



Published in final edited form as:

Genet Med. 2020 March ; 22(3): 538–546. doi:10.1038/s41436-019-0693-9.

***De novo* and inherited variants in *ZNF292* underlie a neurodevelopmental disorder with features of autism spectrum disorder**

A full list of authors and affiliations appears at the end of the article.

Abstract

Purpose: Intellectual disability (ID) and autism spectrum disorder (ASD) are genetically heterogeneous neurodevelopmental disorders. We sought to delineate the clinical, molecular and neuroimaging spectrum of a novel neurodevelopmental disorder caused by variants in the zinc finger protein 292 gene (*ZNF292*).

Methods: We ascertained a cohort of 28 families with ID due to putatively pathogenic *ZNF292* variants that were identified via targeted and exome sequencing. Available data were analyzed to characterize the canonical phenotype and examine genotype-phenotype relationships.

Results: Probands presented with ID as well as a spectrum of neurodevelopmental features including ASD, among others. All *ZNF292* variants were *de novo*, except in one family with dominant inheritance. *ZNF292* encodes a highly conserved zinc finger protein that acts as a transcription factor and is highly expressed in the developing human brain supporting its' critical role in neurodevelopment.

Conclusion: *De novo* and dominantly inherited variants in *ZNF292* are associated with a range of neurodevelopmental features including ID, and ASD. The clinical spectrum is broad, and most individuals present with mild to moderate ID with or without other syndromic features. Our results

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

Correspondence: Ghayda M. Mirzaa, M.D., Center for Integrative Brain Research, Seattle Children's Research Institute, 1900 Ninth Avenue, M/S JMB-10, Seattle WA 98101. ghayda.mirzaa@seattlechildrens.org.

Conflict of interest notification page

Sandra Yang, Rick Person, and Kristin G. Monaghan are employees of GeneDx, Inc., a wholly owned subsidiary of OPKO Health, Inc. Evan E. Eichler, is on the scientific advisory board (SAB) of DNAnexus, Inc. Jennifer Friedman's spouse is founder and principal of Friedman Bioventure which holds a variety of publicly traded and private biotechnology interests. The remaining authors have no competing financial interests.

Web resources

ExAC database, <http://exac.broadinstitute.org/>

GnomAD, <http://gnomad.broadinstitute.org>

Gene, <http://www.ncbi.nlm.nih.gov/gene/>

Online Mendelian Inheritance in Man (OMIM), <http://www.omim.org/>

PDB, <http://www.rcsb.org/pdb/home/home.do>

Combined Annotation Dependent Depletion (CADD), <http://cadd.gs.washington.edu/>

database for nonsynonymous SNPs' functional predictions (dbNSFP), <https://sites.google.com/site/jpopgen/dbNSFP>

REVEL, <https://sites.google.com/site/revelgenomics/>

dbSNP, <http://www.ncbi.nlm.nih.gov/SNP/>

MyGene2, NHGRI/NHLBI University of Washington-Center for Mendelian Genomics (UW-CMG), Seattle, WA (URL: <http://www.mygene2.org> [accessed January 2019]).

suggest that variants in *ZNF292* are likely a recurrent cause of a neurodevelopmental disorder manifesting as ID with or without ASD.

Keywords

ZNF292; intellectual disability; autism spectrum disorders; next generation sequencing; exome sequencing

Introduction

Knowledge about the genetic architecture of intellectual disability (ID), developmental delay (DD) and autism spectrum disorder (ASD) has increased dramatically over the past decade with the wide application of exome and genome sequencing (ES/GS) methods. As these genetic tools are becoming increasingly available in both the clinical diagnostic and research settings, a growing number of children with neurodevelopmental disorders (NDV) are now identified to have genetic variants that arise either *de novo*, or inherited as autosomal dominant, X-linked, or, less commonly, autosomal recessive traits. Discovery of more than 1000 genes underlying ID and / or ASD, to date, has markedly informed the diagnosis for families with ID/ASD and has further also led to the identification and characterization of multiple cellular pathways involved in human brain development, behavior, learning, and memory.¹⁻⁴ Such gene discovery efforts are important as the developmental roles of many of these pathways would not have otherwise been predicted from *in vitro* and model organism studies. Herein, we describe the clinical, neuroimaging and molecular features of 28 individuals with ID and / or ASD due to *de novo* or inherited variants in the zinc finger protein 292 gene (*ZNF292*; MIM: 616213).

ZNF292 encodes a highly conserved zinc finger protein that acts as a transcription factor. *ZNF292* is composed of eight exons, the last of which is the largest and encodes all 16 highly conserved zinc fingers of the predicted 2723-residue protein (canonical transcript in GenBank: [NM_015021.2](#)). Three of these zinc fingers (10–12) bind DNA at the promoter of growth hormone where it cooperates with POU1F1, a member of the POU family of transcription factors known to activate transcription in somatotrophs.⁵ Accordingly, the *ZNF292* protein was originally described as an enhancer of growth hormone (GH) expression in the pituitary gland of a rat animal model.⁵ Its' role was further delineated as a tumor suppressor with critical roles in tumor development and progression.⁶ However, the role of *ZNF292* in neurodevelopment is virtually unknown.

Materials and methods

Cohort ascertainment.

We identified 28 families with *de novo* (N = 27) or inherited (N = 1) pathogenic variants in *ZNF292* using a combination of trio-based ES (20 families) and multi-gene panels (8 families), in both clinical diagnostic and research settings. Families were identified across twenty institutions in six countries with data shared via nodes of the MatchMaker Exchange (MME) network, including MyGene2, GeneMatcher, PhenomeCentral, and by querying investigators with large cohorts of patients with ID and / or ASD.⁷ We identified 12

additional families (15 affected persons total) with variants in *ZNF292* that were considered likely pathogenic in a diagnostic setting, or “suspected” pathogenic in a research setting but for which our confidence in the pathogenicity of these variants was limited due to either incomplete parental testing or the identification of a missense *ZNF292* variant of unclear significance. Therefore, we excluded these families to avoid confounding description of both the canonical phenotype and genotype-phenotype relationships (Supplemental Subjects and Methods). We collected and reviewed detailed clinical data including medical records, facial photographs, and magnetic resonance images (MRI), when available, from affected individuals (summarized in Table 1 and S1). The institutional review board of the University of Washington approved this study, and patient consents were obtained from all individuals for whom identifiable data are presented.

Molecular Methods.

Twenty families were tested via exome sequencing (ES) in either a clinical or research setting, and eight families had targeted sequencing of multi-gene panels. Of persons tested via a targeted multi-gene panel, five underwent targeted capture of a panel for ID and three underwent testing via a single molecule molecular inversion probe (smMIP) based panel of more than 100 genes associated with ID / ASD.⁸ Targeted and exome sequencing methods are further provided in the Supplementary Methods.

