	0	-	Damaging (1)	Damaging (0.002)	Deleterious (-2.67)	Disease causing (0.81)	Low (0.225)	0.883	25.6	0.85

S382P (ZF8-9 linker)	16:49638008 A/G	215 (7.6 e-4)	Uncertain significance	Possibly damaging (0.72)	Tolerated (0.254)	Neutral (-1.65)	Disease causing (0.345)	Low (0.498)	0.682	24.4	0.92
R760C (ZF18)	16:49636874 G/A	11 (3.8 e-5)	-	Probably damaging (0.82)	Damaging (0.001)	Deleterious (-4.83)	Disease causing (0.588)	Low (0.304)	0.904	28.9	0.74
P913L (ZF21)	16:49636414 G/A	47 (1.7e-4)	Pathogenic	Probably damaging (0.92)	Damaging (0.010)	Deleterious (-3.87)	Disease causing (0.81)	Medium (0.833)	0.842	28.5	0.89
Q1008H (ZF24)	16:49636128 C/G	7 (2.8e-5)	-	Probably damaging (0.88)	Damaging (0.000)	Deleterious (-3.61)	Disease causing (0.457)	Medium (0.53)	0.848	25.3	0.93
Y1064C (C4)	16:49635961 T/C	1 (4.2e-6)	-	Probably damaging (0.75)	Damaging (0.031)	Deleterious (-5.02)	Disease causing (0.548)	Low (0.246)	0.815	28.1	0.74
K1071Q (C4)	16:49635941 T/G	0	-	Probably damaging (0.65)	Damaging (0.036)	Neutral (-1.52)	Disease causing (0.465)	Low (0.246)	0.577	28.5	0.92
E1124K (ZF26)	16:49635782 C/T	8 (3.2e-5)	Likely pathogenic	Possibly damaging (0.54)	Tolerated (0.092)	Neutral (-1.64)	Disease causing (0.548)	Medium (0.562)	0.883	23.6	0.98
H1277Y (ZF30)	16:49491301 G/A	0	Pathogenic	Benign (0.04)	Damaging (0.002)	Deleterious (-3.93)	Disease causing (0.53)	Medium (0.924)	0.951	24.5	0.92

We recovered 12 out of 15 designed mutations–10 of 13 intended substitution alleles and each of two patient frame-shift variants with humanized codons–as well as designed silent control edits and a large number of collateral mutations (Figure 1A, S3_Table) and verified transmission by genotyping (S4_Table). Substitution alleles included sites that are highly conserved among vertebrates, including seven that were invariant among 165 curated orthologs [9]. Most lie in or adjacent to zinc finger (ZF) domains (Figure 1B) or a non-motif region containing a quartet of conserved cysteine residues with potential to form a C4-class zinc finger and for which common annotation tools showed wide disagreement on functional predictions (Table 1). Collateral variants included predicted loss of function (pLOF) alleles at several positions in the coding sequence and in-frame deletions that allowed observations of protein stability and domain-specific function in vivo. We included several variants that remove specific zinc fingers or conserved regions between zinc fingers in subsequent analyses. A simple measure of cerebellar vermis

width from surface views of the brain was sensitive to heterozygosity for presumed null alleles and specific to variants that removed critical residues or decreased protein abundance, while comparison among non-mutant littermate pairs suggested high sensitivity and power to detect modest differences in 10-15 sample pairs (Figure 1C, S5_Table, and results below). Importantly, none of the variant effect predictions (Table 1) correctly predicted outcomes with respect to disease-relevant phenotypes (Figure 1C and below).


A
Binding
CEP290
PARP
Self
GCACCC
BRE
SMAD
EBF
C4
Patient SNV
R89H
G132V
S382P
R760C
Q1008H
E1124K
H1277Y
P913L
Y1064C
K1071Q
K761X
K1071H*2
PTC
D70Vfs*6
H88Wfs*3
Q158ter
L125Rfs*23
N507Tfs*43
H503Qfs*6, *11
K430Rfs*145
D986X
R991Pfs
A1067Vfs*54
N1056Qfs*29
M1009Tfs*116
In-frame deletions
Δexon3
Δ63
Δ18
Δ57
Δ111
Δ18
Δ147
Δ261
Δ75
Δ12, 15, 18
Δ6
Δ12
Δ9
Δ15
Δ210
Δexon8
B
R89H
P913L
Q1008H
E1124K
H1277Y
R760C
Zn
C
H
Y
L
F
C
H
C
H
C
H

Frame-shift PTC variants show relative length of altered reading frame (pink bars). N507Tfs*43 and N1056Qfs*28 were humanized to encode the same frame-shifted peptide as reported patient variants. For in-frame deletions (green bars), the number of deleted base pairs is shown. (B) Schematic of a C2H2 zinc finger shows relative positions of patient substitution alleles relative to consensus hydrophobic (blue) residues and required zinc-coordinating residues (grey background). (C) Ratio of mutant to control vermis width measured from surface views of same-sex littermate pairs. Each dot represents one littermate pair. Female pairs, red, male pairs, blue. A few pairs were measured as freshly dissected, unfixed material (open circles), which did not noticeably affect the relative measure. PTC heterozygous, H1277Y homozygotes, homozygotes for each in-frame deletion except N507d111, and all PTC homozygous variants were significantly different from both the expected null model and empirical wild-type:wild-type comparisons.

Simple measures are highly sensitive to *Zfp423* variants.

To test variants for pathogenic potential at moderate to large scale, we looked for phenotypes that were easily obtained and robust across trials. We observed home cage behavior and videographed mice walking across an open stage to assess gross locomotor activity for each variant subjectively. We quantified viability, weight, and anatomical parameters of brains in both surface view and block face photographs as quantitative phenotypes (S6_Table). A subset of null-allele heterozygotes was tested in more detail for locomotor behaviors (S7_Table). To avoid potential confounding effects of age, sex, maternal care, or other factors that might vary across a large colony, all tests were conducted on co-housed, same-sex littermates and scored as mutant/control ratio (anatomy) or differences (behavioral latencies). Variants were studied concurrently by investigators blinded to genotypes.

Each measure showed large effects of *Zfp423* premature termination codon (PTC) alleles in homozygotes, with gross ataxia and cerebellar vermis hypoplasia showing complete penetrance. Predicted null variants had no detectable Zfp423 protein on Western blots, with a detection limit ≤5% of control littermate levels, except D70Vfs*6, which showed a reduced level of a lower molecular weight protein unique to that allele (Figure 2A and S1_Fig). For protein-null alleles (PTC variants except D70Vfs*6), homozygotes were recovered at reduced frequency and several null animals identified at P10-P15 died prior to assessment, showing a relatively broad window for lethal events (Figure 2B). Among survivors, locomotor disability was severe and never overlapped controls (Supporting videos). None of 72 *Zfp423*-null brains had measureable vermis in surface photographs (Figure 1C). Nearly half of those with mid-sagittal block face views (30/66) showed a small amount of cerebellar tissue, largely due to compression of lateral hemispheres toward the midline (Figure 2C-D). Midline fiber tract (corpus callosum and anterior commissure) and cortical thickness measures distinguished mutant from control groups, but with smaller relative magnitude and incomplete penetrance (Figure 2E-J). Among forebrain phenotypes, corpus callosum had the largest effect and highest penetrance, while anterior commissure effects were only evident across a larger population of animals. PTC homozygotes were smaller than littermates, with lower weight at sacrifice (Figure 2K-L). These results generalized findings from earlier studies and put quantitative parameters on magnitude and penetrance for hindbrain, forebrain, and weight measures.

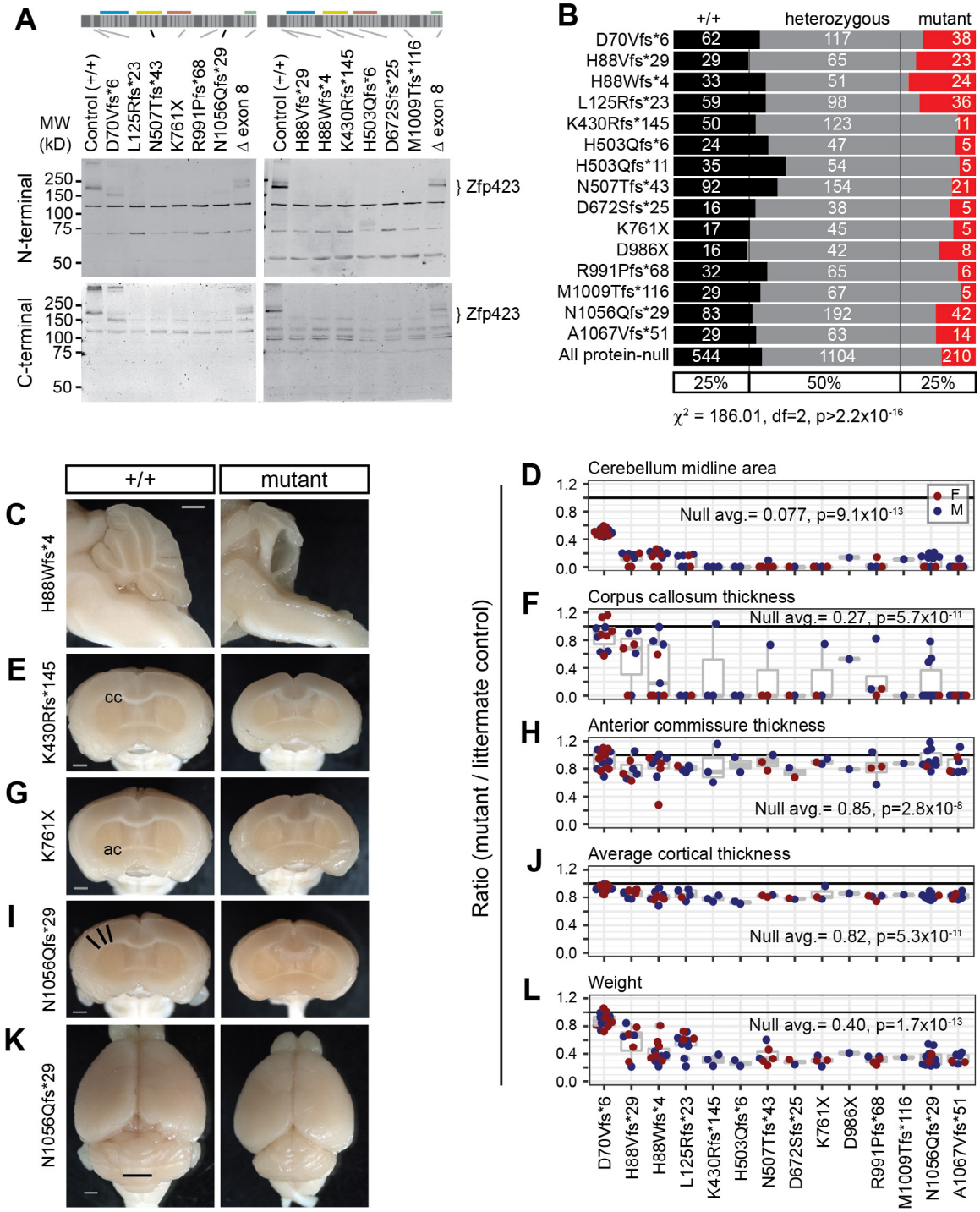


Figure 2. *Zfp423* frameshift and nonsense mutations are effectively null, except D70Vfs*6. (A) Western blots detected full length Zfp423 in perinatal cerebellum of control samples. Among PTC homozygotes, D70Vfs*6 and a deletion of terminal exon 8 showed altered-size proteins at reduced levels. No consistent evidence for residual

protein was seen for any other PTC variants, detection threshold ≤5% wild-type level. N-terminal, Bethyl A304-017A. C-terminal, Millipore ABN410. Cross-reacting background bands were independent of PTC position. (B) Reduced frequency of homozygotes for each PTC at biopsy (P10-P20) from breeding records. Summary chi-square is for protein-null alleles (excluding D70Vfs*6). (C) Mid-sagittal images showed variable amount of midline cerebellar tissue in mutants, with residual tissue attributable to hemispheres. (D) Ratio of midline cerebellum area (mutant/control) from block face images. Coronal block face images showed abnormal forebrains, including (E, F) disrupted or reduced corpus callosum (cc), (G, H) reduced anterior commissure (ac), and (I, J) reduced cortical thickness, measured as the average radial distance at 15°, 30° and 45° from midline (black lines). (K) Representative surface views. Vermis width as plotted in Figure 1 is indicated (black line). (L) PTC mutants had reduced body weight at sacrifice. (D, F, H, J, L) Averages and Wilcoxon signed-rank test p-values for ratio = 1 for combined data from all PTC alleles excluding D70Vfs*6 are shown. Female pairs red, male blue. Scale bars, 1 mm.

Loss of function heterozygotes showed reduced vermis, weight, and stride without evidence for dominant negative activity.

