

Disruption of *POGZ* Is Associated with Intellectual Disability and Autism Spectrum Disorders

Holly A.F. Stessman,^{1,28} Marjolein H. Willemsen,^{2,3,28,*} Michaela Fenckova,^{2,3} Osnat Penn,¹ Alexander Hoischen,^{2,3} Bo Xiong,¹ Tianyun Wang,⁴ Kendra Hoekzema,¹ Laura Vives,¹ Ida Vogel,⁵ Han G. Brunner,^{2,3,6} Ineke van der Burgt,² Charlotte W. Ockeloen,² Janneke H. Schuurs-Hoeijmakers,² Jolien S. Klein Wassink-Ruiter,⁷ Connie Stumpel,⁶ Servi J.C. Stevens,⁶ Hans S. Vles,⁸ Carlo M. Marcelis,² Hans van Bokhoven,^{2,3} Vincent Cantagrel,^{9,10} Laurence Colleaux,^{9,10} Michael Nicouleau,^{9,10} Stanislas Lyonnet,^{9,10,11} Raphael A. Bernier,¹² Jennifer Gerdtts,¹² Bradley P. Coe,¹ Corrado Romano,¹³ Antonino Alberti,¹³ Lucia Grillo,¹⁴ Carmela Scuderi,¹⁵ Magnus Nordenskjöld,¹⁶ Malin Kvarnung,¹⁶ Hui Guo,⁴ Kun Xia,⁴ Amélie Piton,^{17,18} Bénédicte Gerard,¹⁸ David Genevieve,¹⁹ Bruno Delobel,²⁰ Daphne Lehalle,²¹ Laurence Perrin,²² Fabienne Prieur,²³ Julien Thevenon,²¹ Jozef Gecz,²⁴ Marie Shaw,²⁴ Rolph Pfundt,² Boris Keren,^{25,26} Aurelia Jacqueline,²⁵ Annette Schenck,^{2,3} Evan E. Eichler,^{1,27,29,*} and Tjitske Kleefstra^{2,3,29}

Intellectual disability (ID) and autism spectrum disorders (ASD) are genetically heterogeneous, and a significant number of genes have been associated with both conditions. A few mutations in *POGZ* have been reported in recent exome studies; however, these studies do not provide detailed clinical information. We collected the clinical and molecular data of 25 individuals with disruptive mutations in *POGZ* by diagnostic whole-exome, whole-genome, or targeted sequencing of 5,223 individuals with neurodevelopmental disorders (ID primarily) or by targeted resequencing of this locus in 12,041 individuals with ASD and/or ID. The rarity of disruptive mutations among unaffected individuals (2/49,401) highlights the significance ($p = 4.19 \times 10^{-13}$; odds ratio = 35.8) and penetrance (65.9%) of this genetic subtype with respect to ASD and ID. By studying the entire cohort, we defined common phenotypic features of *POGZ* individuals, including variable levels of developmental delay (DD) and more severe speech and language delay in comparison to the severity of motor delay and coordination issues. We also identified significant associations with vision problems, microcephaly, hyperactivity, a tendency to obesity, and feeding difficulties. Some features might be explained by the high expression of *POGZ*, particularly in the cerebellum and pituitary, early in fetal brain development. We conducted parallel studies in *Drosophila* by inducing conditional knockdown of the *POGZ* ortholog *row*, further confirming that dosage of *POGZ*, specifically in neurons, is essential for normal learning in a habituation paradigm. Combined, the data underscore the pathogenicity of loss-of-function mutations in *POGZ* and define a *POGZ*-related phenotype enriched in specific features.

Introduction

Intellectual disability (ID) and autism spectrum disorders (ASD [MIM: 209850]) are frequent, yet extremely heteroge-

neous, disorders. It is estimated that 10%–40% of persons with ID have ASD as a comorbidity.^{1–5} Although ID and ASD often present jointly in the same individual, genetic studies searching for rare causative variants have been

¹Department of Genome Sciences, University of Washington, Seattle, WA 98105, USA; ²Department of Human Genetics, Radboud University Medical Center, 6500 HB Nijmegen, the Netherlands; ³Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, 6500 HB Nijmegen, the Netherlands; ⁴The State Key Laboratory of Medical Genetics, School of Life Sciences, Central South University, Changsha, Hunan 410078, China; ⁵Department of Clinical Genetics, Aarhus University Hospital, Skejby, 8000 Aarhus, Denmark; ⁶Department of Clinical Genetics, Maastricht University Medical Center+, 6229 HX Maastricht, the Netherlands; ⁷Department of Genetics, University of Groningen and University Medical Center Groningen, 9700 RB Groningen, the Netherlands; ⁸Department of Neurologie, Maastricht University Medical Center+, 6229 HX Maastricht, the Netherlands; ⁹INSERM UMR 1163, Laboratory of Molecular and Pathophysiological Bases of Cognitive Disorders, Imagine Institute, 75015 Paris, France; ¹⁰Paris Descartes University – Sorbonne Paris Cité, Imagine Institute, 75015 Paris, France; ¹¹Departments of Genetics, Hôpital Necker – Enfants Malades, Assistance Publique – Hôpitaux de Paris, 75015 Paris, France; ¹²Department of Psychiatry and Behavioral Sciences, University of Washington, Seattle, WA 98105, USA; ¹³Pediatrics & Medical Genetics, Istituto di Ricovero e Cura a Carattere Scientifico Associazione Oasi Maria Santissima, 94018 Troina, Italy; ¹⁴Laboratory of Medical Genetics, Istituto di Ricovero e Cura a Carattere Scientifico Associazione Oasi Maria Santissima, 94018 Troina, Italy; ¹⁵Neuromuscular Diseases Unit, Istituto di Ricovero e Cura a Carattere Scientifico Associazione Oasi Maria Santissima, 94018 Troina, Italy; ¹⁶Department of Molecular Medicine and Surgery, Center for Molecular Medicine, Karolinska Institutet, 171 77 Solna, Sweden; ¹⁷Department of Translational Medicine and Neurogenetics, Institut Génétique Biologie Moléculaire Cellulaire, 67400 Illkirch, France; ¹⁸Laboratoire de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, 67000 Strasbourg, France; ¹⁹Département de Génétique Médicale, Centre de Référence Maladies Rares Anomalies du Développement et Syndromes Malformatifs Sud-Languedoc Roussillon, Hôpital Arnaud de Villeneuve, 34090 Montpellier, France; ²⁰Centre de Génétique Chromosomique, Hôpital Saint Vincent de Paul, Groupement des Hôpitaux de l'Institut Catholique de Lille, 59000 Lille, France; ²¹Équipe EA4271-GAD, Centre de Référence Anomalies du Développement et Syndromes Malformatifs de l'Est, Fédération Hospitalo-Universitaire TRANSLAD, 21079 Dijon, France; ²²Service de Génétique Clinique, Hôpital Robert Debré, 75019 Paris, France; ²³Service de Génétique Clinique Chromosomique et Moléculaire, Centre Hospitalier Universitaire de Saint-Étienne, 42100 Saint-Étienne, France; ²⁴Robinson Research Institute, University of Adelaide and Women's and Children's Hospital, North Adelaide SA 5006, Australia; ²⁵Département de Génétique, Hôpital de la Pitié-Salpêtrière, Assistance Publique – Hôpitaux de Paris, 75013 Paris, France; ²⁶INSERM UMR75 and U1127, CNRS UMR7225, Université Pierre et Marie Curie (Paris 6) and Institut du Cerveau et de la Moelle épinière, 75013 Paris, France; ²⁷Howard Hughes Medical Institute, Seattle, WA 98195, USA