RNA isolation, RT-PCR analysis and Sanger sequencing to analyze biallelic expression. We extracted total RNA from blood lymphocytes using the PAXgene Blood System (Becton Dickinson, Franklin Lakes, NJ, USA). First-strand cDNA synthesis was performed with the Superscript II Reverse Transcriptase Kit (Invitrogen, Carlsbad, CA, USA). If no exon spanning primers (see Figures S7 and S8 for sequences) could be designed, we performed DNase digestion of the RNA with the RNase-Free DNase Set (Qiagen, Venlo, Netherlands) prior to cDNA synthesis. All procedures were performed according to the manufacturer’s instructions. Resulting RT-PCR products were subsequently bidirectionally Sanger sequenced on an ABI 3730 sequencer (Applied Biosystems, Waltham, Massachusetts, USA) using the same primers and standard methods. RT-PCR Sanger traces were compared to DNA Sanger traces for biallelic expression at heterozygous variant positions with the Sequencher 5.1 software package (Gene Codes Corporation, Ann Arbor, MI, USA). Primer sequences used for segregation and RT-PCR are shown in the legends of figures S1 and S2.

RESULTS

We identified 23 *de novo* predicted loss of function (pLoF) variants (pLoFs; nonsense, frameshift, or splice) in 27 families and one transmitted (i.e., inherited from an affected mother) pLoF in one family (18–003) for a total of 24 putatively pathogenic variants in *ZNF292* (Table S1, Figure 1) in 28 families. Two variants were observed in multiple families: c.6160_6161del (p.Glu2054Lysfs*14) found in four unrelated individuals, and c.3066_3069del (p.Glu1022Aspfs*3) found in two unrelated individuals, one of whom previously published in a series of 96 individuals with NDV by a group of our authors (B.P., C.T., J.H., A.R., C.Z.).⁹ Two individuals in our cohort were also recently reported in a large cohort of individuals with ASD / ID: 17–022 with c.2490_2494dup (p. Ser832Ilefs*28) and

17–023 with c.4417dup (p. Ser1473Phefs*5).⁸ All *ZNF292* variants identified were absent in population controls (gnomAD release v2.1), with the exception of one variant that was present at a very low frequency: c.1360C>T (p.Arg454*) in 1 out of 248786 alleles (mean allele frequency 0.00000402). CADD (v1.4)¹⁰ scores (Phred-scaled) for the six nonsense variants in our series ranged from 35 to 42 with a median of 38.¹¹ Most variants (22/24) are located in exon 8, the last and largest exon of *ZNF292*, which encodes a large DNA binding domain of the protein (Figure 1). The majority of families (20/28, 71%) had pLoFs that were either insertions or deletions. Accordingly, we sought to explore whether local sequence context contributed to regional instability of this gene. At least seven of these 20 insertion/deletion events appear likely to have been influenced by sequence context, including five events within palindromic repeat sequences either flanking or directly adjacent to the breakpoints and two in which the variant occurred within a mononucleotide repeat sequence (Figure S3). This local sequence complexity of *ZNF292* may partially explain the high frequency of somatic mutations observed in *ZNF292* in tumor tissues as well.^{12–14}

Ten rare pLoFs in *ZNF292* are present in the “controls-only” subset of gnomAD (release v2.1), with six frameshifts and four nonsense variants that are predicted to affect the canonical transcript. However, manual review of many of these pLoFs suggest that they may be false positive calls, consistent with our observation of multiple palindromic sequences in *ZNF292* that complicates read alignment. For example, manual review of the BAM files available in the gnomAD browser reveals that two frameshifts, c.2574_2575delTC and c.2576_2577insAG, are observed only once and adjacent to one another suggesting they represent a single miscalled variant. Another variant that was annotated as nonsense, c.2690C>A, actually consists of two adjacent single-base substitutions and should have been annotated as a missense variant rather than a pLoF. A third variant, c.4592delC, was listed in the gene-overview gnomAD interface but was not actually called in the heterozygous or homozygous state in any individuals (Figure S4). Overall, only half (i.e. two frameshifts and three nonsense variants) in gnomAD controls appear to be of high quality. These findings are consistent with other reports, including recent guidance from the gnomAD consortium, on the need for manual curation and review of pLoFs called in gnomAD (*unpublished data; Minikel EV, Karczewski KJ, Martin HC, et al. Evaluating potential drug targets through human loss-of-function genetic variation. bioRxiv. 2019:530881*). Furthermore, none of the high quality variants in gnomAD are located between AA1588–2649, which correspond to zinc fingers 10–16 plus a putative coiled coil region and final nuclear localization signal, contrasting with variants in our cohort most of which are within these critical domains.^{15,16} *ZNF292* has a pLI score of 1.0 suggesting it is highly intolerant to loss-of-function (LoF) variants.¹⁷ Assessing the statistical significance of observing 28 families with pLoFs in *ZNF292* (23/24 variants being *de novo*) is challenging due to ascertainment bias as families were collected via matchmaking, rather than sequencing of a single cohort. Accordingly, the true denominator (i.e. number of cases sequenced worldwide and either available via MME or in large ID/DD/ASD cohorts) is unknown.¹⁸ This is a long-standing challenge for all studies of rare Mendelian disorders in general in which cases are ascertained and studied via matchmaking. Nevertheless, to estimate the probability of ascertaining 27 families with *de novo* pLoFs in *ZNF292*, we tested for enrichment of *de novo* variants in this gene.¹⁹ Specifically, we approximated the number of families with ID/DD/ASD that have been

sequenced worldwide and assumed that candidate genes for each family were either published or available for matchmaking via MME. The lower bound (of 100,000 individuals) is three-fold larger than the number of families with ID/DD/ASD who have reportedly undergone ES by GeneDx (*personal communication, K. Retterer, GeneDx, February 4, 2019*) and is the sum of families who underwent clinical sequencing from large diagnostic laboratories in the U.S. and Europe as well as those sequenced via research studies of large ID/DD/ASD cohorts. For an upper bound (300,000 individuals), we assumed that the ~800,000 persons with rare conditions estimated to have been sequenced worldwide likely represent ~350,000 families (~2 exomes per family; *unpublished data: Birney E, Vamathevan J, Goodhand P. Genomics in healthcare: GA4GH looks to 2022. bioRxiv. 2017*), and the primary indication for ~70% of those families was ID/DD/ASD. This is likely an overestimate, but it therefore serves as a conservative upper bound. Under these assumptions, the identification of *de novo* pLoFs in 27 independent families yields a significant enrichment of between 8.4-fold ($p=1.88\times 10^{-16}$ if $N=300,000$) to 25.3-fold ($p=1.93\times 10^{-28}$ if $N=100,000$) compared to an exome-wide significance cutoff of $p<2.7\times 10^{-6}$ under a Bonferroni adjustment for ~18,500 tests/genes. These data show that *ZNF292* variants are likely a recurrent cause of ID and / or ASD. Notably, *ZNF292* is not significantly depleted of NMD escape variants.²⁰ To examine this further, we performed RT-PCR on total RNA from two individuals, 17–005 with the c.3066_3069del (p.Glu1022Aspfs*3) variant and 19–011 with the c.1360C>T, (p.Arg454* variant), which showed bi-allelic expression of the normal and termination codon containing transcript, indicating that these transcripts are not degraded by nonsense-mediated mRNA decay (Figure S1 and S2).