Having a large set of protein-negative PTC alleles across the full of the coding sequence allowed us to compare presumptive null alleles and distinguish potential haploabnormalities from dominant negative effects. PTC variants whose RNAs escape nonsense-mediated decay often enough to produce detectable levels of variant protein can have dominant negative genetic properties, by decoupling functional domains relative to the non-mutant protein [33]. Early PTCs can result in protein translation from an alternative initiation codon [34, 35] or exon skipping, while late PTCs could produce truncated proteins if they escape nonsense-mediated decay (NMD). *Zfp423* has a highly

unusual gene structure, including 4-kb internal coding exon (Figure 1) and variant effects in situ might differ from assays performed in cell culture on compact gene structures according to the proposed “long exon rule” [35, 36]. In Chaki et al. [7], one of us (B.A.H.) speculated that JBTS19 patients carrying one PTC variant heterozygous to an apparently normal allele might have dominant negative activity, supported by transfection assays with a corresponding cDNA in a human cell line. To test this in vivo and to learn more about how PTC variants behave in the context of a very large exon structure, we examined 16 mouse lines carrying different PTC variants at distinct positions relative to exon boundaries and encoded protein domains.

Heterozygotes for PTC variants that do not produce detectable protein should also be a good test of sensitivity, since multiple labs previously reported only recessive phenotypes from several alleles [10-12, 27]. Physical measures were slightly decreased on average in heterozygotes compared to sex-matched littermates across PTC variants for which homozygotes were protein-negative. Vermis width (Figure 3A) and weight at sacrifice (Figure 3B) each had a ~3-4% decrease with strong statistical support ($p<10^{-6}$), while anterior commissure measure had a 1.4% decrease with nominal support ($p=0.046$ after Benjamini-Hochberg correction for false discovery). In contrast, average thickness of cortex ($p=0.11$) and corpus callosum ($p=0.38$) showed little evidence for heterozygote effects, perhaps due to the modest effect size on cortex even in null animals and the comparatively high variance on corpus callosum measures.

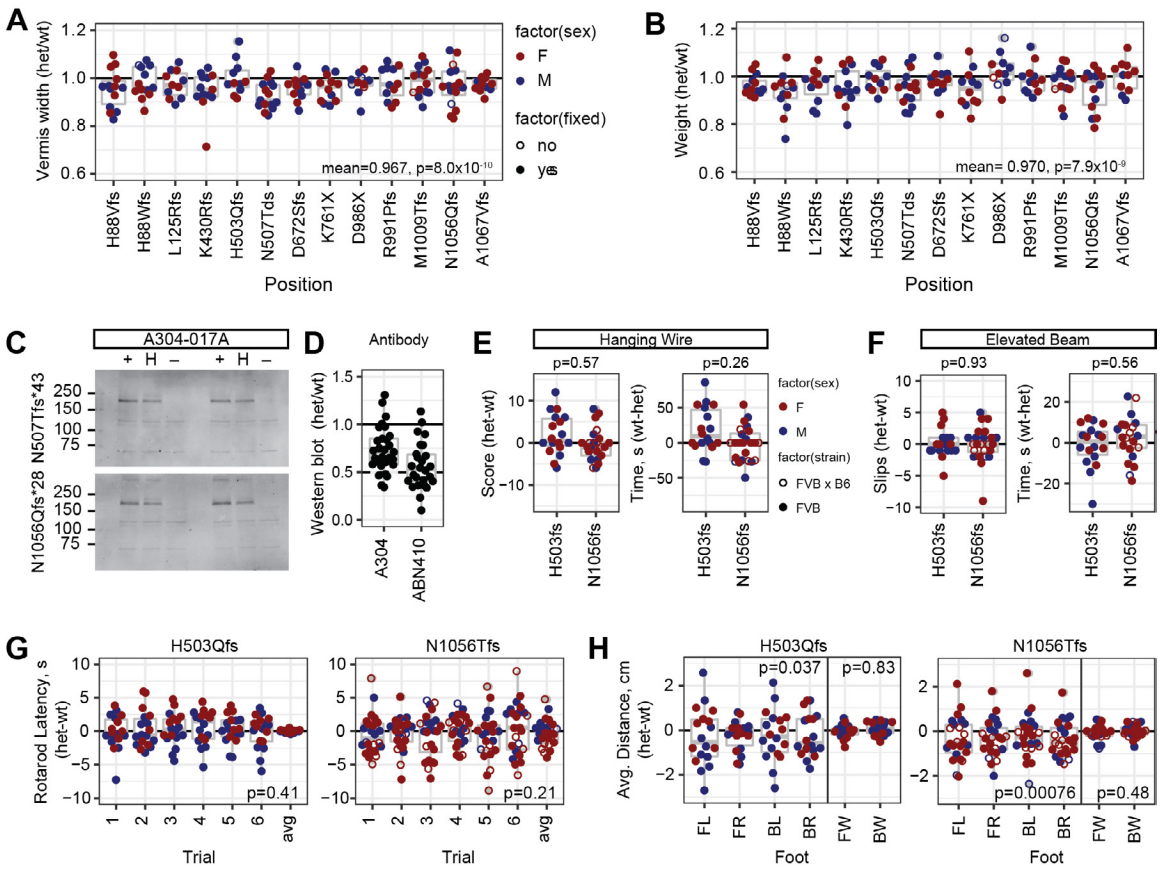


Figure 3. PTC heterozygotes have mild haploabnormalities. (A) Vermis width in heterozygotes for individual protein-negative PTC variants relative to control littermates. Y-axis was shifted relative to Figures 1-2 to emphasize distribution within the range of observed values. Bottom right corner, mean and p-value for all PTC heterozygotes, combined N=165, one-sample t-test for true ratio=1. (B) Weight at sacrifice, combined N=163. (C) Western blots showed reduced steady-state level of Zfp423 protein in neonatal cerebellum of PTC heterozygotes relative to reference littermate. (D) Normalized values among PTC heterozygotes with independent antibodies showed typical values between 0.5 and 1 relative to control littermates in neonatal cerebellum. (E-H) Behavioral tests for locomotor coordination on littermate pairs for two PTC heterozygotes, H503Qfs*11 (N=18) and N1056Qfs*29 (N=24), either coisogenic on FVB (filled circles) or as B6xFVB F1 hybrids (open circles). (E) Hanging wire task

performance scores and fall latencies. Wilcoxon signed rank test, N=43. (F) Elevated beam escape task. (G) Rotating rod fall latencies across six successive trials per pair and average ratio across all six trials. P-value is one-sample t-test for average difference being 0. (H) Footprint analysis for stride length in each paw and width between left and right paws. FL front left, FR front right, BL back left, BR back right, FW front paw stride width, BW back paw stride width. P-values for one-sample t-tests for average differences equal to 0.

Infrared fluorescence Western Blots from neonatal cerebellum (Figure 3C and S2_Fig A) supported average expression ratios between 0.56-0.73 with two different antibodies (N=29 comparisons for A304, 26 for ABN410), albeit with substantial variation across experiments (Figure 3D). Expression ratio greater than 0.5 is consistent with a proposed negative autoregulatory activity of Zfp423 [37]; it may also reflect changes in tissue composition of cell types and states, although the strongest source perinatally should be granule cell precursors, which decrease in mutant animals [10, 20]. With these caveats, identification of mild phenotypes in null heterozygotes places limits on the degree of protein functional deficit required to produce phenotypes relevant to disease modeling for structural abnormalities.

Quantified measures of locomotor function, performed by a core facility blind to genotype and experimental goals, supported a modest difference only in stride length for PTC heterozygotes. Joint analysis of two heterozygous variants, H503Tfs*11 (N=18 littermate pairs) and N1056Qfs*29 (N=24 littermate pairs), did not identify significant differences by genotype on a hanging wire task for grip strength and coordination (Figure 3E), elevated beam escape (Figure 3F), nor accelerating rotating rod (Figure 3G) tasks for locomotor coordination. Footprint analysis for gait parameters showed ~5% reduced stride length

that was independently significant for each variant, although potentially confounded by animal size as assessed by weight, with no detected difference in front or back stride width (Figure 3H). These data supported better sensitivity and cost effectiveness of anatomical measures than simple behavioral measures to perturbations in *Zfp423* function in mice.

Differential effects near either end of the open reading frame.

Despite being created on the more sensitive B6 background, D70Vfs*6 showed milder phenotypes than all other PTC variants, while a 91-bp deletion that removed just the nine terminal residues encoded by exon 8 on the FVB background appeared similar to null variants. Western blots showed that D70Vfs*6 produced a detectable pool of lower-molecular weight Zfp423 protein (Figure 4A and S2_Fig B), presumably by translational initiation after the introduced stop codon at position 72 in its first open reading frame. The next available in-frame AUG is at position 118, between ZF1 and ZF2, and would include each of the previously annotated binding domains. The exon 8 deletion produced a protein of near-normal size, but one that must lack both histidine residues from ZF30, which is required for binding EBF proteins [26, 38] and at substantially reduced steady-state level (Figure 4B and S2_Fig C). D70Vfs*6 Homozygotes have mild to moderate ataxia and hypomorphic anatomical features, most notably vermis hypoplasia (Figure 4C, E). In contrast loss of exon 8 had severe anatomical abnormalities, including nearly complete loss of vermis and corpus callosum in all six animals assessed (Figure 4D, E). Each of these mutations had significantly reduced frequency of homozygote offspring from heterozygote crosses (Figure 4F). These two mutations showed somewhat unexpected features of *Zfp423* functional organization and reinforced the need for empirical testing of predicted variant effects.

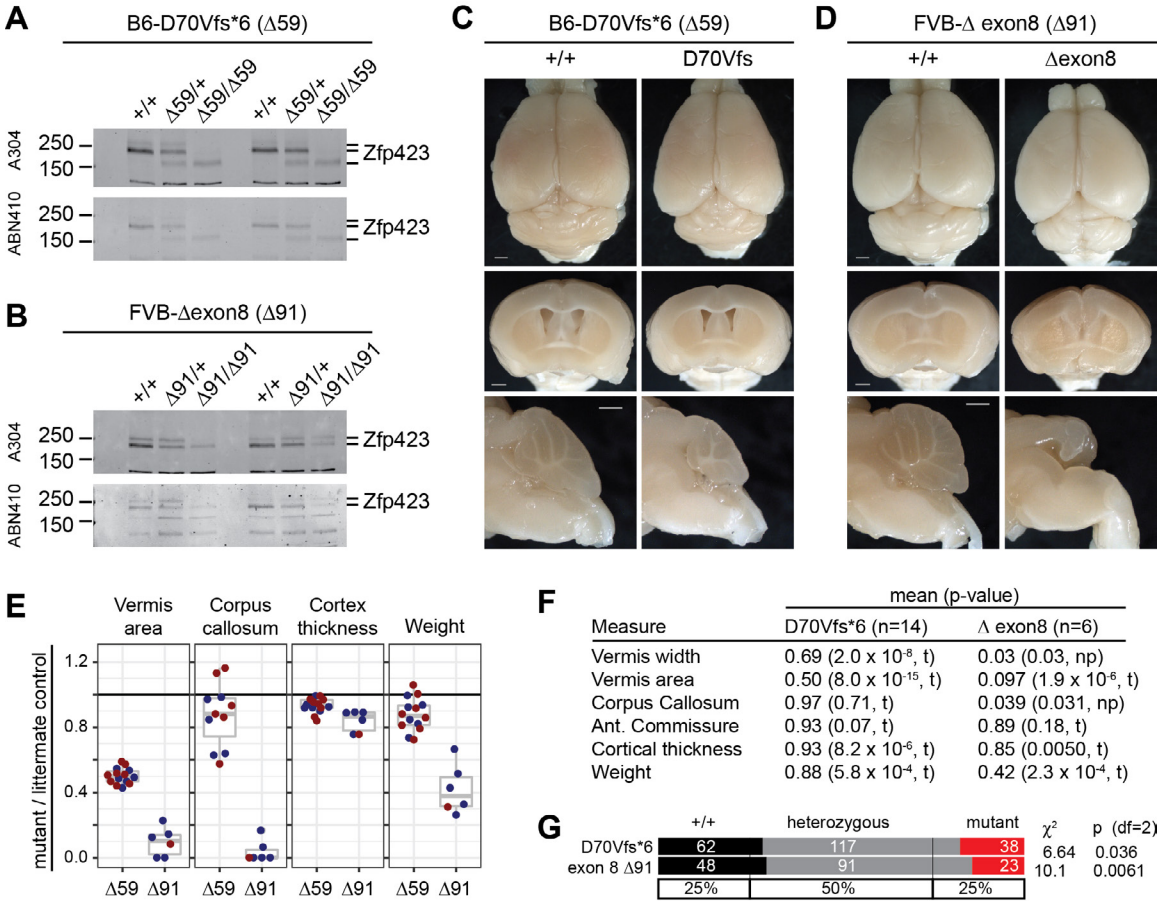


Figure 4. D70Vfs*6 is hypomorphic while deleting exon 8 is approximately null. (A) Westerns blots from neonatal cerebellum from independent trios with two different antibodies against Zfp423 showed a reduced amount of lower molecular weight protein derived from the D70Vfs*6 mutation. (B) Similar blots from two trios for deletion of exon 8, encoding the final nine amino acids of ZF30, showed reduced level of near-full length protein. (C) D70Vfs*6 anatomical phenotypes included reduced cerebellum size. Scale bars, 1 mm. (D) Exon 8 deletion phenotypes approximate those of null alleles despite persistent protein. Scale bars, 1 mm. (E) Quantification of anatomical measures from D70Vfs*6 (Δ59) and exon 8 deletion (Δ91) homozygote and control littermate pairs. (F) P-values from one-sample t-test (t) or non-parametric Wilcoxon Signed Rank test (np) for anatomical measures. (G) Both D70Vfs and exon 8 deletion showed reduced frequencies of homozygotes in crosses.