²⁸These authors contributed equally to this work

²⁹These authors contributed equally to this work

*Correspondence: marjolein.willemsen@radboudumc.nl (M.H.W.), eee@gs.washington.edu (E.E.E.)

<http://dx.doi.org/10.1016/j.ajhg.2016.02.004>. ©2016 by The American Society of Human Genetics. All rights reserved.

Table 1. Clinical Features of Individuals Harboring POGZ Mutations from ID and/or DD Cohorts

	UMCN1	UMCN2	UMCN3	UMCN4	UMCN5	UMCN6	UMCN7	UMCN8	UMCN9
Mutation	c.2590C>T (p.Arg864*)	c.3001C>T (p.Arg1001*)	c.3456_3457del (p.Glu1154 Thrfs*4)	c.2263del (p.Glu755 Serfs*36)	c.1152dup (p.Arg385 Serfs*4)	c.2432+ 1G>A (p.?)	c.2020del (p.Arg674 Valfs*9)	c.3847C>T (p.Gln1283*)	c.3456_3457del (p.Glu1154 Thrfs*4)
Age (years)	5	13	9	6	2	12	5	26	8
Gender	F	M	M	M	M	M	F	M	M
ID/DD	++	++	+/-	+	++	+	+	++	+
Speech or language delay	++	++	++	+	+	++	+	++	++
Motor delay	+	+	+	+	+	+/-	+	+	+/-
ASD	+/-	+	+/-	+/-	+/-	+	+	-	-
Microcephaly	+	+	+	-	+	-	+/-	+	-
Feeding problems	ND	+	+	+	+	-	-	+	+
Vision problems	-	+	+	-	+	+	-	-	+
Obesity tendency	ND	-	-	-	+	+	-	+	+

All HGVS annotations were annotated on RefSeq transcript (GenBank: NM_015100.3). A full clinical description for each individual can be found in Table S1. Abbreviations are as follows: mo., months; ID, intellectual disability; DD, developmental delay; ASD, autism spectrum disorder; M, male; F, female; +, formal diagnosis (mild or moderate); ++, severe presentation; +/-, possessing some features and/or mild presentation; -, not present; ND, no data.
^ainheritance unknown.

focused on cohorts of individuals with either ID or ASD as the main inclusion criteria. At present, more than 700 genes have been implicated across a wide variety of ID syndromes with diverse clinical presentations,⁶ and the total number of ID-associated genes is estimated to be around 2,000.² Furthermore, over 100 genes have been reported in syndromic and non-syndromic forms of ASD.⁷ Not unexpectedly, many of the same genes underlie both ID and ASD, suggesting that the molecular networks and pathways in these disorders are largely overlapping.¹ Therefore, similar genetic defects might lead to either ID or ASD in different individuals or to both conditions in the same person.

The identification of candidate ID- and ASD-associated genes has increased tremendously by the widespread application of next-generation sequencing techniques, including trio-based whole-exome and whole-genome sequencing in research and routine clinical diagnostics.⁸⁻¹² Due to the extensive genetic heterogeneity, interpretation of pathogenicity of gene mutations is challenging because establishment of a conclusive molecular diagnosis is highly dependent on the identification of mutations in the same candidate gene in individuals with similar phenotypes. This is often hampered by the rarity of each individual genetic cause of ID and/or ASD. In addition, even genetic defects in the same gene can be associated with variable clinical features. To gain insight into the clinical spectrum of ID and/or ASD phenotypes associated with genes that have not been previously implicated in these conditions, it is important to collect clinical and molecular data from additional individuals.^{13,14} This can be facilitated by national and international collaboration

and data sharing of clinical and molecular databases. Bundling of knowledge and expertise on rare genetic causes of ID and/or ASD is of great value in the counseling of individuals. Recently, the core phenotypes of several distinct syndromes with both ASD and ID as hallmark features have been reported.¹⁵⁻¹⁸ Here, we report the core phenotype of a ID and/or ASD syndrome or phenotypic gestalt, which is caused by disruption of the pogo transposable element with zinc finger domain (*POGZ* [MIM: 614787]). A small number of mutations in *POGZ* were previously reported in different cohorts of individuals with ASD, ID, or schizophrenia (MIM: 181500); however, most of these studies did not provide further detailed clinical, molecular, or functional information on these individuals, leaving the pathogenicity of these mutations unclear.^{12,19-27} Meta-analyses of both small de novo deletions and/or single-nucleotide variants in families affected by ASD and developmental delay (DD) have shown that *POGZ* is likely a risk gene for neurodevelopmental disorders.^{25,28} Recently, truncating mutations in *POGZ* have been reported among small numbers of unrelated individuals. Although the phenotypes were variable, there is evidence of shared phenotypic features, suggesting that mutations in this gene might represent a distinct ASD and/or ID syndrome.^{29,30}