All individuals in this cohort had ID with or without ASD and ADHD, with the exception of only one individual (17–023) who did not have evidence of ID but had ASD and speech delays at age 6 years. Of the individuals with ID, delays were mild in 11/27 (40%), moderate in 6/27 (22%) and severe in 3/27 (11%). A confirmed or suspected diagnosis of ASD was present in 17/27 (62%) individuals and of ADHD in 9/27 (33%) (Table 1, Figure S5). Speech delays were prominent in this cohort, seen in 26/27 (96%) individuals overall. One proband (17–027) had severe expressive language delays at age seven years, another (17–003) was minimally verbal at age five years and two children (17–013, 17–016) were nonverbal at the ages of four years and 18 years, respectively. Two children (17–003, 17–015) also had regression of speech and language development at ages six years and two years, respectively. Another individual (17–007), a 24-year-old male, had progressive developmental issues including memory problems with a suspicion for developmental regression overall. Most affected children walked prior to age two years with the exception of one child who remained nonambulatory at age four years. Notably, none of the individuals in this series had isolated behavioral issues without ID or ASD (Figure S5).

Growth abnormalities, including short stature, were diagnosed in 11/27 individuals (Figure S6). Tone abnormalities were observed in 13/27 individuals including hypotonia ($N = 10$), hypertonia or mixed tone ($N = 3$). Dysmorphic facial features, most notably micrognathia and hypertelorism, were observed in thirteen individuals (Figure 2). Less common facial features included prominent incisors, protruding ears, and prominent nasal bridge. Overall, these facial features are not characteristic. Ocular abnormalities including nystagmus,

esotropia, and strabismus were found in nine individuals. Four individuals had mild microcephaly with a head occipitofrontal circumference (OFC) of 2–3 standard deviations (SD) below the mean, and one child had an OFC of 4 SD below the mean at age four years. Overall, the observed dysmorphic features were nonspecific leading to likely low clinical recognizability of individuals with pathogenic *ZNF292* variants.

Notable brain abnormalities were detected in three of twelve probands who underwent brain magnetic resonance imaging (Figure 3, Tables S1 and S2). One child (17–009) had complex cerebellar abnormalities with hypoplasia of the cerebellar vermis and hemispheres, with marked cerebellar asymmetry and possible clefting within the cerebellum. There was evidence of asymmetric hemosiderin deposition on imaging suggestive of a previous vascular injury. However, there was no documented history of an *in utero* vascular insult or injury as pregnancy was uneventful and delivery was at term without complications. The other two children also had evidence of vascular injury on brain imaging. 18–003 was delivered at 25 weeks of gestation and had post hemorrhagic hydrocephalus and focal cystic encephalomalacia attributed to prematurity. 17–008 had a lacunar insult in the subcortical white matter with white matter injury but was delivered at term without notable complications during pregnancy. Other brain MRI findings present included ventriculomegaly, callosal abnormalities, and periventricular nodular heterotopia (each observed in one individual).

Finally, we identified 12 additional families (15 individuals total) with *ZNF292* variants in whom pathogenicity was suspected but was less certain because of one or more of the following reasons: (a) insufficient phenotype data from an individual with the candidate variant to determine their affected status; (b) inability to determine whether a candidate variant was inherited or *de novo* due to lack of parental genotype data (i.e. incomplete or absent parental testing information); and (c) the candidate variant was missense with no functional data to support pathogenicity.¹⁷ Nine of these families were identified by ES and six by multi-gene panel testing. The clinical and molecular data on this additional cohort are provided in Tables S3 and S4, and Figures S7 and S8. Notably, these variants were also rare or absent in gnomAD controls (N = 114, 704) and had high CADD scores. One family in this cohort harbored a variant that was also identified as a *de novo* variant in our primary cohort (p.Leu2221Serfs*10). However, this family was tested via a targeted multi-gene panel and there was insufficient data to determine whether the genotype segregated with affected individuals. Therefore, we conservatively assessed this family's clinical affected status to be uncertain. Nevertheless, it is likely that this secondary cohort is enriched for additional pathogenic variants.

DISCUSSION

We discovered variants in *ZNF292* that are likely pathogenic for a neurodevelopmental disorder variably accompanied by ASD and minor dysmorphic facial features that together delineate a novel condition with low clinical recognizability. Indeed, in none of the families with pathogenic *ZNF292* variants was the diagnosis of a syndrome with high clinical recognizability even considered. Accordingly, we anticipate that the number of persons with neurodevelopmental delays caused by pathogenic variants in *ZNF292* identified via

multigene panels, exome- or genome sequencing will continue to grow. If this prediction is correct, it seems reasonable to wonder why *ZNF292* has not been previously reported as a priority candidate gene for NDV in large cohorts of probands or trios with ID/DD/ASD.

ZNF292 first appeared as a possible candidate gene for NDV in 2012 because it was included in a supplemental table of 77 genes in which *de novo* variants were found in a cohort of 100 trios with severe ID.² Over the next seven years, seven additional probands with NDV with *de novo* variants in *ZNF292* were reported across five different ID/DD/ASD cohorts adding up to a total of eight probands with *de novo* variants in *ZNF292* in ~8,800 families with ID/DD/ASD tested^{8,21–24} (Table S5). In only two of these cohorts was more than one proband identified with a *de novo* variant in *ZNF292*, and the largest of such studies (DDD) included only one individual with a pLoF; while the others had missense variants. Accordingly, none of these studies had adequate statistical power to detect a significant enrichment of *de novo* missense variants or pLoFs in *ZNF292*. A recently reported analysis of 187 candidate genes including *ZNF292* by one of our authors (H.G.) detected a significant association ($p=0.016$) with ID/DD/ASD only after combining *de novo* variants identified across 2,926 families with ID/DD/ASD from a previously reported association study.⁸ However, this result would not reach genome-wide significance and the lack of deep phenotyping limited conclusions about both the canonical phenotype and distribution of phenotypic effects in persons with pathogenic variants in *ZNF292*.

Our study nicely illustrates that substantially greater statistical power can be achieved by testing a very large sample size (i.e., putatively all families with ID/DD/ASD tested to date) via effective data sharing and summing rare variants in the same candidate gene across families tested in research and clinical labs with different (i.e., gene discovery vs. diagnostic testing), albeit, complementary motivations. This strategy is expected to be particularly productive for identifying moderate to large effect alleles for genetically heterogeneous conditions with low clinical recognizability and to be much more efficient nowadays with the emergence of multiple platforms that facilitate global sharing of candidate genes over the past several years. Indeed, the process of matching to build confidence that a candidate gene is causal would likely happen much more quickly today than the seven years, beginning in 2012, required to demonstrate that variants in *ZNF292* underlie a neurodevelopmental disorder. However, it should also be noted that most existing platforms for data sharing do not allow public data sharing, encourage sharing of all candidate genes identified in a family and their phenotypic data, or permit direct participation of families in matching.