H1277Y is pathogenic in mice, three other asserted mutations and six VUS are not. We assessed 10 patient-derived substitution alleles: four asserted pathogenic in ClinVar and six rare variants of uncertain significance spanning a range of allele frequencies in public databases from zero to 7.6 x 10^{-4} (Table 1 and Figure 1). Each position is highly conserved across vertebrates, except R89H. Eight lie in or adjacent to C2H2 zinc fingers (Figure 1B) or a putative C4 zinc finger [9]. Among these ten, only H1277Y showed a severe disease-related abnormality while none of the others was distinguishable from control littermates in our assays.

H1277Y was identified in a patient with cerebellar vermis hypoplasia, nephronophthisis, and perinatal breathing abnormalities [7]. By replacing a zinc-coordinating histidine with tyrosine, H1277Y should disrupt the structure of ZF30, which is required for interaction with EBF family proteins [26, 38]. Mice homozygous for the H1277Y variant showed many features of null alleles, including gross ataxia (Supplementary videos), vermis agenesis, reduced cortical thickness, and incomplete corpus callosum (Figure 5A) while a silent control edit was indistinguishable from control littermates (Figure 5B). Zfp423 H1277Y protein has substantially reduced abundance in neonatal cerebellum, similar to the exon 8 deletion and consistent with structural destabilization of the terminal zinc finger (Figure 5C and S2_Fig D). Quantitative measures from multiple same-sex littermate pairs show full penetrance of severe defects in surviving mutants of both sexes for cerebellar vermis hypoplasia, loss of corpus callosum at the midline, cortical thickness, and body weight (Figure 5D). Retrospective analysis of breeding records showed reduced frequency of homozygotes for H1277Y, but not the silent substitution control allele (Figure 5E). These results confirm the pathogenic nature of H1277Y for structural brain abnormalities.

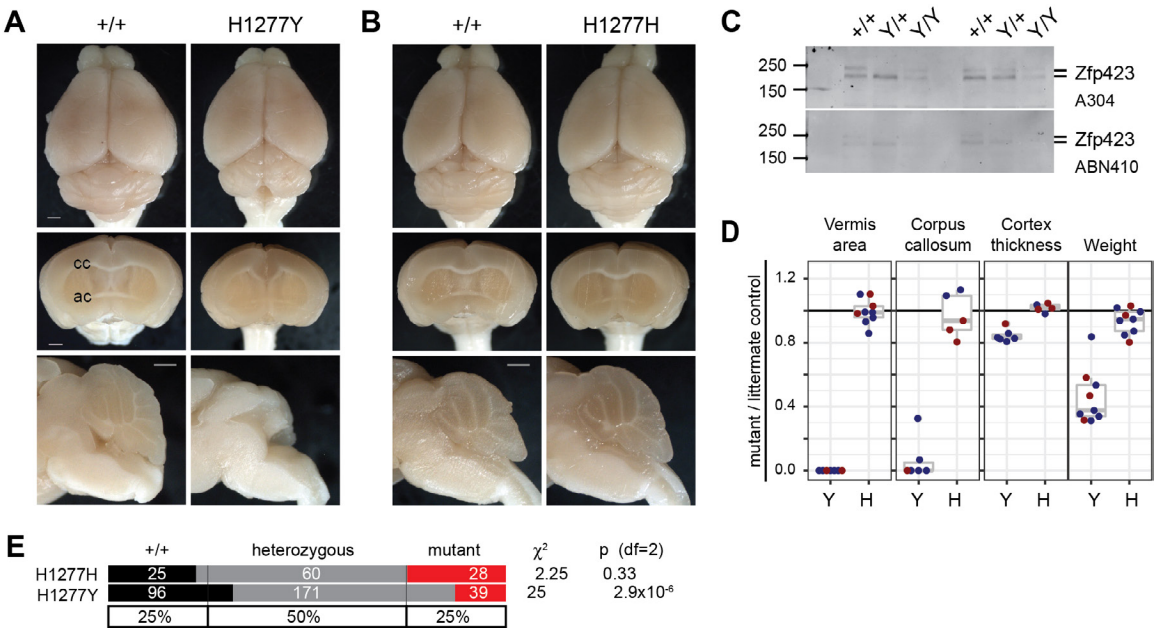


Figure 5. H1277Y is pathogenic and has reduced protein abundance. (A) Surface and block face views of same-sex littermates show H1277Y midline defects similar to null alleles, including complete loss of vermis, loss of corpus callosum crossing, and reduced cortical thickness. Scale bars, 1 mm. (B) Silent substitution H1277H is indistinguishable from control littermates. (C) Western blots with antibodies to amino-terminal (top) or carboxy-terminal (bottom) domains show dramatic loss of Zfp423 protein abundance. Results from two distinct trios shown. (D) Ratios relative to control same-sex littermates quantify loss of vermis, midline corpus callosum, cortex thickness, and body weight for H1277Y, but not H1277H. (E) Breeding records show reduced frequency of H1277Y, but not H1277H, homozygotes from heterozygote crosses.

The other three asserted pathogenic variants were not sufficient to induce JSRD-like or other obvious phenotypes, nor were any of the tested VUS alleles. P913L was identified by homozygosity in a consanguineous patient with cerebellar vermis hypoplasia, nephronophthisis, and situs inversus [7]. R89H and E1124K were found together in a

patient with macrocephaly, extended subarachnoid spaces, and thin corpus callosum [8]. Each of these variants was absent in contemporaneous control subjects. Each was later found in new and larger public databases, but only at low allele frequencies (1.8×10^{-4} to 3.2×10^{-5}, Table 1) and only as heterozygotes. Using a minimum 10 replicate sample pairs in mice, we did not identify any defect in vermis nor in midline forebrain phenotypes for P913L (S3_Fig), nor for R89H or E1124K homozygous mice, nor for R89H/E1124K compound heterozygous mice (S4_Fig). Substitution variants heterozygous to a null allele did not show any effect in smaller sample sizes. For P913L, we created the same mutation on the more sensitive B6 background, again with no evidence for an effect on brain structures typically affected in *Zfp423* mutants nor reduction in protein level (S1_Fig C, S2_Fig E, S3_Fig). Pooling all substitution variants except H1277Y and its silent control H1277H to test for a generalized substitution effect at high power (N=179 littermate pairs) did not produce statistical support for abnormality in any measure before correction for multiple tests ($p>0.15$ all tests). These data confirmed the pathogenic nature of H1277Y, but supported a more benign interpretation of all other non-synonymous substitutions tested.

Deletions that remove SMAD-binding fingers or a potential C4-ZF domain produce intermediate alleles.

We examined several in-frame deletions (Figure 1A) for Zfp423 protein abundance and brain phenotypes. Recovered examples included distinct Zfp423 functional domains (Figure 6A), several of which had notable reductions in cerebellar vermis (Figure 6B). Several variants with strong phenotypes also significantly affected protein level (Figure 6C, S1_Fig C, and S5_Fig) and we could therefore not distinguish between functional requirement for a domain and destabilization of the protein due to awkward breakpoints,

although comparison to null heterozygotes suggested that strong phenotypes associated with >75% of control expression levels probably indicates a sequence-specific function.

Cerebellar vermis was the most sensitive anatomic measure to in-frame deletions as a class (Figures 1C, 6D). Impact of intermediate alleles on corpus callosum measures had high variance that limited power after correction for multiple tests (Figure 6E) and lacked effect magnitude for simple measures of cortical thickness (Figure 6F). As a group, in-frame deletions slightly decreased body weight (Figure 6G).

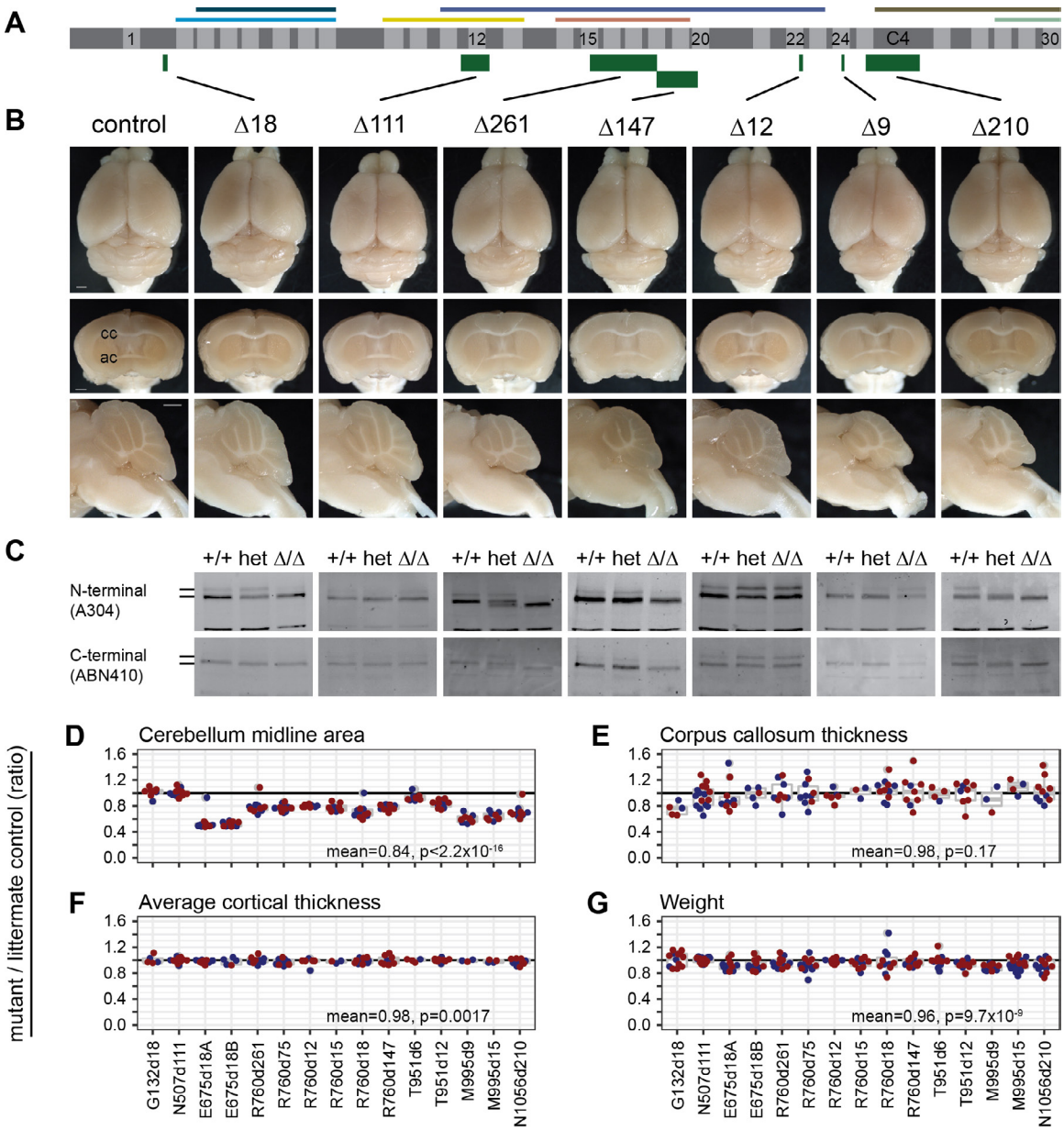


Figure 6. In-frame deletion that remove critical regions or reduce protein abundance are hypomorphic. (A) Schematic as in Figure 1 shows the locations relative to 30 C2H2 zinc fingers and known binding regions of in-frame deletions with example data below. (B) Surface views, coronal, and sagittal block face preparations of brains from typical control and in-frame deletions G132Δ18, N507Δ111, R760Δ261, R760Δ147, T951Δ12, M995Δ9, and N1056Δ210. Scale bars, 1 mm. (C) Western blots show Zfp423 proteins in newborn cerebellum for control, heterozygote, and mutant

littermates for each variant. (D) Cerebellum midline area, (E) Corpus callosum midline thickness, (F) Cortex thickness as the average of three points at 15°, 30°, and 45° from midline, and (G) Weight at sacrifice each expressed as the ratio of mutant to wild-type same-sex littermate controls for 15 in-frame deletion mutations.