POGZ encodes a heterochromatin protein 1 α -binding protein containing a cluster of multiple C2H2-type zinc fingers, a centromere protein (CENP) B-like DNA-binding domain, and a DDE domain that might regulate gene expression.²⁰ *POGZ* is involved in mitosis³¹ and is expressed in the human fetal and adult brain (BrainSpan Atlas and GTEx). It is hypothesized to function as a

UMCN10	EE3	EE4	FR1	FR2	FR3	FR4	FR5	FR6	EE9 ^a
c.3040C>T (p.(Gln1014*))	c.2196_2198 delAG (p.Val733del)	c.2020del (p.Arg674 Valfs*9)	c.2545+1del (p.?)	c.2400dup (p.Lys801 Glnfs*7)	c.2836del (p.Asp946 Metfs*12)	c.2574del (p.His858 Glnfs*13)	c.1810G>T (p.Glu604*)	c.3001C>T (p.Arg1001*)	c.2501del (p.Leu834 Trpfs*20)
11	19	12	4	11	6	6	24	17	11 mo.
F	M	M	M	F	F	F	M	M	M
+	+	+	++	+	++	++	+	++	+/-
++	+	+	++	+	+	++	++	+	ND
+/-	++	+	++	+	+	+	+/-	+	ND
-	-	-	+	-	+	+	ND	+	ND
-	-	-	+	-	+	+	-	-	ND
+	ND	-	+	-	+	ND	-	-	ND
-	+	+	+	-	+	+	ND	+	ND
+	+	-	-	+	-	-	+	-	ND

transcriptional regulator in molecular networks crucial for neuronal function.^{11,19} *POGZ* has been shown to be co-expressed with ASD- and ID-associated genes involved in chromatin remodeling and gene transcription,³² such as *CHAMP1* (MIM: 616327).³³

We collected a cohort of 24 individuals with de novo mutations in *POGZ* by different approaches, including whole-exome sequencing in diagnostic and research cohorts, whole-genome sequencing in a research cohort, and targeted resequencing in research cohorts. The cohorts consisted of individuals diagnosed with various neurodevelopmental disorders, including ID and/or ASD. By studying their phenotype, we were able to further define and establish the core phenotype associated with disruptive mutations in *POGZ*. We provide further support for the importance of *POGZ* in cognitive function by utilizing a *Drosophila* knockdown model of the *POGZ* ortholog *row*. We found that downregulation of *row* expression, specifically in neurons, leads to deficits in habituation, a form of non-associative learning that is highly relevant for both ID and ASD.

Materials and Methods

Individual Selection

Persons with severe *POGZ* mutations (frameshift, nonsense, and splice) were identified by different strategies. 12 individuals with unexplained ID (UMCN1, UMCN3–UMCN10, FR1, FR3, and FR6), out of approximately 2,413 screened, were identified by diagnostic exome sequencing⁸ (Tables 1 and Table S1). One individual (UMCN2) was ascertained and published previously from whole-genome sequencing of 50 individuals with unexplained ID.¹²

Two individuals (EE1 and EE2) were identified and published previously from research exome sequencing of 2,377 simplex families affected by ASD.¹⁰ The remaining ten individuals were identified by a variety of targeted sequencing techniques. One individual (FR5) was identified with the Illumina TruSight One panel (Tables 1 and Table S1) from a cohort of 175 individuals with neurodevelopmental disorders. Two individuals (FR2 and FR4) were identified with a custom SureSelect panel of 275 ID-related genes (Tables 1 and Table S1) from a total of 208 probands with mild to severe ID. The remaining seven individuals (EE3–EE9) were identified with molecular inversion probe (MIP)-based sequencing (see below). Four missense events reported were identified by either research exome sequencing (in individuals EE11–EE13, reported previously)^{10,27} or MIP-based sequencing (in individual EE10). Subsequent phenotypic follow-up was performed by clinical interview with individuals and families. All experiments carried out on these individual samples were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national), and proper informed consent was obtained. For identification of disruptive variation in unaffected individuals, the Exome Aggregation Consortium (ExAC) data were used with neuropsychiatric cases masked (early access to these data were provided by Daniel MacArthur). This dataset is now publicly available for download on the ExAC portal.

Identification of *POGZ* Mutations by MIP Resequencing

Seven individuals with disruptive events and one individual with a missense event in *POGZ* were identified with MIP resequencing as previously described³⁴ from a research cohort of 12,041 individuals with unexplained ID and/or ASD (EE3–EE10). 4,025 sibling control individuals from the Simons Simplex Collection (SSC)³⁵ and The Autism Simplex Collection (TASC)³⁶ were also screened as unaffected individuals in this study. In total, we designed 53 small-molecule MIPs to tile across the coding regions and splice

junctions of the human *POGZ* locus. The design and concentration of each MIP probe used in the pool, as well as their individual performance, are detailed in Table S2. Secondary Sanger validation was performed on all individuals and their parents.