ZNF292 is highly expressed in the developing human brain (and is among the most highly expressed ZNFs in the BrainSpan dataset), especially the cerebellum, with the highest expression identified during the prenatal period (Figure 4). However, the mechanism by which variants in *ZNF292* disrupt human brain development and behavior are unclear. The two most likely possibilities by which pLoF might underlie disease are escape from nonsense-mediated decay (NMD), leading to expression of a truncated protein that has either a gain of function or dominant negative effect, or simple haploinsufficiency. Eighteen of the 20 unique pLoFs in our cohort are predicted to escape NMD,²⁰ and in contrast to the remaining handful of high quality pLoF calls in gnomAD, ten variants in our series overlap a residue between 1588–2649, which form zinc fingers 10–16, a putative coiled coil region,

and the final nuclear localization signal. This is a potentially significant concentration of variants overlapping those residues (Fisher exact test $p=0.004$).^{15,16} However, the seeming concentration of pLoFs in NMD-escape regions of the gene is consistent with the distribution expected by random chance (Fisher exact test $p=1$) and not necessarily indicative of the pathogenic mechanism as the last exon is very large (7,152 bp, or ~88% of total coding transcript length).²⁵ Although the prematurely truncated transcript is expressed, it is still possible that the pathogenic mechanism is that of haploinsufficiency, depending what functions are retained by the truncated transcript. Notably, deletions of the 6q locus containing *ZNF292* have been identified in individuals with a range of developmental issues including ID and ASD, further supporting the role of this gene in neurobehavioral phenotypes.²⁶

Finally, one affected parent in our cohort had mild ID that was diagnosed as an adult, suggesting that affected persons may go undiagnosed or be diagnosed later in life. This is consistent with the observation of five pLoFs in the gnomAD “control” group that appear to be valid. These observations suggest that some pathogenic *ZNF292* genotypes are incompletely penetrant and/or they underlie mild ID/DD/ASD.

In summary, this study demonstrates that *de novo* and dominantly inherited variants in *ZNF292* are associated with a spectrum of neurodevelopmental features including ID, ASD, ADHD, among others. The clinical spectrum of individuals with *ZNF292* variants is broad, with evidence of incomplete penetrance. This cohort shows that variants in *ZNF292* are a recurrent cause of ID with or without ASD and other neurodevelopmental features.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Ghayda Mirzaa, M.D.^{1,2}, Jessica X. Chong, Ph.D.^{2,3}, Amélie Piton, M.D.^{4,5}, Bernt Popp, M.D.⁶, Kimberly Foss, M.S.¹, Hui Guo, Ph.D.⁷, Ricardo Harripaul, M.Sc.^{8,9}, Kun Xia, Ph.D.⁷, Joshua Scheck, B.S.¹, Kimberly A. Aldinger, Ph.D.¹, Samin A. Sajan, Ph.D.¹⁰, Sha Tang, Ph.D.¹¹, Dominique Bonneau, M.D., Ph.D.^{12,13}, Anita Beck, M.D., Ph.D.², Janson White, Ph.D.³, Sonal Mahida, M.G.C., C.G.C.¹⁴, Jacqueline Harris, M.D.¹⁴, Constance Smith-Hicks, M.D., Ph.D.¹⁴, Juliane Hoyer, M.D.⁶, Christiane Zweier, M.D., Ph.D.⁶, André Reis, M.D.⁶, Christian T. Thiel, M.D.⁶, Rami Abou Jamra, M.D.¹⁵, Natasha Zeid, M.S., C.G.C.¹⁶, Amy Yang, M.D.¹⁷, Laura S. Farach, M.D.¹⁸, Laurence Walsh, M.D.¹⁹, Katelyn Payne, M.S., C.G.C.¹⁹, Luis Rohena, M.D.^{20,21}, Milen Velinov, M.D., Ph.D.²², Alban Ziegler, M.D.^{12,23}, Elise Schaefer, M.D., Ph.D.²⁴, Vincent Gatinois, M.D., Ph.D.^{25,26,27}, David Geneviève, M.D., Ph.D.^{25,26,27}, Marleen E.H. Simon, M.D.²⁸, Jennefer Kohler, M.S.²⁹, Joshua Rotenberg, M.D.³⁰, Patricia Wheeler, M.D.³¹, Austin Larson, M.D.³², Michelle E. Ernst, M.S., C.G.C.³³, Cigdem I. Akman, M.D.^{33,34}, Rachel Westman, M.S., C.G.C.³⁵, Patricia Blanchet, M.D.²⁶, Lori-Anne Schillaci, M.D.³⁶, Catherine Vincent-Delorme, M.D.³⁷, Karen W. Gripp, M.D.³⁸, Francesca Mattioli, Ph.D.³⁹, Gwenaél Le

Guyader, M.D., Ph.D.⁴⁰, Bénédicte Gerard, M.D.⁴, Michèle Mathieu-Dramard, M.D.⁴¹, Gilles Morin, M.D.⁴², Roksana Sasanfar, M.D.⁴², Muhammad Ayub, M.D.⁴³, Nasim Vasli, Ph.D.⁴⁴, Sandra Yang, M.S., C.G.C.⁴⁵, Rick Person, Ph.D.⁴⁵, Kristin G. Monaghan, Ph.D.⁴⁵, Deborah A. Nickerson, Ph.D.³, Ellen van Binsbergen, M.D.²⁸, Gregory M Enns, M.B.Ch.B.^{29,46}, Annika M Dries, B.S.²⁹, Leah J Rowe, M.S., C.G.C.³², Anne CH Tsai, M.D.³², Shayna Svihovec, M.S., C.G.C.³², Jennifer Friedman, M.D.^{47,48}, Zehra Agha, M.D.⁴⁹, Raheel Qamar, M.D.⁴⁹, Lance H. Rodan, M.D.^{50,51}, Julian Martinez-Agosto, M.D., Ph.D.⁵², Charlotte W. Ockeloen, M.D.⁵³, Marie Vincent, M.D.⁵⁴, William James Sunderland, Ph.D.⁵⁵, Jonathan A. Bernstein, M.D., Ph.D.^{29,46}, Undiagnosed Diseases Network⁵⁶, Evan E. Eichler, Ph.D.^{3,57}, John B. Vincent, Ph.D.^{8,9}, University of Washington Center for Mendelian Genomics (UW-CMG)³, Michael J. Bamshad, M.D.^{2,3}