Mutations that substantially reduced protein level also had the strongest effects on vermis measures. Overlapping deletions in ZF15 (E675Δ18A, E675Δ18B) and ZF24 (M995Δ9, M995Δ15), and a single small deletion in ZF18 (R760Δ18), each of which is predicted to destabilize the C2H2 structure by removing critical residues, all reduced protein expression level and reduced cerebellar vermis midline area by approximately half.

Four mutations removed significant protein-coding sequences without reducing measured protein level: G132Δ18 p.del(Leu125-Glu130), which removes 6 amino acids in the sequence between ZF1 and ZF2; N507Δ111 p.del(Arg500-Ile536), which removes ZF12 in the BRE-binding region; R760Δ261 p.del(E675-K761), which fuses ZF15 to ZF18 while deleting ZF16-ZF17 in the SMAD-binding region (R760Δ147 and Δ75 also delete fingers within the SMAD region); and N1056Δ210 p.del(Thr1032-Gly1102)>Arg, which removes part of ZF25 and all of the C4 ZF-like sequence. Surprisingly, deletion of ZF12 (N507Δ111) in the annotated BRE-binding region had no measurable effect. The small deletion between ZF1 and ZF2 (G132Δ18) had a nominal effect (mean = 0.97, $p = 0.035$, one-sample t-test, N = 12). However, deletions in the SMAD-binding region that did not disrupt C2H2 structural elements and retained near-normal protein levels (R760Δ261, R760Δ147, and R760Δ75) each showed a strong intermediate phenotype, consistent with independent SMAD-domain deletions reported by Casoni et al. [27]. Deletion of ZF25 and putative C4 ZF (N1056Δ210) showed a similar degree of vermis

hypoplasia, providing the first evidence for organic function of these domains. These results showed that *Zfp423* brain structural phenotypes were sensitive to most in-frame deletions, often including reduced protein levels, and that different ZF domains or clusters had different degrees of sensitivity. That in-frame deletions were hypomorphic rather than effectively null reinforced the idea that Zfp423 coordinates activities among its interaction partners rather than being an essential component of one pathway.

ZF1 deletion has a reproducibly mild phenotype.

The D70Vfs*6 long open reading frame lacks N-terminal residues including the first zinc finger. To determine how much of this phenotype is attributable to loss of ZF1 rather than other N-terminal residues or reduced Zfp423 protein level, we examined three distinct deletions that remove ZF1. One in-frame deletion recovered in the course of modeling patient substitution variants on the FVB background removed all of exon 3 (R89 d345, Figure 1A, C), including ZF1 and ~40 other residues. To test the requirement for ZF1 specifically we constructed two smaller deletions (Δ57, Δ63) targeting just the ZF domain in the more sensitive B6 background (Figure 7A). Both targeted alleles expressed Zfp423 protein at levels similar to control littermates (Figure 7B, C) and differed only slightly in deletion breakpoints relative to the coding sequence (Figure 7D). All three mutations resulted in slightly smaller cerebellar vermis (Figure 7A, E), but less severe than D70Vfs*6 and comparable to reductions seen in null allele heterozygotes. This showed that while ZF1 contributed to Zfp423 function in hindbrain development, it had a smaller effect size than features required for protein production and stability or SMAD binding.

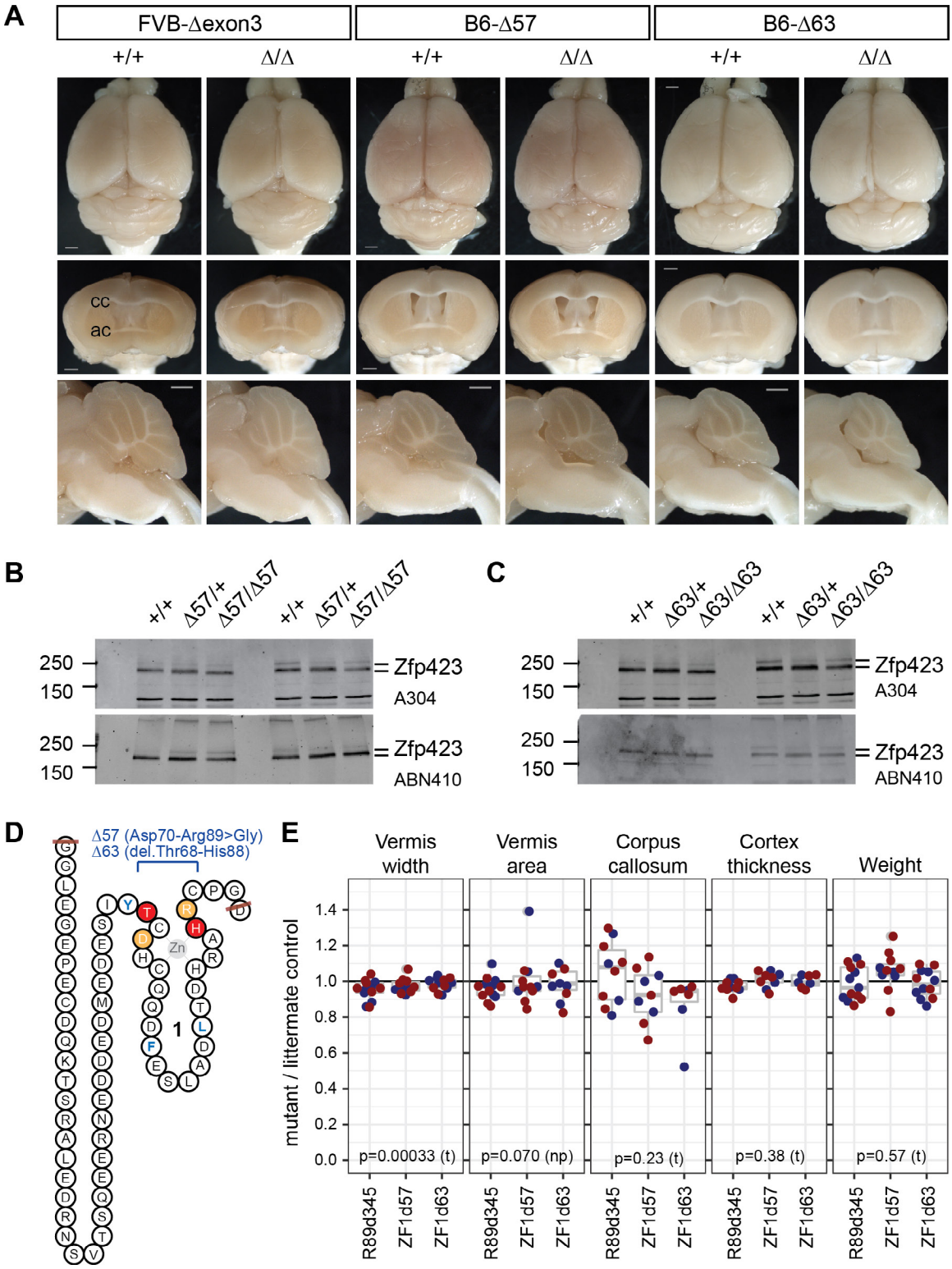


Figure 7. ZF1 deletions had modest, but measureable impact. (A) Surface views and block face views from control and mutant same-sex littermate pairs for in-frame deletions

removing 345 bp (Δexon3) or smaller fragments within exon 3 (Δ57, Δ63) on FVB or B6 coisogenic backgrounds. Scale bars, 1 mm. (B, C) Western blots show similar levels of Zfp423 protein in control, heterozygote, and mutant newborn cerebellum for Δ57 and Δ63 mutations. (D) Peptide sequence encoded by exon 3, with zinc finger 1 schematized and deletion boundaries indicated in light orange (Δ57) or red (Δ63). Structural hydrophobic residues are blue as in Figure 1B. (E) Ratios between mutant and same-sex littermate controls for vermis width (redrawn from Figure 1C), vermis midline area, corpus callosum midline thickness, average cortical thickness, and weight at sacrifice are shown. P-values are shown for one-sample t-test (t) or non-parametric Wilcoxon signed rank test (np).

DISCUSSION

Genomic medicine for rare disorders is often limited by the ability to interpret rare variants. Databases such as ClinVar catalog clinical variants from multiple sources, but report pathogenic assertions based on varied, evolving, and sometimes unclear standards of evidence. Algorithmic predictions and cell models have many benefits, including potential to score all possible single variants, but are prone to errors if the input-output relationships of the assessment do not scale with the impact protein function on the relevant organ system. Attempts to validate effect predictions using clinical variant databases may be somewhat circular if the clinical variants were classified in part on the same criteria as the classifier, such as evolutionary constraint or physic-chemical properties of the substituted residue. Here we showed that current variant effect predictors failed to predict major outcomes for patient *ZNF423* variants accurately. While a minority of variant effect predictors (VEST3, Mutation Assessor) correctly ranked H1277Y as the most likely to be deleterious, neither of these

categorically separated this from more benign variants. In contrast, we showed that simple quantitative phenotypes based on domain expertise of a small team can rapidly assess a large array of variants in whole-animal models in a cost-effective manner.

We confirmed H1277Y as a pathogenic variant, while providing evidence against three other variants previously asserted pathogenic or likely pathogenic and six rare/singleton VUSs (Table 1). The quantification of heterozygote phenotypes, as well as inclusion of collateral variants that remove domains or reduce protein abundance, demonstrated sensitivity of simple assays based on prior genetic analyses to relatively modest genetic perturbation. Simple measures with low variance, such as vermis width, allowed adequate sampling for high statistical power and detection of structural changes less than 5% of mean values in the brains in heterozygous animals. This strengthens the interpretation that substitution alleles with no abnormality are benign with respect to major brain phenotypes. By examining in-frame deletion variants, we confirmed the importance of the SMAD-binding ZF cluster, showed that deletion of ZF25 plus a conserved non-motif segment with potential to form a C4-class zinc finger in a region previously implicated in ZNF423 protein self-association is sufficient to cause intermediate decrease in vermis size, and that ZF1 had detectable, but limited, impact on studied phenotypes despite strong conservation across species. Surprisingly, deletion of ZF12 in the BRE-interacting zinc finger cluster had no detectable effect in any of our outcome measures.

Limitations

While ZNF423-homologous proteins are 98.5% identical in amino acid sequence and previous studies from several laboratories established strong homology between mouse

Zfp423 and human JSRD phenotypes, the degree of sensitivity to mutation could be non-linear. Human and mouse brains, while homologous in structure, develop on different physical and temporal scales. Amino acid substitutions that alter a binding surface could potentially have differential effects depending on conservation of specific binding partners, although this seems unlikely given strong homology of known partners. In order to test many variants rapidly at high power, we focused on simple and less-expensive measures. It is possible that more intensive studies on any specific variant might identify a phenotype, but these are unlikely to be severe in the context of laboratory mice. We also have not examined impacts on other systems, such as olfactory epithelium and adipose tissue, where *Zfp423* null mutants have known phenotypes. While acknowledging these caveats in principle, we nonetheless found high sensitivity of the mouse brain models to even modest genetic perturbation, including heterozygosity for loss-of-function alleles.

Prospects

Well-powered results from experimental models should inform and modify clinical interpretation of rare alleles. Previous studies identified several variants studied here as pathogenic or likely pathogenic in subjects with JSRD or other neurodevelopmental abnormalities based on being rare variants in a gene with known phenotypic overlap and other properties typically associated with causal variants. However, new evidence should update our expectations and interpretations. For example, P913L was proposed as causal based on its being homozygous in an affected child from a consanguineous pairing, not detected in a control population, and conserved across available vertebrate sequences. Subsequent studies, however, found this allele at modest frequency in the general population, which should reduce confidence in a pathogenic role. In our models,

we see no evidence, even on the most sensitive genetic background, for pathogenic consequences of this allele, which should further reduce confidence in a pathogenic interpretation. This variant should now be regarded as likely benign. Similarly, R89H (which is not highly conserved) and E1124K (which is) were reported as likely pathogenic based on being rare or novel in humans and found together in a rare patient. Neither allele when homozygous, nor the two together in trans, nor E1124K in trans to a null allele showed any significant effect on anatomical measures. Interpretation of these variants should also now be modified to likely benign. In contrast, H1277Y, which was strongly predicted to destroy a critical structural residue in the final C2H2 domain, both reduced protein abundance and showed essentially null phenotypes. Interpretation of this variant should now be updated to experimentally supported pathogenic.