Calculation of De Novo Significance and Penetrance

De novo significance was calculated as previously described by Samocha et al.³⁷ with the R package *denovolyzeR* (v.3.1.0). To calculate the significance and penetrance of *POGZ* likely gene-disruptive (LGD) mutations in ID and/or DD, we utilized two control sets as a baseline: the first, representing 45,376 samples, was composed of the ExAC database with psychiatric cases removed (ExAC v.0.3) and the second was composed of the 4,025 control individuals sequenced by MIPs (see above). All significance calculations were based on the two LGD events found in this composite collection of 49,401 control individuals and the 25 LGD events found in 17,264 case individuals included in this study. Case-control significance was calculated with a standard two-tailed Fisher's exact test, and penetrance and its confidence bounds were calculated with the model described in Rosenfeld et al.,³⁸ specifically:

$$P(D | G) = \frac{P(G | D) P(D)}{P(G | D) P(D) + P(G | \bar{D}) P(\bar{D})}$$

where D = disease, G = genotype (the presence of an LGD event in *POGZ*), and \bar{D} = absence of disease. The general population incidence of ID and/or DD in our cohort was assumed to match that described in Rosenfeld et al. ($P(D) = 5.12\%$), given that our cohort composition has a similar representation of youth-onset diseases with an important genetic component with broad exclusion of chromosomal disorders.

Analysis of GTEx Data

Reads per kilobase of transcript per million mapped reads (RPKM) values were downloaded from the GTEx portal v.6 on 11 November 2015 and were used to generate boxplots for each of the *POGZ* isoforms included in GENCODE v.19 (NCBI browser GRCh37.p13). Ensembl identifiers from GENCODE were converted to RefSeq identifiers with the Ensembl portal (release 83). All RefSeq identifiers referenced here correspond to RefSeq release 73. To test the significance of the difference between the average expression of the pituitary and brain subtissues, permutation tests were conducted for every pair of subtissues. For each pair, we calculated the difference in means, d , and performed 10,000 permutations. In each permutation, the subtissue labels were randomly swapped and the difference in means calculated. We then calculated the average of the difference in means across all shuffles, m . We calculated p values as $(R + 1) / (N + 1)$ where R is the number of times the permuted difference in the mean is further from m than d and N is the number of permutations (10,000). Bonferroni's correction was used to correct for multiple testing.

Drosophila Knockdown Experiments

Fly stocks were kept on a standard *Drosophila* diet (cornmeal, sugar, and yeast) at 25°C and 45%–60% humidity at a 12 hr light-dark cycle. The inducible RNAi line against the *POGZ* *Drosophila* ortholog *row* (*vdrc28196*, no predicted off-targets) and its genetic background control line (*vdrc60000*) were obtained from the Vienna *Drosophila* Resource Center.³⁹ For the habituation experiments, RNAi was induced with the GAL4-UAS system and the panneuronal *elav-Gal4* driver line *w¹¹¹⁸*; *2xGMR-wIR*; *elav-Gal4*, *UAS-Dicer-2*. The genetic elements in this line suppress eye color as required for

the light-off jump response (*2xGMR-wIR*)^{40,41} and increase RNAi efficiency (*UAS-Dicer2*), respectively.³⁹ Flies were reared and tested at 25°C and 70% humidity. For qPCR experiments, RNAi was induced with the ubiquitous actin-Gal4 driver line (*w¹¹¹⁸*; *actin-Gal4/CyOGFP*).

Drosophila Light-Off Jump Reflex Habituation Assay

The light-off jump habituation assay was performed as previously described.⁴² Habituation of the startle jump response toward repeated light-off stimulus of 3- to 7-day-old individual male flies was tested in two independent 16-unit light-off jump habituation systems. 32 flies (16 flies per system) were simultaneously exposed to a series of 100 short (15 ms) light-off pulses with a 1 s inter-pulse interval. The noise amplitude of wing vibration after every jump response was recorded for 500 ms after the start of light-off pulse, and an appropriate threshold was applied to filter out the background noise. Data were collected and analyzed by a custom Labview Software (National Instruments). High initial jumping response to light-off pulse decreased with a growing number of trials, and flies were considered habituated when they failed to jump in five consecutive trials (no-jump criterion). Habituation was scored as the mean number of trials required to reach the non-jump criterion (trials to criterion [TTC]). The main effect of the genotype (mutant versus control, corrected for experimental day and system) on log-transformed TTC values was tested via linear model regression analysis (lm) with R statistical software (v.3.0.0).

Results

Identification of De Novo Mutations in *POGZ*

Through diagnostic whole-exome sequencing or targeted exome sequencing, we identified 15 individuals, out of approximately 2,796 case individuals, with unexplained ID and LGD mutations in *POGZ*. In order to identify additional individuals with *POGZ* mutations, we targeted and sequenced the coding portions of this gene by using MIPs (see Materials and Methods) on a large research cohort of 12,041 individuals with an ASD and/or ID diagnosis and identified seven additional events. Combined with three events that we previously identified,^{10,12} we collected 19 individuals with a primary diagnosis of ID (Tables 1 and Table S1) and 6 individuals with a primary diagnosis of ASD (Table 2) with LGD mutations in *POGZ* (Figure 1 and Table S3). These include 10 nonsense, 12 frameshift, 1 in-frame deletion, and 2 splice events.

In order to determine the significance of LGD mutations at this locus in case individuals, we examined several control cohorts for similar mutations. Using the MIP-based approach,³⁴ we targeted and sequenced the *POGZ* locus in 4,025 unaffected sibling control individuals and identified no LGD events and one de novo missense event that has been previously published¹⁰ (Table S3). Furthermore, when we queried the ExAC database (with neuropsychiatric cases masked), we identified two LGD events out of 45,376 individuals. Among 29,085 DD cases with published copy-number variant (CNV) data,⁴³ we observed

Table 2. Clinical Features of POGZ Individuals from ASD Cohorts

	EE1	EE2	EE5	EE6	EE7	EE8
Mutation	c.3600_3607dupTGATGACG (p.Glu1203Valfs*28)	c.3022C>T (p.Arg1008*)	c.1212C>A (p.Tyr404*)	c.538C>T (p.Gln180*)	c.3139G>T (p.Glu1047*)	c.2291del (p.Pro764Leufs*27)
Age (years)	8	14	21	7	7	6
Gender	M	M	F	F	M	F
ID/DD	+/-	+	+	+/-	+	+/-
Speech or language delay	+	+/-	+	++	++	+
Motor delay	+	++	ND	ND	+	+
ASD	+	+	+	+	+	+
Microcephaly	-	+/-	-	-	-	-
Feeding problems	-	-	ND	ND	-	+
Vision problems	+	ND	ND	ND	+	+
Obesity tendency	+	+	+/-	-	+	+

All HGVS annotations were annotated on RefSeq transcript NM_015100.3. A full clinical description for each individual can be found in [Table S1](#).