Affiliations

¹Center for Integrative Brain Research, Seattle Children's Research Institute, Seattle, Washington, USA ²Division of Genetic Medicine, Department of Pediatrics, University of Washington School of Medicine, Seattle, Washington, USA ³Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington, USA ⁴Molecular Genetic Unit, Strasbourg University Hospital, Strasbourg, France ⁵Institute of Genetics and Molecular and Cellular Biology, Université de Strasbourg, Illkirch, France ⁶Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander-Universität (FAU) Erlangen-Nürnberg, Erlangen, Germany ⁷Center for Medical Genetics & Hunan Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, Changsha, Hunan, China ⁸The Campbell Family Mental Health Research Institute, Centre for Addiction & Mental Health (CAMH), Toronto, ON, M5T 1R8, Canada ⁹Institute of Medical Science, University of Toronto, Toronto, ON, M5S 1A8, Canada ¹⁰Department of Clinical Genomics, Ambry Genetics, Aliso Viejo, California, USA ¹¹WuXi NextCODE, Cambridge, Massachusetts, USA ¹²Département de Biochimie et de Génétique, CHU d'Angers, Angers, France ¹³UMR INSERM 1083 CNRS 6015, Angers, France ¹⁴Department of Neurogenetics, Kennedy Krieger Institute, Baltimore, Maryland, USA ¹⁵Institute of Human Genetics, University Medical Center Leipzig, Leipzig, Germany ¹⁶Yale New Haven Health, New Haven, Connecticut, USA ¹⁷Department of Molecular and Medical, Oregon Health and Science University, Portland, Oregon, USA ¹⁸Department of Pediatrics, McGovern Medical School at the University of Texas Health Sciences Center, Houston, Texas, USA ¹⁹Indiana University Health at Riley Hospital for Children, Indianapolis, Indiana, USA ²⁰Division of Genetics, Department of Pediatrics, San Antonio Military Medical Center, San Antonio, Texas, USA ²¹Department of Pediatrics, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA ²²New York State Institute for Basic Research in Developmental Disability, Staten Island, New York, USA ²³Service de Génétique Médicale, Centre hospitalier, le Mans, France ²⁴Service de Génétique Médicale, Hôpitaux Universitaires de Strasbourg, Institut de Génétique Médicale d'Alsace, Strasbourg, France ²⁵Service de génétique clinique, Département de Génétique Médicale, Maladies Rares et Médecine Personnalisée,

France ²⁶Centre de Référence Maladies Rares Anomalies du Développement et Syndromes Malformatifs Sud-Ouest Occitanie Réunion, Hôpital Arnaud de Villeneuve, Montpellier, France ²⁷Université Montpellier, Unité Inserm U1183, Montpellier, France ²⁸Department of Genetics, University Medical Center Utrecht, 3508 GA Utrecht, The Netherlands ²⁹Stanford Center for Undiagnosed Diseases, Stanford University, Stanford, California, USA ³⁰Memorial Hermann Memorial City Medical Center, Houston, Texas, USA ³¹Arnold Palmer Hospital for Children, Orlando, Florida, USA ³²Section of Genetics, Department of Pediatrics, University of Colorado School of Medicine and Children's Hospital Colorado, Aurora, Colorado, USA ³³Institute for Genomic Medicine, Columbia University Irving Medical Center, New York, USA ³⁴Division of Pediatric Neurology, Columbia University Irving Medical Center, New York, USA ³⁵Division of Genetics, St. Luke's Clinic, Boise, Idaho, USA ³⁶Department of Genetics and Genome Sciences, Case Western Reserve University, University Hospitals Cleveland Medical Center, Cleveland, Ohio, USA ³⁷Service de Génétique Clinique Guy Fontaine Centre de référence maladies rares Anomalies du développement, Hôpital Jeanne de Flandre Lille, France ³⁸Department of Pediatrics, Al duPont Hospital, Wilmington, Delaware, USA ³⁹Institut de Genetique et de Biologie Moleculaire et Cellulaire, Illkirch-Graffenstaden, France ⁴⁰Service de Génétique Clinique, Centre de compétence Maladies rares Anomalies du développement, CHU de Poitiers, Poitiers, France ⁴¹Service de Génétique Clinique Centre de référence maladies rares Anomalies du développement, CHU Amiens-Picardie, Amiens, France ⁴²Children's Medical Center, UMass Memorial Medical Center, Worcester, Massachusetts, USA ⁴³Department of Psychiatry, Queen's University, Kingston, ON, K7M 8A6, Canada ⁴⁴Division of Clinical & Metabolic Genetics, Hospital for Sick Children, Toronto, ON, M5G 1X8, Canada ⁴⁵GeneDx, Gaithersburg, Maryland, USA ⁴⁶Division of Medical Genetics, Department of Pediatrics, Lucile Packard Children's Hospital, Stanford University, Stanford, California, USA ⁴⁷Departments of Neurosciences and Pediatrics, University of California San Diego and Division of Neurology, Rady Children's Hospital, San Diego, California, USA ⁴⁸Rady Children's Institute for Genomic Medicine, California, USA ⁴⁹Department of Biosciences, COMSATS University, Islamabad, Pakistan ⁵⁰Division of Genetics and Genomics, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA ⁵¹Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA ⁵²David Geffen School of Medicine at UCLA, Los Angeles, California, USA ⁵³Department of Human Genetics, Radboud University Medical Center, Nijmegen, the Netherlands ⁵⁴CHU de Nantes, Service de génétique médicale, Nantes, France ⁵⁵University of Washington Foundation Board, University of Washington, Seattle, Washington, USA ⁵⁶NIH Undiagnosed Diseases Network, Office of the Director and the National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA ⁵⁷Howard Hughes Medical Institute, University of Washington, Seattle, Washington, USA