More generally, the ability of a single research team to harness domain-specific knowledge to test multiple variants in parallel should improve model-based evidence for causal variants. Our results reinforce the idea that evolutionary constraint is sensitive to much smaller effect sizes than Mendelian disorders and predictions based primarily on constraint are likely prone to false positive calls. Indeed, seven of ten missense positions studied here were otherwise invariant across 165 diverse vertebrates [9]. Pathogenic variant H1277Y was unique among these in removing a structural requirement for its C2H2 domain, which also reduced protein abundance. A similar logic may limit the predictive value of bulk replacement data from mutational scanning experiments. Incorporating domain-level constraints, like zinc-coordinating residues in zinc fingers, might improve the specificity of useful algorithms. While animal models do not scale sufficiently to test all possible variants, we showed here that mouse models can scale adequately to test plausible variants identified in patients for a rare disorder in order to

refine molecular diagnoses. In addition, features that distinguish model-pathogenic from nonpathogenic variants may add to predictive power for untested variants.

MATERIALS AND METHODS

Genome editing. All editing experiments used CRISPR/Cas9 ribonucleoproteins (RNPs) based on *S. pyogenes* Cas9. Guide sequences were selected for limited off-target potential using public, web-based tools [39, 40]. Modified (AltR) crRNA guides and tracrRNA were purchased from IDT. Standard and Hi-Fi variant Cas9 proteins were purchased from IDT and New England Biolabs. Single-stranded oligonucleotide donors for homology-dependent repair (Ultramers and Megamers) were purchased from IDT. Injections of FVB/NJ and C57BL/6N one-cell stage embryos were performed in the Rebecca and John Moores UCSD Cancer Center Transgenic Mouse Shared Resource. Guide sequences, predicted scores, RNP assembly conditions, and editing results are given in S1_Table and S2_Table.

Mutation discovery and validation. Pups derived from injected embryos were screened for developmentally early mutations by PCR-based Sanger sequencing of 500-bp to 700-bp PCR products from crude tail tip lysis DNA preparations. Screening primer sequences are given in S3_Table. Transmission to F1 offspring was confirmed by allele-specific PCR and/or additional DNA sequencing (S4_Table). For mutations with large effects, predicted off-target sites were sequenced to reduce potential for false-positive effects. All variants were also studied across multiple lines and/or backcross generations to further guard against undetected collateral variants by segregation.

Variant effect prediction. PolyPhen2 [41], SIFT [42], PROVEAN [43], MutationTaster [44], MutationAssessor [45], VEST3 [46], CADD [47], and others were run on VCF files

for studied patient variants using wANNOVAR [48] from its web interface (http://wannovar.wglab.org/). A subset of predictions were re-run through their stand-alone web pages for validation. Envision scores [49] were obtained from the Envision web site (https://envision.gs.washington.edu/shiny/envision_new/). Categorical calls (if any) and scores are listed in Table 1. Scales differ across algorithms; original references should be consulted for interpretation.

Stock maintenance. Mice were maintained by backcross to FVB/NJ or C57BL/6J and by intercrosses to obtain desired genotypes. Mice were maintained in a specific pathogen free (SPF) facility on 12 h light, 12 h dark cycle in high-density racks with HEPA-filtered air and ad libitum access to water and food (LabDiet 5P06).

Western blots. Cerebellums were manually dissected from young litters (P0-P4) and individually frozen prior to genotyping. Samples from littermate pairs and trios with desired genotypes were homogenized in RIPA buffer supplemented with protease inhibitors (Millipore Sigma P8340) using a small glass dounce (20 strokes). Protein extracts were quantified with BCA assays. 52 µg samples run through Laemmli SDS-PAGE gels before transfer to nitrocellulose membranes (Bio-Rad 1620112). A subset of blots were incubated with Ponceau-S to visualize protein transfer and subjected to image analysis to quantify bulk protein per lane as a reference for subsequent measures. Zfp423 protein was detected using an antibody to residues 250-300 (A304-017A, Bethyl Labs) or residues near the carboxylterminal end of the protein (ABN410, Millipore) with IR-700 conjugated goat anti-rabbit secondary antibody (Rockland 611130122) and detection on a LiCor Odyssey fluorescence imaging station. Processed blots were reprocessed with a cocktail of anti-phosphoprotein antibodies (Millipore Sigma, P3430, P3300, and P3555; with an IR-800 conjugated secondary antibody) as a proxy for total

protein. Gel images were quantified in ImageJ and measurements were recorded as background-corrected Zfp423 signal normalized to either Ponceau-S or phosphoproteins.

Anatomical measures. Samples were prepared, photographed, and measured by an investigator blinded to genotypes. For fixed preparations, deeply anesthetized animals were perfused with phosphate buffered saline followed by 4% paraformaldehyde and brains were removed into fresh 4% paraformaldehyde at 4°C for 12-24 h and 15-30% sucrose for 24-48 h. Brains were imaged through a dissecting microscope (Zeiss Stemi 2000-C) with a digital camera (Nikon DS-Fi1) using standardized zoom and distance settings and a standard ruler in frame to verify scale. Paired samples were processed together and imaged consecutively. Anatomical features were measured in ImageJ (v1.52a). For surface images, brains were aligned on a swivel-mount platform and photographed dorsal side up. Vermis width was measured at the middle of the folium-tuber lobule (VII). Cerebellar hemisphere height was measured as a vertical line drawn from the dorsal-most point of the simple lobule. Coronal and sagittal block face preparations were made using a standard mouse brain matrix (Zinc Instrument) with the sample aligned anteriorly. Coronal cuts were made at the rostral end of the optic chiasm, through the striatum. Sagittal cuts were made at the midline. Cut brains were mounted on a rotating platform to hold each surface perpendicular to the lens. For cortical thickness, three lines were drawn using the ImageJ ROI Manager, one each at 15 degrees, 30 degrees, and 45 degrees counter-clockwise from vertical, starting at a point where the line would be perpendicular to the angle of the brain surface, and ending at the dorsal side of the corpus callosum; the average of these three measurements was used for each animal. The thickness of the corpus callosum and anterior commissure were measured with vertical lines at the midline and the width of the brain was measured with a horizontal line across the coronal surface at its widest point. Vermis area was

measured from midline sagittal block face image using the polygon selections tool in ImageJ to manually define the region of interest.

Locomotor assays. Gross locomotor function was assessed in home cages and by allowing each test animal and same-sex littermate control to walk freely across a small stage with video recordings for a minimum of three crossings. Rotating rod, footprint pattern, hanging wire, and beam walking tests were performed on same-sex littermate pairs in the Scripps Research Institute Animal Models Core Facility by staff blinded to genotype and hypothesis and following standard protocols. Animals included roughly equal numbers of male pairs and female pairs for each genotype tested.

Rotating rod test. Latency to fall from an accelerating rotating rod assessed a combination of proprioceptive, vestibular, fine motor, and motor learning capabilities required to avoid falling [50]. Animals were placed on the apparatus (Roto-rod Series 8, IITC Life Sciences, Woodland Hills, CA) prior to acceleration. Latency to fall was recorded by sensing platforms below the rotating rod. Mice were tested in two sets of 3 trials separated by 2 hours.

Footprint pattern test. Footprint pattern analysis assessed basic gait parameters [50-52]. Non-toxic paint was applied to each paw, with front and back paws distinguished by color. Each mouse was placed at one end of a runway covered in paper and allowed to walk until their paws no longer left marks. Forelimb and hindlimb stride lengths (left and right) and front and back leg stride widths were measured the average of three full strides was used for each mouse's values. Mice that did not make 3 measurable strides were excluded.

Hanging wire test. The hanging wire test assessed grip strength and coordination [53, 54]. Mice were held so that only their forelimbs contact an elevated metal bar (2 mm diameter, 45 cm long, 37 cm above the floor) parallel to the ground and released to hang. Each mouse had three trials separated by 30 seconds. Each trial was scored 0 (mouse fell off), 1 (hung onto the wire by two forepaws), 2 (hung onto the wire by two forepaws and attempted to climb onto the wire), 3 (hung onto the wire by two forepaws plus one or both hindpaws), 4 (hung onto the wire by all four paws plus tail wrapped), or 5 (escaped to the ring stand holding the bar or climbed down the stand to the table). Latency to falling off was measured up to a maximum of 30 s.

Elevated beam test. Escape latency and observed foot slips during escape on a narrow beam further assessed locomotor coordination [50]. Three successive trials were recorded. Average escape time and total number of slips were compared between genotypes.

Statistical analyses. Target samples sizes were estimated from literature and refined according to power calculations based on the observed standard deviations for wild-type littermate pairs as an empirical null model. Retrospective analysis in the R package pwr (v1.2-2, https://github.com/heliosdrm/pwr) or stand-alone software G*power (v3.1.9.2, [55]). Hypothesis tests were performed in R (v3.5.1 [56]). A one-sample t-test was performed for same-sex littermate ratios = 1, or the non-parametric Wilcoxon Signed Rank test if the data distribution showed significant departure from normality by the Shapiro-Wilk test. Multiplicity corrections for family-wise error rate and false discovery rate with mutant classes and false discovery rate across all genotypes are given in S5_Table. Graphical output was in R base graphics or ggplot2 (v3.1.0 [57]) with ggbeeswarm (v0.6.0 https://github.com/eclarke/ggbeeswarm).

To estimate statistical power, we first analyzed non-mutant same-sex littermate pairs, for which data accumulated more quickly than for any specific variant. Vermis width and cortical thickness measures both showed paired sample ratios with low variance and approximately normal distribution. Power calculation for a one-sample t-test estimated 90% power to detect a 10% difference with 5-6 paired samples for a nominal alpha=0.05 or 10-12 samples after Bonferroni correction for ~50 variants tested for each measure. Other measured values had larger variance (and in some cases incomplete penetrance) and required transformation to meet normality. Vermis width was also the most sensitive measure for variants with any effect (see below). Based on these observations, we set vermis as the primary outcome and 10 sex-matched littermate pairs as a target minimum sample size for testing quantitative effects of *Zfp423* variants.

Ethics statement. All animal experiments were approved by the University of California San Diego Institutional Animal Care and Use Committee (UCSD-IACUC) under protocol S00291.

ACKOWLEDGEMENTS

We thank Professors Joseph Gleeson and Friedhelm Hildebrandt for communicating anonymized patient variants and Dr. Kevin Ross for comments on a draft manuscript and Professor Amit Majithia for helpful conversations. We thank families who shared their variants in MyGene2 and ClinVar. We thank Ella Kothari and Jun Zhao for assistance with mouse zygote injections. We thank Dr. Amanda Roberts for advice and coordination of behavioral tests. We thank Abraham Ramirez for excellent animal care.

SUPPORT

This work was supported by grant R01 NS097534 from the National Institute of Neurological Disorders and Stroke.

AUTHOR CONTRIBUTIONS

BAH designed and performed experiments, analyzed data, and wrote the manuscript. OD, RZ, RL, and DC refined methods, performed experiments, and analyzed data. All authors reviewed and agreed to the manuscript.

COMPETING INTERESTS STATEMENT

The authors declare no competing interests.

REFERENCES

1. Kim YE, Ki CS, Jang MA. Challenges and Considerations in Sequence Variant Interpretation for Mendelian Disorders. Ann Lab Med. 2019;39(5):421-9. doi: 10.3343/alm.2019.39.5.421. PubMed PMID: 31037860; PubMed Central PMCID: PMCPMC6502951.

2. Rivera-Munoz EA, Milko LV, Harrison SM, Azzariti DR, Kurtz CL, Lee K, et al. ClinGen Variant Curation Expert Panel experiences and standardized processes for disease and gene-level specification of the ACMG/AMP guidelines for sequence variant interpretation. Hum Mutat. 2018;39(11):1614-22. doi: 10.1002/humu.23645. PubMed PMID: 30311389; PubMed Central PMCID: PMCPMC6225902.

3. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, et al. Analysis of protein-coding genetic variation in 60,706 humans. Nature. 2016;536(7616):285-91. doi: 10.1038/nature19057. PubMed PMID: 27535533; PubMed Central PMCID: PMCPMC5018207.

4. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. Variation across 141,456 human exomes and genomes reveals the spectrum of loss-of-function intolerance across human protein-coding genes. bioRxiv. 2019:531210. doi: 10.1101/531210.

5. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. Genome-wide genetic data on ~500,000 UK Biobank participants. bioRxiv. 2017:166298. doi: 10.1101/166298.

6. Walsh R, Thomson KL, Ware JS, Funke BH, Woodley J, McGuire KJ, et al. Reassessment of Mendelian gene pathogenicity using 7,855 cardiomyopathy cases and

60,706 reference samples. Genet Med. 2017;19(2):192-203. doi: 10.1038/gim.2016.90. PubMed PMID: 27532257; PubMed Central PMCID: PMCPMC5116235.

7. Chaki M, Airik R, Ghosh AK, Giles RH, Chen R, Slaats GG, et al. Exome capture reveals ZNF423 and CEP164 mutations, linking renal ciliopathies to DNA damage response signaling. Cell. 2012;150(3):533-48. doi: 10.1016/j.cell.2012.06.028. PubMed PMID: 22863007; PubMed Central PMCID: PMCPMC3433835.