ID, intellectual disability; DD, developmental delay; ASD, autism spectrum disorder; M, male; F, female; +, formal diagnosis (mild or moderate); ++, severe presentation; +/-, possessing some features and/or mild presentation; -, not present; ND, no data.

one 8.3 Mbp duplication that subsumed the *POGZ* locus, but no gene-breaking CNVs ([Figure S1](#)). A similar analysis of 19,584 control individuals showed no CNVs at this locus. Previously published studies reported 2 de novo LGD events among probands with schizophrenia,^{21,24} 16 LGD events (including one gene-breaking CNV) among cases of ID, DD, and ASD,^{20,25,26,29,30} and 1 missense event in a case of ASD;²² however, none of these overlap directly with the 25 events reported in this study ([Table S3](#)). On the basis of these data, we estimate that protein-truncating mutations in *POGZ* are significantly enriched in ASD and/or ID individuals in comparison to the general population ($p = 4.19 \times 10^{-13}$, two-tailed Fisher's exact test; odds ratio = 35.8; 95% confidence interval [CI] = 8.9–311.9; significance calculations computed solely on the 25 cases characterized in this study). On the basis of these data, we estimate the penetrance of *POGZ* LGD mutations to be 65.9% (95% CI = 36.4%–89.8%) given the incidence of ID (5.12%) in the general population.³⁸

24 of the 25 *POGZ* truncating events were shown to be de novo, and there was one individual with an LGD event for whom inheritance status could not be determined (EE9). Using a previously described statistical framework for identifying excesses of de novo mutation by gene,³⁷ we estimate the probability of detecting 24 or more *POGZ* de novo LGD events in 17,264 cases to be $p = 5.85 \times 10^{-39}$. In addition to these LGD events, we also identified a number of missense variants. Exome sequencing studies previously reported four de novo missense variants among families with ASD^{10,22,27} ([Figure 1](#) and [Table S3](#)). Two of these events were in probands and one was in an unaffected sibling control individual. We identified one additional de novo ([Table S4](#)) and two private, inherited ([Table S5](#)) missense events through our MIP sequencing efforts.

Clinical Description and Phenotype Analysis

The clinical features of the 25 individuals with LGD mutations in *POGZ* are summarized in [Tables 1](#) and [2](#) with accompanying photographs ([Figure 2](#)). Detailed case reports are described in the [Supplemental Note](#). Though most individuals exhibit facial dysmorphisms, there is no significant pattern that could guide to a recognizable facial gestalt. However, individuals with disruptive mutations in *POGZ* showed a clear overlap in several other clinical features. Whereas all had some degree of DD varying from borderline to severe impairment, most individuals had mild ID. Speech and language were generally classified as more severely affected than motor development. Strikingly, in several individuals, the ability to speak in sentences and write and read simple language started very late, but was eventually acquired. In addition, a distinct neurobehavioral phenotype could be recognized, including either a formal ASD diagnosis or features of ASD and, in many cases, a seemingly contrary overly social and overly friendly demeanor. Hyperactivity and sleeping problems were frequently observed as well. Other medical concerns included infections, mainly of the upper respiratory tract at a younger age, and nonspecific ocular problems. In contrast to the LGD variants, the missense variants are not clearly associated with impaired intellectual capacity. However, the missense variants seem to be associated with a behavioral phenotype, including ASD or autistic-like features, reminiscent of that seen in individuals carrying LGD variants²² ([Table S4](#)).

To clarify the phenotype of individuals with the *POGZ* mutation, we compared several key phenotypic variables in the individuals with a *POGZ* mutation and an ASD diagnosis to those in a cohort of 2,718 children with ASD from the SSC.³⁵ The rate of reported problems that emerged through comprehensive individual follow-up, including

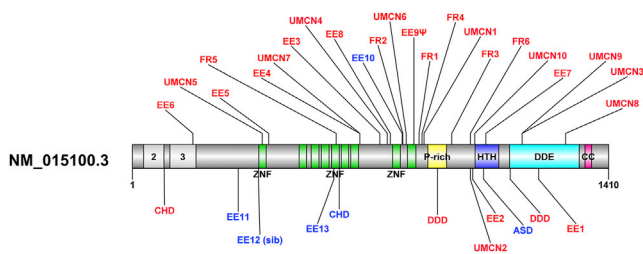


Figure 1. Protein Model of *POGZ* with Currently Identified Mutations Indicated

All mutations (indicated by individual identifiers that correspond to Tables 1 and 2 and Tables S1, S3, and S4) have been annotated on the RefSeq transcript (GenBank: NM_015100.3) (*POGZ*). Events in red are LGD and blue are missense. Mutations listed on the top of the protein structure have not been previously identified or reported. Mutations listed on the bottom of the protein structure have been published previously. Protein domains are indicated on the structure. Abbreviations are as follows: ZNF, zinc finger; HTH, helix-turn-helix; CC, coiled coil; CHD, congenital heart defect; ASD, autism spectrum disorder; DDD, developmental delay; sib, sibling. Light-gray shaded portions indicate amino acids omitted by alternatively spliced *POGZ* transcripts (3, *POGZ* isoform; 2, *POGZ* isoform 2). ^ψMutations for which inheritance is unknown. All other mutations are de novo.