Acknowledgements

We thank the families and referring physicians for their participation in this study. We also thank people from the Unit of Molecular Genetics and the Centre National de Recherche en Génomique Humaine, especially Claire Feger, Elsa Nourrisson and Céline Cuny for the library preparation, DNA sequencing and bioinformatics analysis for respective individuals. Analysis of ES data performed at the University of Washington Center for Mendelian Genomics (UW-CMG) was funded by the National Human Genome Research Institute (NHGRI) and the National Heart, Lung and Blood Institute grants UM1 HG006493 and U24 HG008956 (to D.A.N. and M.J.B.). This work was supported by grants U01HL131003, UM1HL098147, UM1HL098123, UM1HL128761, UM1HL128711, and UM1HL098162 in support of the Pediatric Cardiac Genomics Consortium from the National Heart, Lung, and Blood Institute and the Eunice Kennedy Shriver National Institute of Child Health and Human Development (to D.A.N., and M.J.B.), the Intramural Research Program of the NHGRI and by the National Institutes of Health (NIH) through the Office of Strategic Coordination and Office of the NIH Director under Award Numbers U01 HG007708, U01 HG007709, U01 HG007703, U01 HG007530, U01 HG007942, U01 HG007690, U01 HG007674, U01 HG007672, U01 TR001395, U01 HG007943, and U54 NS093793, the Stanford Clinical and Translational Science Award (CTSA) to Spectrum (UL1 TR001085), the CAMH Foundation (to J.V.), a NARSAD New Investigator Award (to N.V.), a Peterborough K.M. Hunter Charitable Foundation Graduate Scholarship (to R.H.), the Special Education Organization of Iran for additional support for family recruitment (to R.S.), the DFG grant ZW184/6-1 (to C.Z.), the DFG grant PO2366/2-1 (to B.P.), the Agence de la Biomédecine and the CREGEMES, the Canadian Institutes of Health Research (#MOP-102758) (to J.V.), the National Natural Science Foundation of China (31671114) (to H.G.), the DFG393/2-2 ward (to R.A.J.), the National Natural Science Foundation of China (81525007) (to K.X.), the National Institute of Neurological Disorders and Stroke (NINDS) grant K08NS092898 (to G.M.M.), NIH R01 MH101221 (to E.E.E.), and Jordan's Guardian Angels (to G.M.M.). E.E.E. is an investigator of the Howard Hughes Medical Institute. Members of the Undiagnosed Disease Network (UDN) are listed in the Supplementary Data.

References

1. Chong JX, Yu JH, Lorentzen P, 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):788–795. [PubMed: 26656649]
2. de Ligt J, Willemsen MH, van Bon BW, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med.* 2012;367(20):1921–1929. [PubMed: 23033978]
3. Kochinke K, Zweier C, Nijhof B, et al. Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am J Hum Genet.* 2016;98(1):149–164. [PubMed: 26748517]
4. O’Roak BJ, Deriziotis P, Lee C, et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat Genet.* 2011;43(6):585–589. [PubMed: 21572417]
5. Lipkin SM, Naar AM, Kalla KA, Sack RA, Rosenfeld MG. Identification of a novel zinc finger protein binding a conserved element critical for Pit-1-dependent growth hormone gene expression. *Genes Dev.* 1993;7(9):1674–1687. [PubMed: 8370519]
6. Takeda H, Wei Z, Koso H, et al. Transposon mutagenesis identifies genes and evolutionary forces driving gastrointestinal tract tumor progression. *Nat Genet.* 2015;47(2):142–150. [PubMed: 25559195]
7. Philippakis AA, Azzariti DR, Beltran S, et al. The Matchmaker Exchange: a platform for rare disease gene discovery. *Hum Mutat.* 2015;36(10):915–921. [PubMed: 26295439]
8. Guo H, Wang T, Wu H, et al. Inherited and multiple de novo mutations in autism/developmental delay risk genes suggest a multifactorial model. *Mol Autism.* 2018;9:64. [PubMed: 30564305]
9. Popp B, Ekici AB, Thiel CT, et al. Exome Pool-Seq in neurodevelopmental disorders. *Eur J Hum Genet.* 2017;25(12):1364–1376. [PubMed: 29158550]
10. 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–D894. [PubMed: 30371827]
11. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med.* 2015;17(5):405–423. [PubMed: 25741868]

12. Camacho N, Van Loo P, Edwards S, et al. Appraising the relevance of DNA copy number loss and gain in prostate cancer using whole genome DNA sequence data. *PLoS genetics*. 2017;13(9):e1007001. [PubMed: 28945760]
13. Lee JH, Song SY, Kim MS, Yoo NJ, Lee SH. Frameshift mutations of a tumor suppressor gene ZNF292 in gastric and colorectal cancers with high microsatellite instability. *APMIS: acta pathologica, microbiologica, et immunologica Scandinavica*. 2016;124(7):556–560.
14. Yang P, Qiu Z, Jiang Y, et al. Silencing of cZNF292 circular RNA suppresses human glioma tube formation via the Wnt/beta-catenin signaling pathway. *Oncotarget*. 2016;7(39):63449–63455. [PubMed: 27613831]
15. Kosugi S, Hasebe M, Tomita M, Yanagawa H. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proc Natl Acad Sci U S A*. 2009;106(25):10171–10176. [PubMed: 19520826]
16. Wolf E, Kim PS, Berger B. MultiCoil: a program for predicting two- and three-stranded coiled coils. *Protein Sci*. 1997;6(6):1179–1189. [PubMed: 9194178]
17. Lek M, Karczewski KJ, Minikel EV, et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536(7616):285–291. [PubMed: 27535533]
18. Akle S, Chun S, Jordan DM, Cassa CA. Mitigating false-positive associations in rare disease gene discovery. *Hum Mutat*. 2015;36(10):998–1003. [PubMed: 26378430]
19. Ware JS, Samocha KE, Homsy J, Daly MJ. Interpreting de novo Variation in Human Disease Using denovolyzeR. *Curr Protoc Hum Genet*. 2015;87:7.2521–15.
20. Coban-Akdemir Z, White JJ, Song X, et al. Identifying Genes Whose Mutant Transcripts Cause Dominant Disease Traits by Potential Gain-of-Function Alleles. *Am J Hum Genet*. 2018;103(2):171–187. [PubMed: 30032986]
21. De Rubeis S, He X, Goldberg AP, et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature*. 2014;515(7526):209–215. [PubMed: 25363760]
22. Deciphering Developmental Disorders S Prevalence and architecture of de novo mutations in developmental disorders. *Nature*. 2017;542(7642):433–438. [PubMed: 28135719]
23. Homsy J, Zaidi S, Shen Y, et al. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science*. 2015;350(6265):1262–1266. [PubMed: 26785492]
24. Wang T, Guo H, Xiong B, et al. De novo genic mutations among a Chinese autism spectrum disorder cohort. *Nat Commun*. 2016;7:13316. [PubMed: 27824329]
25. Hsu MK, Lin HY, Chen FC. NMD Classifier: A reliable and systematic classification tool for nonsense-mediated decay events. *PLoS One*. 2017;12(4):e0174798. [PubMed: 28369084]
26. Engwerda A, Frentz B, den Ouden AL, et al. The phenotypic spectrum of proximal 6q deletions based on a large cohort derived from social media and literature reports. *Eur J Hum Genet*. 2018;26(10):1478–1489. [PubMed: 29904178]