8. Karaca E, Harel T, Pehlivan D, Jhangiani SN, Gambin T, Coban Akdemir Z, et al. Genes that Affect Brain Structure and Function Identified by Rare Variant Analyses of Mendelian Neurologic Disease. Neuron. 2015;88(3):499-513. doi: 10.1016/j.neuron.2015.09.048. PubMed PMID: 26539891; PubMed Central PMCID: PMCPMC4824012.

9. Hamilton BA. ZNF423 orthologs are highly constrained in vertebrates but show domain-level plasticity across invertebrate lineages. bioRxiv. 2020:2020.03.09.984518. doi: 10.1101/2020.03.09.984518.

10. Alcaraz WA, Gold DA, Raponi E, Gent PM, Concepcion D, Hamilton BA. Zfp423 controls proliferation and differentiation of neural precursors in cerebellar vermis formation. Proc Natl Acad Sci U S A. 2006;103(51):19424-9. doi: 10.1073/pnas.0609184103. PubMed PMID: 17151198; PubMed Central PMCID: PMCPMC1748242.

11. Cheng LE, Zhang J, Reed RR. The transcription factor Zfp423/OAZ is required for cerebellar development and CNS midline patterning. Dev Biol. 2007;307(1):43-52. doi: 10.1016/j.ydbio.2007.04.005. PubMed PMID: 17524391; PubMed Central PMCID: PMCPMC2866529.

12. Warming S, Rachel RA, Jenkins NA, Copeland NG. Zfp423 is required for normal cerebellar development. Mol Cell Biol. 2006;26(18):6913-22. doi: 10.1128/MCB.02255-05. PubMed PMID: 16943432; PubMed Central PMCID: PMCPMC1592861.

13. Casoni F, Croci L, Vincenti F, Podini P, Massimino L, Cremona O, et al. ZFP423 regulates early patterning and multiciliogenesis in the hindbrain choroid plexus. bioRxiv. 2020:2020.03.04.975573. doi: 10.1101/2020.03.04.975573.

14. Cheng LE, Reed RR. Zfp423/OAZ participates in a developmental switch during olfactory neurogenesis. Neuron. 2007;54(4):547-57. doi: 10.1016/j.neuron.2007.04.029. PubMed PMID: 17521568; PubMed Central PMCID: PMCPMC2866517.

15. Massimino L, Flores-Garcia L, Di Stefano B, Colasante G, Icoresi-Mazzeo C, Zaghi M, et al. TBR2 antagonizes retinoic acid dependent neuronal differentiation by repressing Zfp423 during corticogenesis. Dev Biol. 2018;434(2):231-48. doi: 10.1016/j.ydbio.2017.12.020. PubMed PMID: 29305158.

16. Gupta RK, Arany Z, Seale P, Mepani RJ, Ye L, Conroe HM, et al. Transcriptional control of preadipocyte determination by Zfp423. Nature. 2010;464(7288):619-23. doi: 10.1038/nature08816. PubMed PMID: 20200519; PubMed Central PMCID: PMCPMC2845731.

17. Gupta RK, Mepani RJ, Kleiner S, Lo JC, Khandekar MJ, Cohen P, et al. Zfp423 expression identifies committed preadipocytes and localizes to adipose endothelial and perivascular cells. Cell Metab. 2012;15(2):230-9. doi: 10.1016/j.cmet.2012.01.010. PubMed PMID: 22326224; PubMed Central PMCID: PMCPMC3366493.

18. Shao M, Hepler C, Vishvanath L, MacPherson KA, Busbuso NC, Gupta RK. Fetal development of subcutaneous white adipose tissue is dependent on Zfp423. Mol Metab. 2017;6(1):111-24. doi: 10.1016/j.molmet.2016.11.009. PubMed PMID: 28123942; PubMed Central PMCID: PMCPMC5220400.

19. Plikus MV, Guerrero-Juarez CF, Ito M, Li YR, Dedhia PH, Zheng Y, et al. Regeneration of fat cells from myofibroblasts during wound healing. Science. 2017;355(6326):748-52. doi: 10.1126/science.aai8792. PubMed PMID: 28059714; PubMed Central PMCID: PMCPMC5464786.

20. Hong CJ, Hamilton BA. Zfp423 Regulates Sonic Hedgehog Signaling via Primary Cilium Function. PLoS Genet. 2016;12(10):e1006357. doi: 10.1371/journal.pgen.1006357. PubMed PMID: 27727273; PubMed Central PMCID: PMCPMC5065120.

21. Hata A, Seoane J, Lagna G, Montalvo E, Hemmati-Brivanlou A, Massague J. OAZ uses distinct DNA- and protein-binding zinc fingers in separate BMP-Smad and Olf signaling pathways. Cell. 2000;100(2):229-40. doi: 10.1016/s0092-8674(00)81561-5. PubMed PMID: 10660046.

22. Huang S, Laoukili J, Epping MT, Koster J, Holzel M, Westerman BA, et al. ZNF423 is critically required for retinoic acid-induced differentiation and is a marker of neuroblastoma outcome. Cancer Cell. 2009;15(4):328-40. doi: 10.1016/j.ccr.2009.02.023. PubMed PMID: 19345331; PubMed Central PMCID: PMCPMC2693316.

23. Ku M, Howard S, Ni W, Lagna G, Hata A. OAZ regulates bone morphogenetic protein signaling through Smad6 activation. J Biol Chem. 2006;281(8):5277-87. doi: 10.1074/jbc.M510004200. PubMed PMID: 16373339.

24. Masserdotti G, Badaloni A, Green YS, Croci L, Barili V, Bergamini G, et al. ZFP423 coordinates Notch and bone morphogenetic protein signaling, selectively up-regulating Hes5 gene expression. J Biol Chem. 2010;285(40):30814-24. doi: 10.1074/jbc.M110.142869. PubMed PMID: 20547764; PubMed Central PMCID: PMCPMC2945575.

25. Signaroldi E, Laise P, Cristofanon S, Brancaccio A, Reisoli E, Atashpaz S, et al. Polycomb dysregulation in gliomagenesis targets a Zfp423-dependent differentiation network. Nat Commun. 2016;7:10753. doi: 10.1038/ncomms10753. PubMed PMID: 26923714; PubMed Central PMCID: PMCPMC4773478.

26. Tsai RY, Reed RR. Cloning and functional characterization of Roaz, a zinc finger protein that interacts with O/E-1 to regulate gene expression: implications for olfactory neuronal development. J Neurosci. 1997;17(11):4159-69. PubMed PMID: 9151733.

27. Casoni F, Croci L, Bosone C, D'Ambrosio R, Badaloni A, Gaudesi D, et al. Zfp423/ZNF423 regulates cell cycle progression, the mode of cell division and the DNA-damage response in Purkinje neuron progenitors. Development. 2017;144(20):3686-97. doi: 10.1242/dev.155077. PubMed PMID: 28893945; PubMed Central PMCID: PMCPMC5675449.

28. Fowler DM, Araya CL, Fleishman SJ, Kellogg EH, Stephany JJ, Baker D, et al. High-resolution mapping of protein sequence-function relationships. Nat Methods. 2010;7(9):741-6. doi: 10.1038/nmeth.1492. PubMed PMID: 20711194; PubMed Central PMCID: PMCPMC2938879.

29. Fowler DM, Fields S. Deep mutational scanning: a new style of protein science. Nat Methods. 2014;11(8):801-7. doi: 10.1038/nmeth.3027. PubMed PMID: 25075907; PubMed Central PMCID: PMCPMC4410700.

30. Majithia AR, Tsuda B, Agostini M, Gnanapradeepan K, Rice R, Peloso G, et al. Prospective functional classification of all possible missense variants in PPARG. Nat Genet. 2016;48(12):1570-5. doi: 10.1038/ng.3700. PubMed PMID: 27749844; PubMed Central PMCID: PMCPMC5131844.

31. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 2018;46(D1):D1062-D7. doi: 10.1093/nar/gkx1153. PubMed PMID: 29165669; PubMed Central PMCID: PMCPMC5753237.

32. Chong JX, Yu JH, Lorentzen P, Park KM, Jamal SM, Tabor HK, et al. Gene discovery for Mendelian conditions via social networking: de novo variants in KDM1A cause developmental delay and distinctive facial features. Genet Med. 2016;18(8):788-95. doi: 10.1038/gim.2015.161. PubMed PMID: 26656649; PubMed Central PMCID: PMCPMC4902791.

33. Herskowitz I. Functional inactivation of genes by dominant negative mutations. Nature. 1987;329(6136):219-22. doi: 10.1038/329219a0. PubMed PMID: 2442619.

34. Cohen S, Kramarski L, Levi S, Deshe N, Ben David O, Arbely E. Nonsense mutation-dependent reinitiation of translation in mammalian cells. Nucleic Acids Res. 2019;47(12):6330-8. doi: 10.1093/nar/gkz319. PubMed PMID: 31045216; PubMed Central PMCID: PMCPMC6614817.

35. Lindeboom RG, Supek F, Lehner B. The rules and impact of nonsense-mediated mRNA decay in human cancers. Nat Genet. 2016;48(10):1112-8. doi: 10.1038/ng.3664. PubMed PMID: 27618451; PubMed Central PMCID: PMCPMC5045715.

36. Hoek TA, Khuperkar D, Lindeboom RGH, Sonneveld S, Verhagen BMP, Boersma S, et al. Single-Molecule Imaging Uncovers Rules Governing Nonsense-Mediated mRNA Decay. Mol Cell. 2019;75(2):324-39 e11. doi: 10.1016/j.molcel.2019.05.008. PubMed PMID: 31155380; PubMed Central PMCID: PMCPMC6675935.

37. Cho YW, Hong CJ, Hou A, Gent PM, Zhang K, Won KJ, et al. Zfp423 binds autoregulatory sites in P19 cell culture model. PLoS One. 2013;8(6):e66514. doi:

10.1371/journal.pone.0066514. PubMed PMID: 23762491; PubMed Central PMCID: PMCPMC3675209.

38. Tsai RY, Reed RR. Identification of DNA recognition sequences and protein interaction domains of the multiple-Zn-finger protein Roaz. Mol Cell Biol. 1998;18(11):6447-56. doi: 10.1128/mcb.18.11.6447. PubMed PMID: 9774661; PubMed Central PMCID: PMCPMC109231.

39. Haeussler M, Schonig K, Eckert H, Eschstruth A, Mianne J, Renaud JB, et al. Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. Genome Biol. 2016;17(1):148. doi: 10.1186/s13059-016-1012-2. PubMed PMID: 27380939; PubMed Central PMCID: PMCPMC4934014.

40. Labun K, Montague TG, Gagnon JA, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res. 2016;44(W1):W272-6. doi: 10.1093/nar/gkw398. PubMed PMID: 27185894; PubMed Central PMCID: PMCPMC4987937.

41. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7(4):248-9. doi: 10.1038/nmeth0410-248. PubMed PMID: 20354512; PubMed Central PMCID: PMCPMC2855889.

42. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. 2012;40(Web Server issue):W452-7. doi: 10.1093/nar/gks539. PubMed PMID: 22689647; PubMed Central PMCID: PMCPMC3394338.

43. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the functional effect of amino acid substitutions and indels. PLoS One. 2012;7(10):e46688. doi: 10.1371/journal.pone.0046688. PubMed PMID: 23056405; PubMed Central PMCID: PMCPMC3466303.

44. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014;11(4):361-2. doi: 10.1038/nmeth.2890. PubMed PMID: 24681721.

45. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. 2011;39(17):e118. doi: 10.1093/nar/gkr407. PubMed PMID: 21727090; PubMed Central PMCID: PMCPMC3177186.

46. Carter H, Douville C, Stenson PD, Cooper DN, Karchin R. Identifying Mendelian disease genes with the variant effect scoring tool. BMC Genomics. 2013;14 Suppl 3:S3. doi: 10.1186/1471-2164-14-S3-S3. PubMed PMID: 23819870; PubMed Central PMCID: PMCPMC3665549.

47. Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. Nucleic Acids Res. 2019;47(D1):D886-D94. doi: 10.1093/nar/gky1016. PubMed PMID: 30371827; PubMed Central PMCID: PMCPMC6323892.

48. Yang H, Wang K. Genomic variant annotation and prioritization with ANNOVAR and wANNOVAR. Nat Protoc. 2015;10(10):1556-66. doi: 10.1038/nprot.2015.105. PubMed PMID: 26379229; PubMed Central PMCID: PMCPMC4718734.

49. Gray VE, Hause RJ, Luebeck J, Shendure J, Fowler DM. Quantitative Missense Variant Effect Prediction Using Large-Scale Mutagenesis Data. Cell Syst. 2018;6(1):116-24 e3. doi: 10.1016/j.cels.2017.11.003. PubMed PMID: 29226803; PubMed Central PMCID: PMCPMC5799033.

50. Carter RJ, Morton J, Dunnett SB. Motor coordination and balance in rodents. Curr Protoc Neurosci. 2001;Chapter 8:Unit 8 12. doi: 10.1002/0471142301.ns0812s15. PubMed PMID: 18428540.