sleep problems, feeding problems, vision problems, hyperactivity, obesity, and microcephaly, were compared across groups with the Fisher's exact test (Table S6). It should be noted that different approaches were used to assess sleep, feeding, and vision problems in this cohort than were used in the SSC. For the individuals with a *POGZ* mutation, the presence of hyperactivity and sleep, feeding, and vision problems were identified through formal clinical assessment. For the SSC cohort, sleep, feeding, and vision problems were established through standardized medical history interview. Hyperactivity was defined as meeting clinical cut-off on parent reporting on the Child Behavior Checklist Externalizing Domain T score.⁴⁴ Obesity and microcephaly were determined through examination of standardized BMI and head circumference values. Sleep disturbance was not statistically enriched in the *POGZ* cohort. Although feeding problems and obesity showed borderline significance ($p = 0.045$ and 0.07 , respectively), significant increases in vision problems ($p = 0.000067$), microcephaly ($p = 0.003$), and hyperactivity ($p = 0.004$) were observed in the individuals with a *POGZ* mutation and ASD relative to the SSC comparison cohort (Table S6).

In this case series, 12 of the 25 individuals (48%) were identified as clinically obese at some point in the developmental trajectory. This is in contrast to the feeding problems, including problems with chewing and swallowing and aversion to solid foods, that were reported in 10 of the 25 individuals (40%; Tables 1 and 2). For one individual (EE8), longitudinal growth data were available and showed that, although other growth measurements (e.g., height) were within normal ranges, body weight already exceeded three SDs above the mean starting at two-and-a-half years of age. This trend has been maintained through the age of six (the age of the last recording;

Figure S2A) and is in contrast to this individual's unaffected sibling, who consumes the same diet and falls within normal body mass ranges (Figure S2B). The rate of obesity (i.e., two SDs above the mean) among children with sporadic ASD is approximately 17% (based on the SSC), in comparison to 40% in our study (calculation performed only on individuals for whom these data were available; Table S6).

POGZ Expression Analysis

We explored the normal expression patterns of *POGZ* across human tissues by using the publicly available GTEx database. Of the ten isoforms identified in GTEx, only two are highly expressed across the majority of tissues: *POGZ* isoform 3 (GenBank: NM_145796.3) and isoform 2 (GenBank: NM_207171.2) (Figures 3A and 3B). The longest annotated isoform (GenBank: NM_015100.3) (Figure 3A) does not appear to be constitutively expressed (not shown). In adult tissues, *POGZ* isoforms 2 and 3 consistently show significantly increased expression in the cerebellum and the pituitary gland ($p < 0.01$, permutation test; Tables S7 and S8) in comparison to expression in all other brain subregions (Figure 3B). The most abundantly expressed RNA isoform 3 (GenBank: NP_665739.3) (Figure 3B, right panel) differs from the longest annotated isoform of *POGZ* only by omission of amino acids 112–189 encoded by two in-frame exons that are alternatively spliced out of the mature transcript. All individual mutations would affect protein isoform 3 of *POGZ* with the exception of that of individual EE6 (p.Gln180*; Figure 1). Isoform 2 (GenBank: NP_997054.1), which shows the second highest RNA expression pattern (Figure 3B, left panel), differs from the longest isoform by omission of amino acids 34–94. This isoform (GenBank: NM_207171.2) would contain all individual mutations (Figure 1). Further exploration of *POGZ* expression in fetal brain tissues via the BrainSpan Atlas shows that expression is highest during early embryonic development (9 post-conception weeks) and decreases gradually until birth, after which low-level expression is maintained into adulthood (Figure 3C).

Modeling *POGZ* Partial Loss of Function in *Drosophila*

To gain independent support for the implication of *POGZ* in ID/ASD pathologies and address whether the protein is required directly in neurons for normal functioning, we turned to the fruit fly *Drosophila melanogaster*, an organism that has already been successfully exploited to study specific aspects of the observed human pathologies.⁴⁶ The *Drosophila* genome harbors a single gene ortholog representing the human *POGZ* family (*POGZ*, ZNF280A-D); this ortholog is called relative of woc (*row*).^{47,48} According to the ModEncode and FlyAtlas systematic expression databases,^{49,50} *row* is expressed across developmental stages and tissues, and the highest expression is found in the larval CNS.^{38,39} We modeled the *POGZ* loss-of-function condition by conditional knockdown of *row* in *Drosophila*



Figure 2. Clinical Photographs of Individuals with De Novo LGD Mutations in *POGZ*

Individuals harboring *POGZ* mutations show an overlap in facial features, including brachycephaly (not shown) and a broad forehead, a high nasal bridge, hypertelorism, and a thin upper lip in some. However, the facial phenotype is not very specific or recognizable.

(A and B) Individual UMCN1 at the ages of 1 year (A) and 3 years (B).

(C) Individual UMCN2 at the age of 9 years.

(D and E) Individual UMCN3 at the ages of 4 years (D) and 8 years (E).

(F and G) Individual UMCN4 at the ages of 4.5 years (F) and 5 years, 2 months (G).

(H and I) Individual UMCN6 at the ages of 6 months (H) and 11 years (I).

(J) Individual UMCN7 at the age of 4 years.

(K and L) Individual UMCN8 as a child (K) and at the age of 26 years (L).

(M) Individual UMCN9 at the age of 8 years.

(N and O) Individual UMCN10 at the ages of 4 years (N) and 11 years (O).

(P) Individual EE4 at the age of 12 years.

(Q and R) Individual FR3 at the ages of 7 months (Q) and 6 years (R).

(S) Individual FR4 at the age of 6 years.

(T) Individual EE2 at the age of 14 years.

(U) Individual EE6 at the age of 7 years.