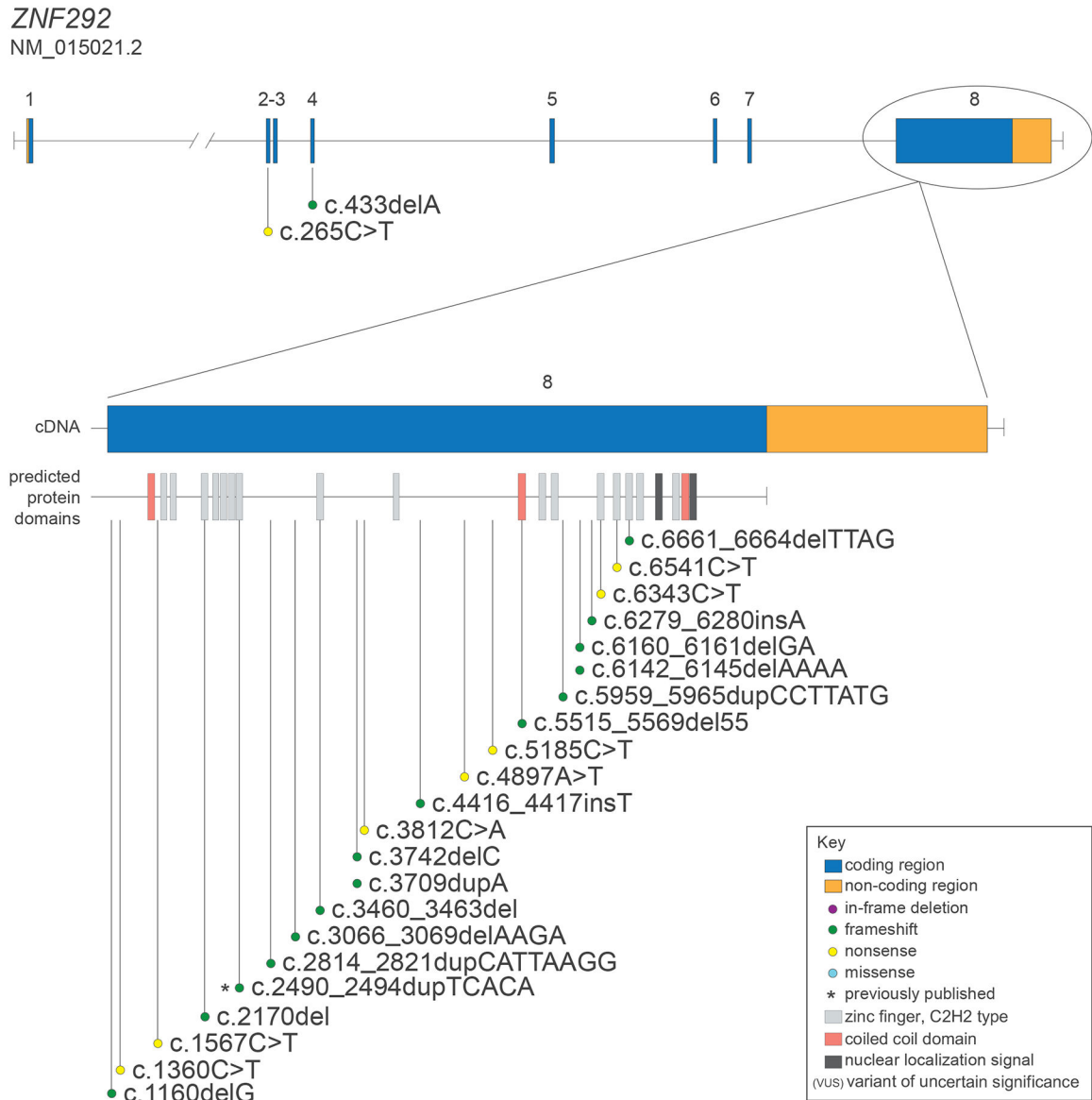


Figure 1. Genomic structure and distribution of variants in ZNF292.

Most of the identified variants in *ZNF292* are truncating (frameshift, nonsense) located within the largest and most terminal exon (8) of the gene that encodes a ZNF292 DNA binding domain. Several of these variants lie within zinc finger regions (depicted in gray) and coiled coil domains (depicted in pink) upstream of the nuclear localization signal (NLS, depicted in black). The cDNA panel shows the coding and non-coding regions of the gene (in blue and yellow, respectively). The bottom panel shows the predicted protein domains including the zinc finger (C2H2 type) regions (shown in gray), the coiled coil domain (pink) and the nuclear localization signal (black). *ZNF292* variants in the main cohort are shown, color-coded by type with nonsense variants shown in yellow, and frameshift variants in green. C-terminal coiled coil regions were calculated using multicoil2 (<http://cb.csail.mit.edu/cb/multicoil2/cgi-bin/multicoil2.cgi>)¹⁶, and NLS regions were mapped using cNLS mapper (http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi)¹⁵

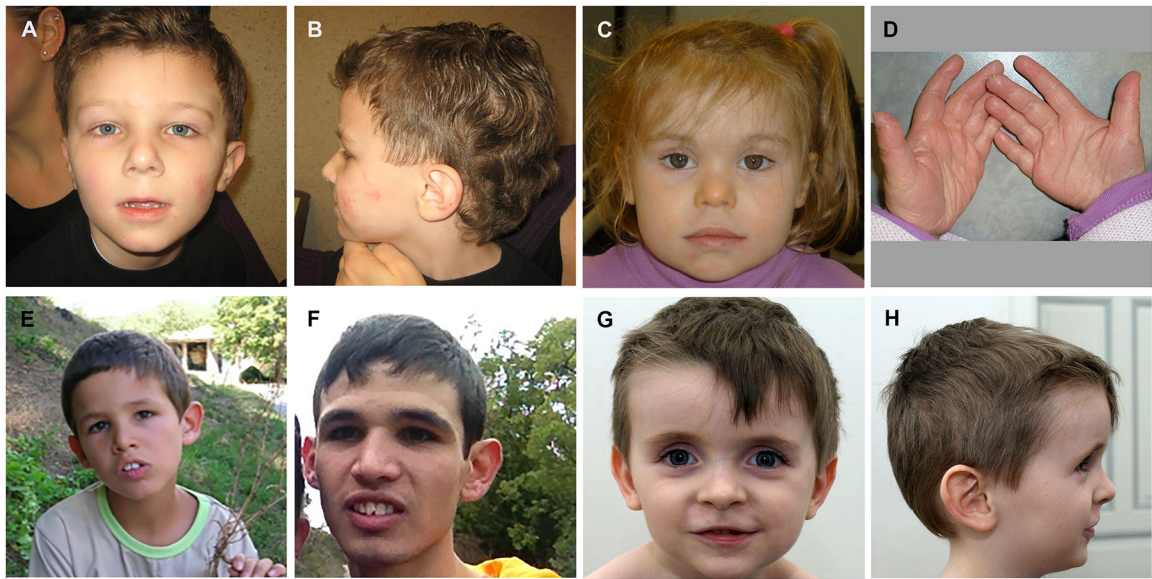


Figure 2. Facial features of individuals with pathogenic *ZNF292* variants.