51. Crawley JN, Paylor R. A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. Horm Behav. 1997;31(3):197-211. doi: 10.1006/hbeh.1997.1382. PubMed PMID: 9213134.

52. Gurney ME, Pu H, Chiu AY, Dal Canto MC, Polchow CY, Alexander DD, et al. Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation. Science. 1994;264(5166):1772-5. doi: 10.1126/science.8209258. PubMed PMID: 8209258.

53. Crawley JN. Behavioral phenotyping of transgenic and knockout mice: experimental design and evaluation of general health, sensory functions, motor abilities, and specific behavioral tests. Brain Res. 1999;835(1):18-26. doi: 10.1016/s0006-8993(98)01258-x. PubMed PMID: 10448192.

54. Freitag S, Schachner M, Morellini F. Behavioral alterations in mice deficient for the extracellular matrix glycoprotein tenascin-R. Behav Brain Res. 2003;145(1-2):189-207. doi: 10.1016/s0166-4328(03)00109-8. PubMed PMID: 14529817.

55. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods. 2009;41(4):1149-60. doi: 10.3758/BRM.41.4.1149. PubMed PMID: 19897823.

56. R Core Team. R: A language and environment for statistical computing.: R Foundation for Statistical Computing; 2017. Available from: https://www.R-project.org/.

57. Wickham H. ggplot2: Elegant Graphics for Data Analysis: Springer-Verlag New York; 2016.

SUPPORTING TABLE CAPTIONS

S1_Table. Embryo injections and recovered mutations. Cas9 editing formulations and frequency of recovered founders are given.

S2_Table. Single-stranded donors used for homology-dependent repair (HDR). Ultramer and Megamer sequences. Essential variants are shown in red, silent substitutions intended to bias repair or improve genotyping in blue.

S3_Table. Screening primers. Cas9 guide sequences, target sites, and PCR primers used for screening and sequencing potential founders.

S4_Table. Genotyping assays and strains deposited to MMRRC. Primers sequences, restriction enzymes if needed, product sizes, and gel conditions for all variants analyzed. MMRRC stock numbers are included for nine strains accepted by the repository.

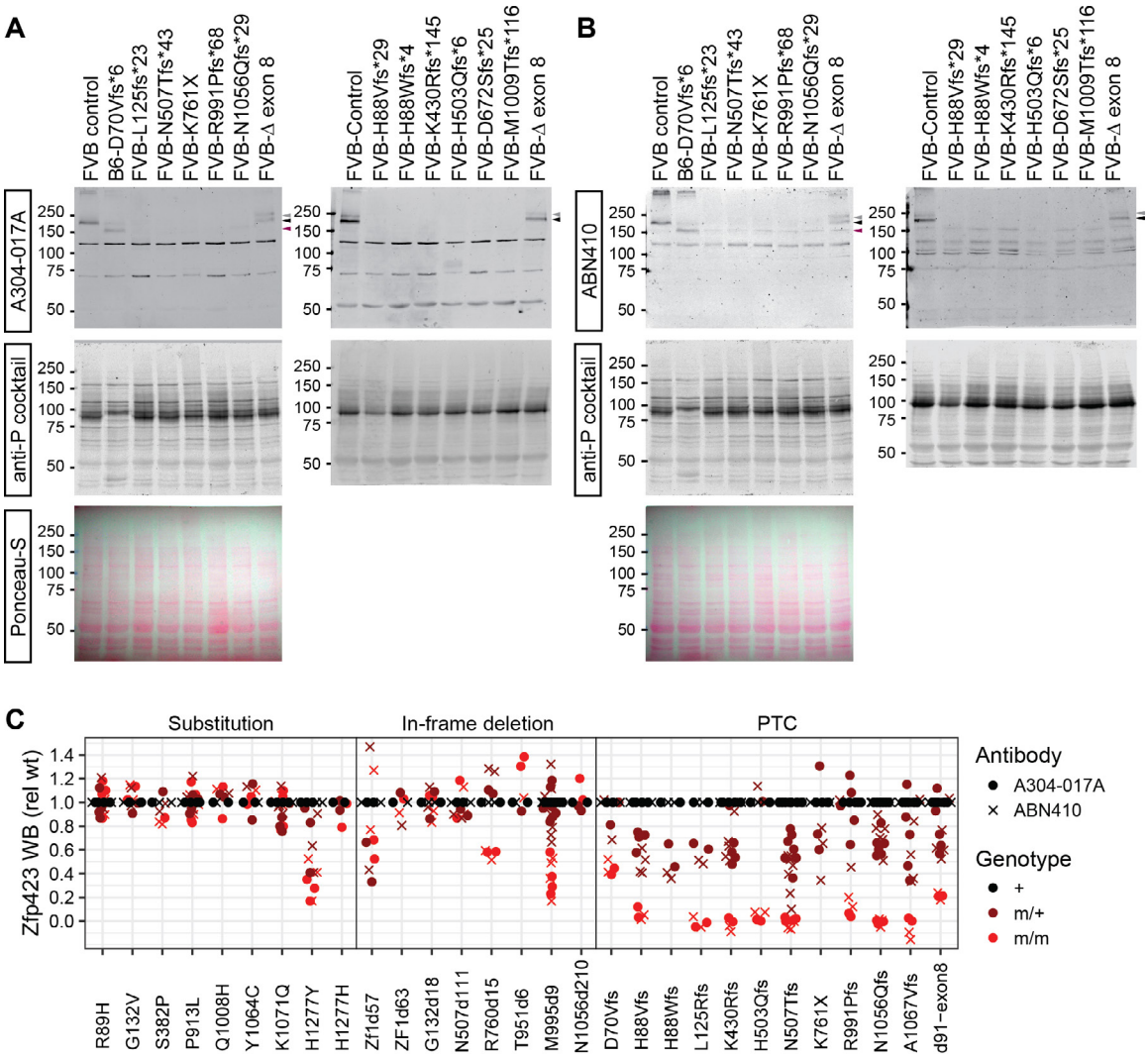
S5_Table. Nominal and corrected p-values for each genotype. One-sample tests of mutant/control ratios for each mutation and seven phenotypes. Family-wise error rate (FWER) and false discovery rate (FDR) corrections are listed within each mutational class (i.e., substitutions, in-frame deletions, premature termination codons) for each phenotype. False discovery rate across all genotypes for a phenotype are also listed.

S6_Table. Physical measures of sex-matched littermate pairs. Primary physical measures for each tested animal, organized by littermate pair.

S7_Table. Locomotor measures sex-matched littermate pairs. Primary behavioral measures for each tested animal, organized by littermate pair.

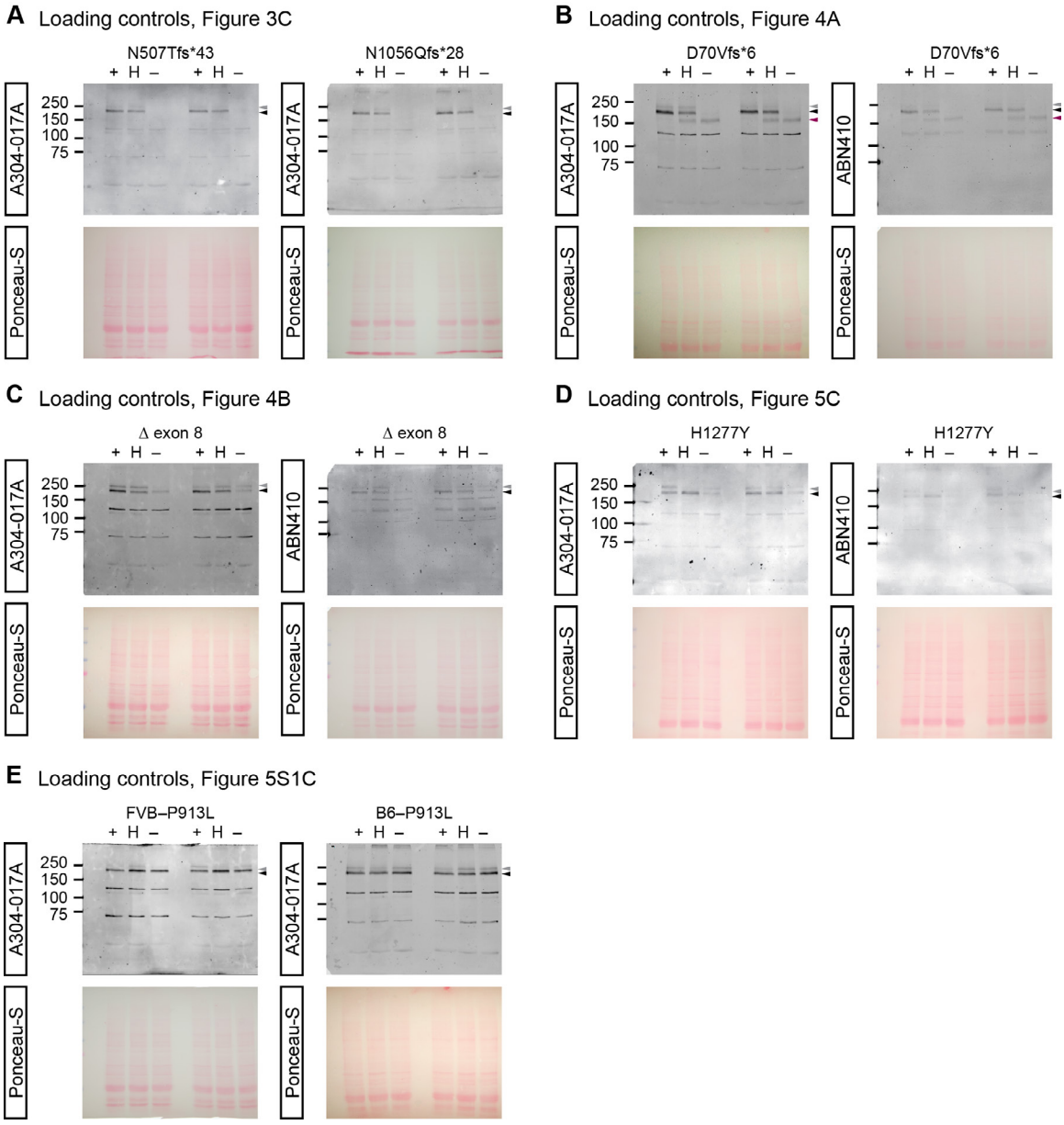
SUPPORTING FIGURES

S1_Fig. Related to Figure 2

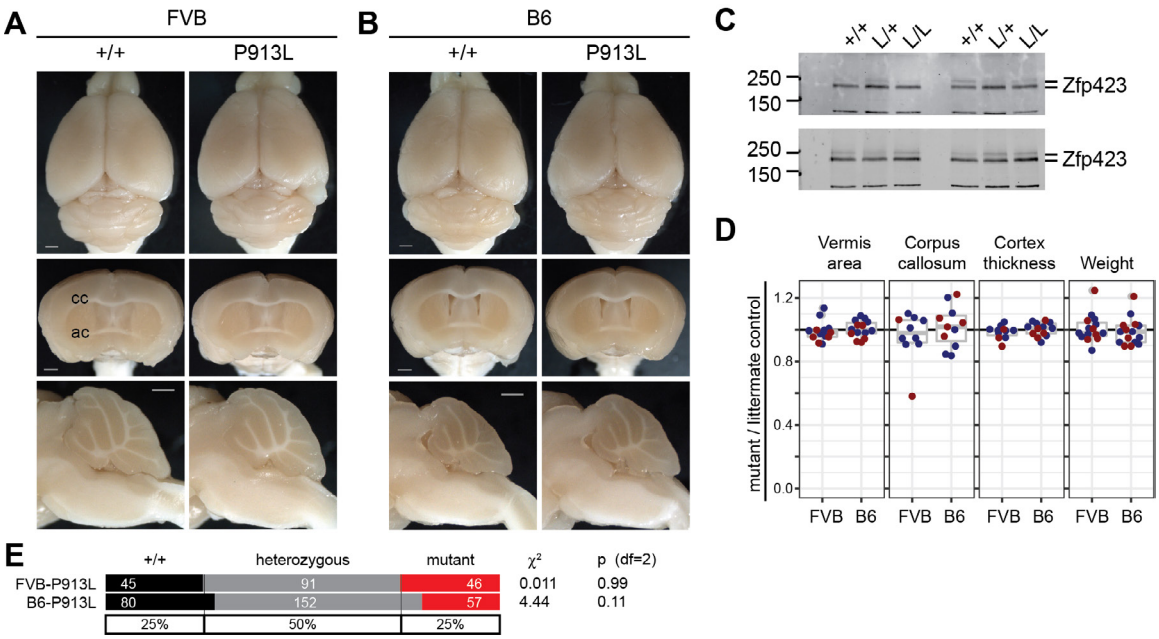


S1_Fig. Related to Figure 2. Western blot loading controls and quantification. (A) Full images for N-terminal blots in Figure 2A. A304-017A antibody detects Zfp423 (black arrowhead) and a variable conformational isomer (gray arrowhead) in control samples as well as a D70Vfs*6-specific protein (purple arrowhead). Molecular weight (kDa) of size marker bands is shown to the left. Membranes re-probed with a cocktail of antibodies againt phosphoserine, phosphothreonine, and phosphotyrosine show approximately even loading. The major band difference in D70Vfs sample in other blots to be a difference between B6 and FVB strain backgrounds. Ponceau-S staining of the