(V) Individual EE7 at the age of 7 years.

by using the GAL4-UAS system⁵¹ and an inducible RNAi line.^{39,51} The efficacy of the *row* construct was confirmed by qRT-PCR upon ubiquitous knockdown with the *actin-Gal4* promoter line. The relative mRNA expression compared to that of the appropriate genetic background control (the promoter line crossed to the isogenic background of the RNAi line) was 25% ($p = 0.0003$; Student's *t* test; Table S9).

In order to assess the consequences of *row* knockdown in neurons, we crossed the inducible *row* RNAi line to the panneuronal *elav-Gal4* promoter line.^{40,41} The isogenic host strain of the RNAi construct crossed to the same promoter line served as genetic background controls. We assessed habituation, a non-associative form of learning and behavioral plasticity, in the light-off jump habituation paradigm as previously described.⁴² In this paradigm, the strong initial reaction to a non-threatening stimulus (light-off) gradually wanes. Established *Drosophila* learning and memory mutants, as well as several *Drosophila* ID-associated gene models, have previously been shown to exhibit deficits in habituation.^{40–42,52} *row*-knockdown flies, as well as their genetic background controls, showed good initial

jump response (70% of flies jumping in the first five trials), indicating an absence of severe neurological defects. Whereas control flies quickly habituated to the repeatedly presented light-off stimuli (mean TTC = 10.2), the plastic behavioral response of *row* knockdown flies was strongly affected. They needed, on average, more than double the number of trials to reach the no-jump criterion ($n = 135$; mean TTC = 21.57; $p = 9.19 \times 10^{-5}$). The average jump response of *row*-knockdown and control flies, as well as the difference between mean TTC values, is depicted in Figure 4. We conclude that partial loss of function of the *POGZ* ortholog *row* in *Drosophila* neurons specifically affects non-associative learning in the light-off jump reflex habituation paradigm.

Discussion

We collected a substantial cohort of 24 individuals with de novo LGD mutations in *POGZ*, enabling us to define the core phenotype. Individuals are characterized by a variable neurodevelopmental or neurobehavioral profile associated

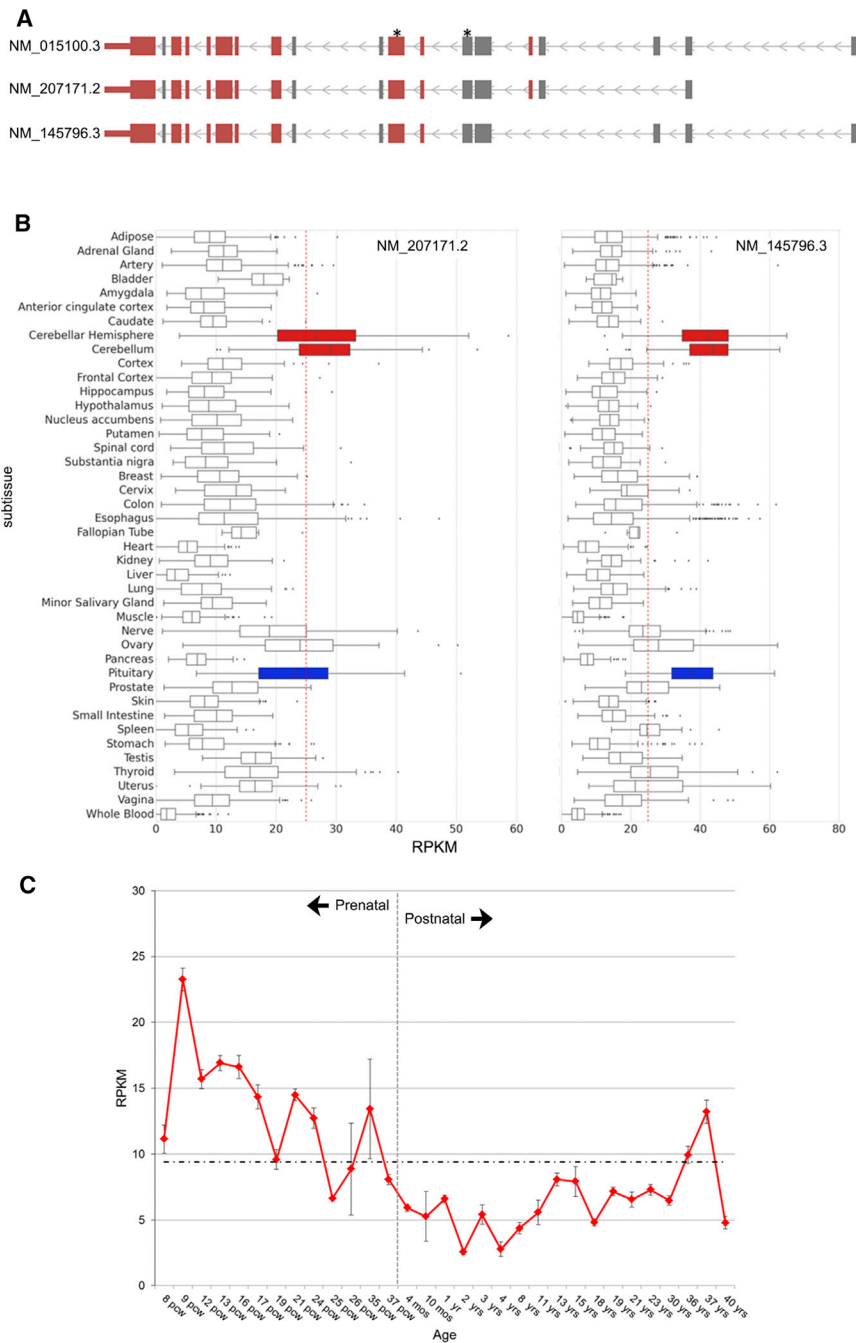


Figure 3. Expression of *POGZ* Is Highest in Cerebellum and Pituitary Tissues

(A) Gene models of three RefSeq *POGZ* isoforms. Isoform 1, the longest isoform, is shown on top (GenBank: NM_015100.3), isoform 2 in the center (GenBank: NM_207171.2), and isoform 3 on the bottom (GenBank: NM_145796.3). Exons are shown as blocks and directionality as light-gray arrows. Exons in red contain a mutation identified in this study. Black stars indicate exons in which disruptive mutations were identified in control individuals.