(A, B) Photos of 17–027 showing a thin upper lip, smooth philtral folds, upturned nasal tip, sparse but long eyebrows with synophrys; (C–D) Photos of 18–007 at age 3.5 years showing epicanthal folds, mildly upslanted palpebral fissures, prominent forehead, and bulbous nose. D, hand photographs of the child showing ichthyosis; (E, F) Photo of 17–013 as a child (E) and as a teenager (F) showing laterally prominent ears, thick lips with a tented upper lip, short philtrum, prominent eyebrows with very prominent brow ridge and deep set eyes; (G, H) Frontal and lateral facial photograph of 17–005 at age 4 years 1 month showing mild micrognathia, short philtrum, and mildly downslanting palpebral fissures. All affected individuals have a prominent chin.

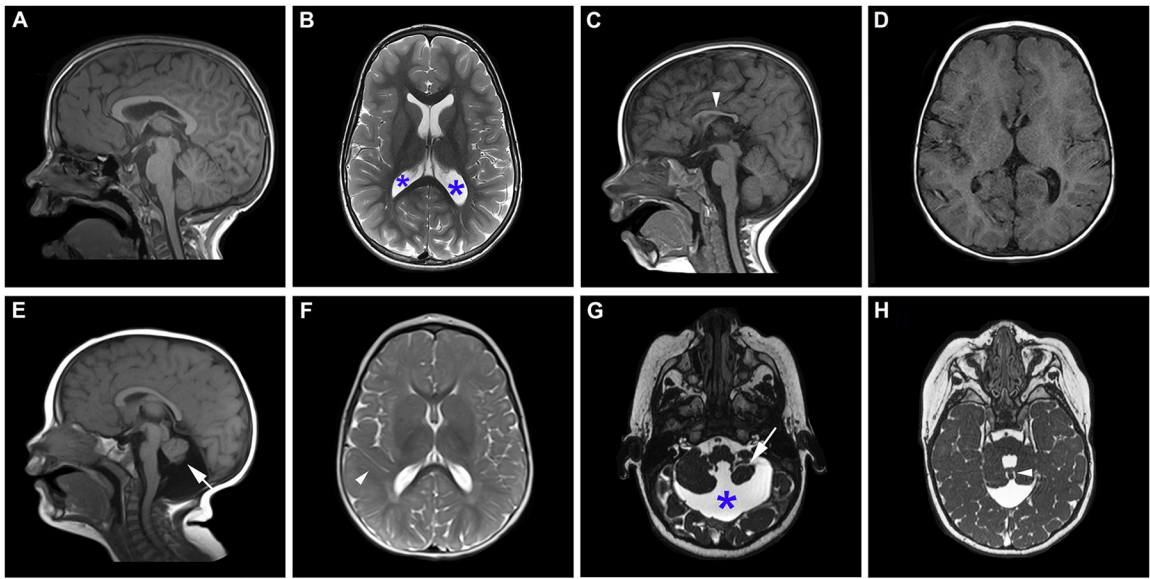


Figure 3. Brain MRI images of individuals with pathogenic variants in *ZNF292*.

(**A-B**) T1-weighted and T2-weighted brain MR images of patient 17-003 showing mildly prominent ventricles (asterisks, image F); (**C-D**) T1-weighted sagittal and axial images of 17-008 showing paucity of the white matter due to an in utero vascular insult and a thin corpus callosum (arrowhead; image G); (**E-H**) T1-weighted and constructive interference in steady state (CISS) images of patient 17-009 showing multiple abnormalities including hypoplasia of the cerebellar vermis and hemispheres, with marked asymmetry (arrow, image I; asterisks, image G), with possible clefting OF the cerebellum (arrow, image G), as well as a deep infold within the cortical surface (arrowhead, image F). Patient also has evidence of possible hemosiderin deposition that is asymmetric, suggesting a previous vascular insult/injury.

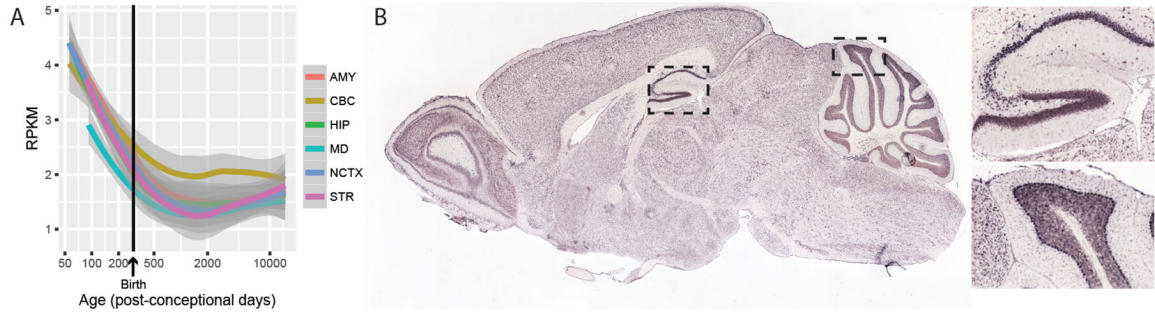


Figure 4. Expression of *ZNF292* in developing human and mouse brains. **A**, *ZNF292* expression in the developing human brain (normalized RPKM data) showing high expression during early prenatal development that diminishes in the postnatal brain. *Data obtained from BrainSpan <http://www.brainspan.org>.* Codes: AMY, amygdala; CBC, cerebellum; HIP, hippocampus; MD, medial dorsal nucleus of the thalamus; NCTX, neocortex; NCTX, neocortex. **B**, *Zpf292* expression in the adult mouse brain showing the highest expression (indicated by higher intensity staining) in hippocampus and Purkinje cells of the cerebellum.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1. Summary of the clinical features of *ZNF292* mutation-positive individuals (N=28)

Feature	Present	Absent	ND
Neurodevelopmental features			
ID/DD	27 (96%)	1 (4%)	0 (0%)
Speech delays	26 (93%)	1 (4%)	1 (4%)
ASD	17 (61%)	10 (36%)	1 (4%)
ADHD	9 (32%)	18 (64%)	1 (4%)
Tone abnormalities	13 (46%)	10 (36%)	5 (18%)
Brain MRI abnormalities	9 (32%)	8 (29%)	11 (39%)
Epilepsy	3 (11%)	23 (82%)	2 (7%)
Other features			
Dysmorphic features	13 (46%)	13 (47%)	2 (7%)
Ocular features	9 (32%)	17 (61%)	2 (7%)
Growth failure (weight/height <2 SD)	11 (39%)	15 (54%)	2 (7%)
Microcephaly (OFC <2SD)	4 (14%)	24 (86%)	0 (0%)
Feeding issues	8 (29%)	18 (64%)	2 (7%)
Constipation	6 (21%)	20 (71%)	2 (7%)
Skeletal abnormalities	5 (18%)	19 (68%)	4 (14%)
Cardiac abnormalities	5 (18%)	20 (71%)	3 (11%)

Abbreviations: ADHD, attention-deficit hyperactivity disorder; ASD, autism spectrum disorder; DD, developmental delays; ID, intellectual disability; ND, no data; OFC, occipitofrontal circumference; SD, standard deviations.