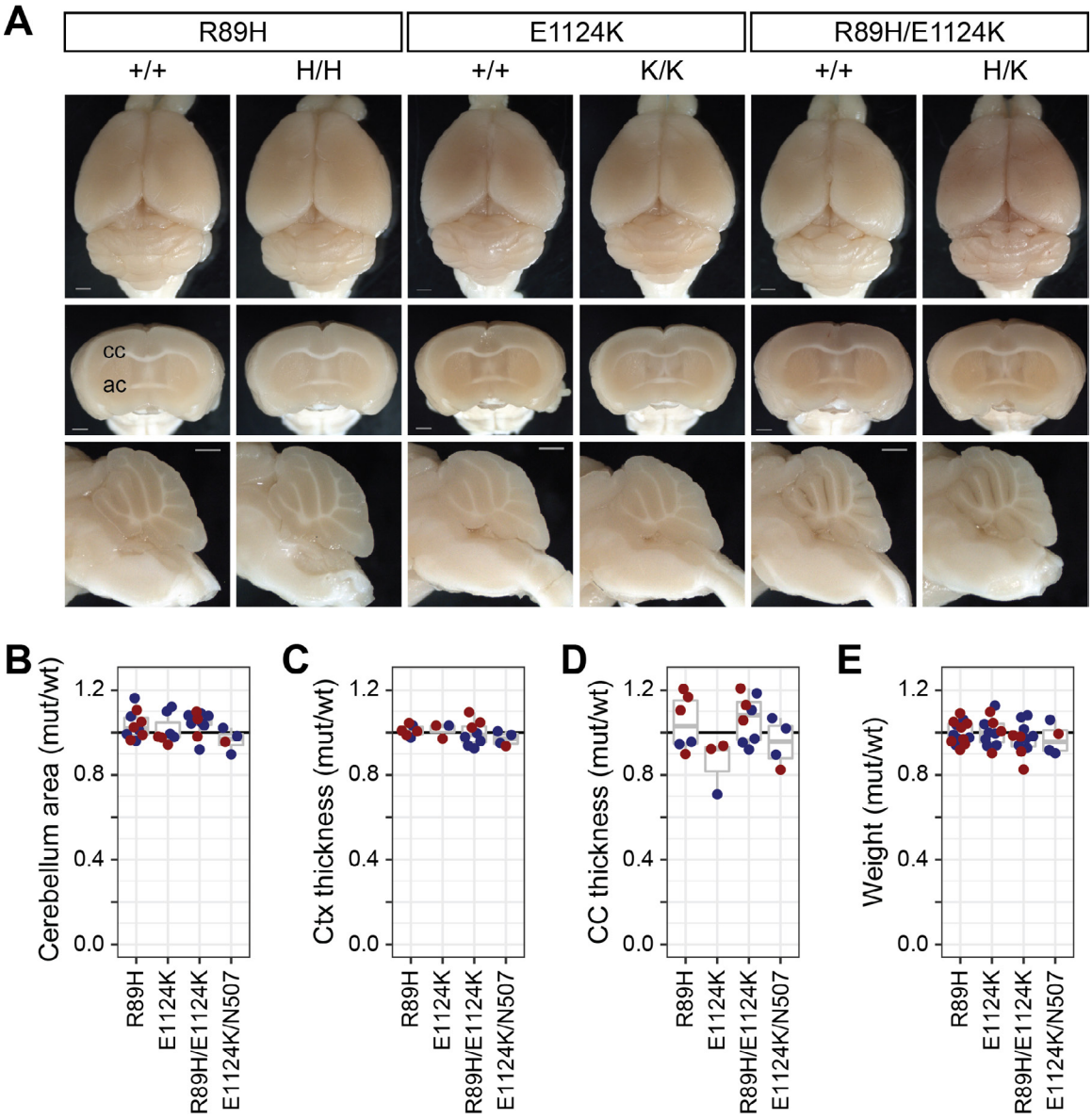
membrane before antibody application also shows approximately equal loading. (B) Western blots and stained membranes used for C-terminal antibody ABN410. (C) Approximate quantification by infrared imaging of all blots in this work. All measures adjusted to loading controls and plotted as ratio to wild-type control sample on the same membrane. Dots are measure from A304-017A, crosses from ABN410. Colors indicate genotypes with non-mutant controls in black, heterozygotes in brown and homozygous mutants in red.



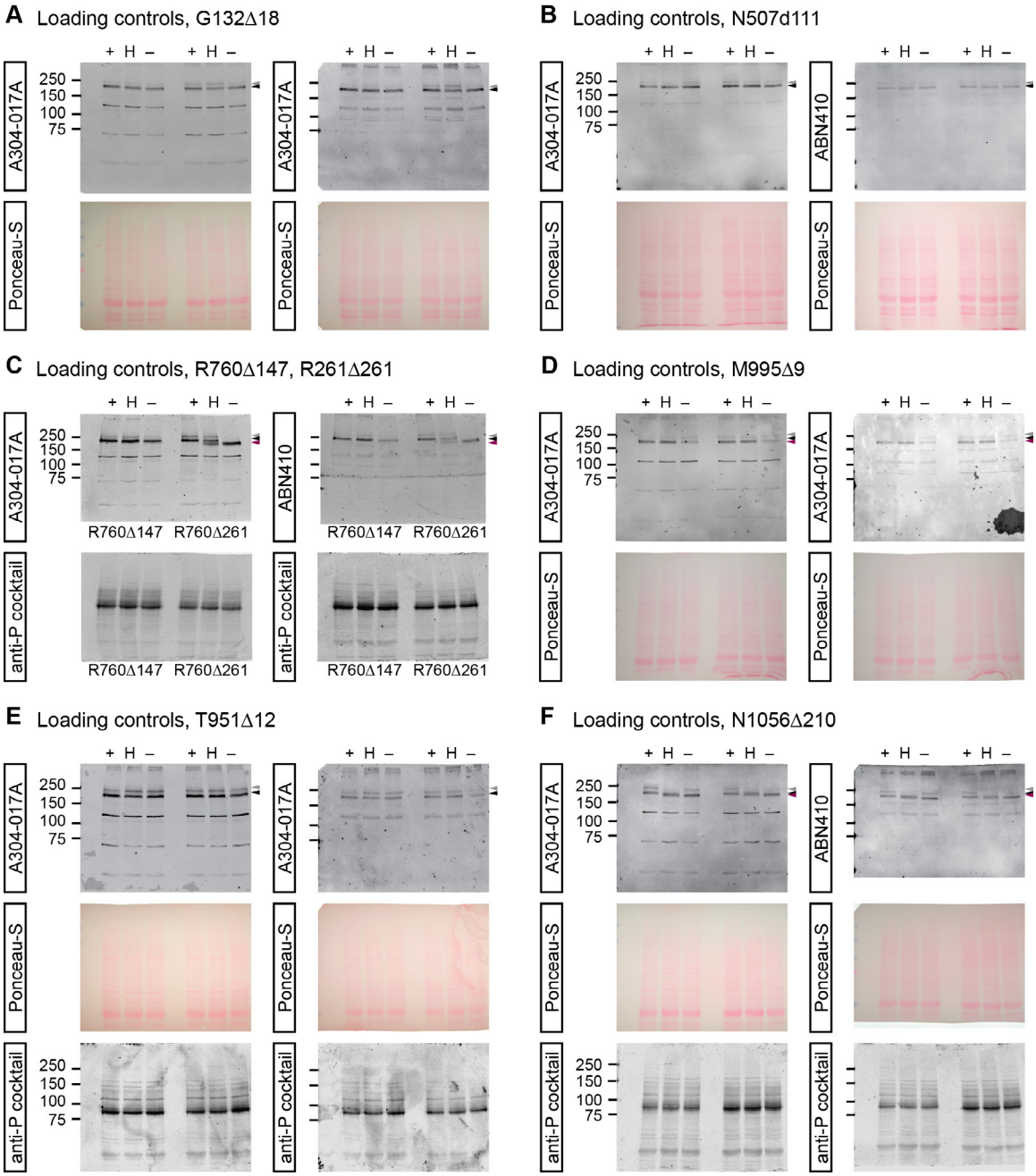
S2_Fig. Related to Figures 3-5. Western blots and loading controls. Full blots and Ponceau-S stained membranes for blots shown in Figure 3C (A), Figure 4A (B), Figure 4B (C), 5C (D) and S3_Fig.C. Size marker molecular weight in kDa is shown to the left. Position of the primary Zfp423 band is indicated by a black arrowhead to the right of the blot and the inconsistent conformational isomer by a gray and any consistently observed mutant specific band is indicated by a purple arrowhead.



S3_Fig related to Figure 5. P913L is not pathogenic in mice. (A) Surface views of brains from control and mutant same-sex littermate pairs shows grossly normal brains for P913L substitution allele edited independently on FVB/NJ and C57BL/6 (B6) strain backgrounds. (B) Forebrain images show apparently normal structure for P913L mutant on both backgrounds, while highlighting different extent of lateral ventricles between strains at this place of section. (C) Western blots for FVB-P913L (top) or B6-P913L (bottom) with antibody against residues 250-300 (A304-017A) show normal Zfp423 protein abundance. Results from two distinct trios shown. Ratios between same-sex littermates for (D) vermis area at midline, corpus callosum thickness at midline, average cortical thickness at 15°, 30° and 45° from midline, and body weight fail to identify defects in P913L homozygotes. Scale bars, 1 mm.

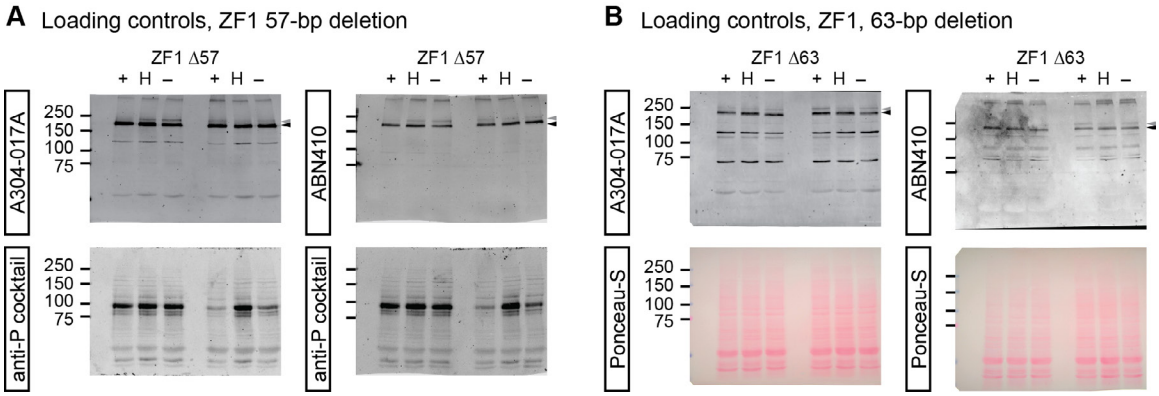


S4_Fig related to Figure 5.. R89H and E1124K are not pathogenic in mice. (A) Dorsal surface, coronal forebrain, and sagittal hindbrain views from control and mutant same-sex littermate pairs showed grossly normal brains for R89H homozygous, E1124K homozygous, or R89H/E1124K compound (trans) heterozygous animals on FVB/NJ background. Ratios between same-sex littermates for (B) vermis midline area, (C) average cortical thickness, (D) midline corpus callosum thickness, and (E) weight at sacrifice fail to identify significant deviations for any of these genotypes. Scale bars, 1 mm.



S5_Fig related to Figure 6. Western blots and loading controls. Full blots and loading control images for in-frame deletion variants G132Δ18 (A), N507Δ111 (B), R760Δ147 and R760Δ261 (C), M995Δ9 (D), T951Δ12 (E), and N1056Δ210 (F). Size marker molecular weight in kDa is shown to the left. Position of the primary Zfp423 band is indicated by a black arrowhead to the right of the blot and the inconsistent

conformational isomer by a gray and any consistently observed mutant specific band is indicated by a purple arrowhead.



S6_Fig related to Figure 7. Western blots and loading controls. Full blots and loading control images for 57-bp (A) and 63-bp (B) in-frame deletions of zinc finger 1. Size marker molecular weight in kDa is shown to the left. Position of the primary Zfp423 band is indicated by a black arrowhead to the right of the blot and the inconsistent conformational isomer by a gray and any consistently observed mutant specific band is indicated by a purple arrowhead.