(B) Expression of *POGZ* shown by isoform (GenBank: NM_207171.2 on the left and GenBank: NM_145796.3 on the right), and subtissue from the GTEx database (v.6).⁴⁵ Cerebellar tissues are shown in red and the pituitary gland is shown in blue. The red dashed line indicates reads per kilobase of transcript per million mapped reads (RPKM) of 25.

(C) *POGZ* average expression of all brain subtissues across brain development (with RNA-seq RPKM values from the BrainSpan Atlas v.10; shown in red from 8 post-conception weeks to 40 years of age). Birth is indicated by a vertical gray dashed line, and the mean expression across all time points is indicated by a horizontal black dashed line.

with other recent reports of smaller cohorts of individuals with *POGZ* mutations.^{29,30}

We repeated the phenotypic analysis, limiting it to those 12 individuals with a diagnosis of autism, and compared it to phenotypic data collected for 2,718 children from the SSC³⁵ ascertained for the presence of an ASD diagnosis. Significant increases in vision problems, hyperactivity, feeding problems, and microcephaly ($p < 0.05$) and a trend toward obesity ($p = 0.07$) were observed in the individuals with a *POGZ* mutation relative to the SSC

comparison cohort (Table S6). As has been observed for other autism risk genes (*CHD8* [MIM: 610528], *DYRK1A* [MIM: 600855], and *ADNP* [MIM: 611386]), these data argue that autistic individuals harboring *POGZ* mutations represent not only a distinct genetic but also a clinical subtype of the condition defined by a gestalt of features. It is interesting that *POGZ* was one of 21 genes recently highlighted as being recurrently mutated in individuals with congenital heart disease (CHD [MIM: 600001]) and neurodevelopmental disorders.⁵³ Homsy et al. considered the fact that *POGZ* mutations were also observed in cohorts of individuals with ASD and/or ID as strong evidence that CHD and neurodevelopmental disorders share a

with ID and/or either autistic features or a formal diagnosis of ASD. As expected, all individuals showed evidence of DD ranging from mild to severe. Language and speech delay presented more prominently than motor delay and coordination problems. 50% of the individuals reported here received a formal diagnosis of ASD; however, several individuals showed an atypical behavioral phenotype including overly friendly and overly social behavior and hyperactivity. Other significant phenotypic features were specific feeding difficulties that seemed to be related to oversensitivity of the oral region in several individuals, a tendency to obesity, nonspecific vision problems, and microcephaly. These clinical findings are also consistent

with ID and/or either autistic features or a formal diagnosis of ASD. As expected, all individuals showed evidence of DD ranging from mild to severe. Language and speech delay presented more prominently than motor delay and coordination problems. 50% of the individuals reported here received a formal diagnosis of ASD; however, several individuals showed an atypical behavioral phenotype including overly friendly and overly social behavior and hyperactivity. Other significant phenotypic features were specific feeding difficulties that seemed to be related to oversensitivity of the oral region in several individuals, a tendency to obesity, nonspecific vision problems, and microcephaly. These clinical findings are also consistent

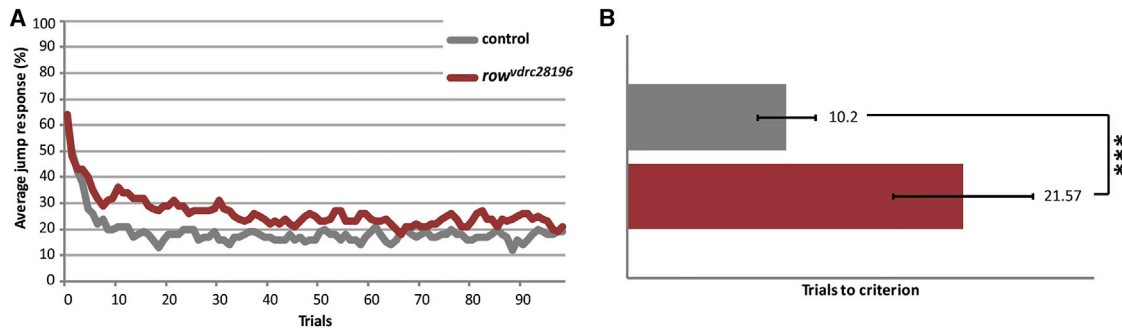


Figure 4. Knockdown of the *Drosophila* POGZ Ortholog *row* Affects Non-Associative Learning in the Light-Off Jump Reflex Habituation Paradigm

Jump responses were induced by repeated light-off pulses (100 trials) with a 1 s inter-trial interval. *row* knockdown flies (*row*^{vdr28196}; genotype: *w*¹¹¹⁸; 2xGMR-wIR/+; UAS-*row*^{vdr28196}/elav-Gal4, UAS-Dicer-2) are plotted in red, and genetic background control flies (control; genotype: *w*¹¹¹⁸; 2xGMR-wIR/+; elav-Gal4, UAS-Dicer-2/+) are plotted in dark gray. Linear regression model analysis revealed that flies with panneuronally-induced *row* knockdown habituated significantly slower (fold-change = 2.1; ****p* < 0.001).

(A) Average jump response (% of jumping flies) across 100 light-off trials.

(B) Mean TTC of *row*^{vdr28196} (TTC = 21.57, *n* = 135) versus mean TTC of control flies (TTC = 10.2, *n* = 107). Error bars indicate SEM.

common genetic etiology. Two de novo variants were identified in the CHD study, including one early splice and one missense event (Figure 1 and Table S3). In our cohort, only one of the 25 individuals (FR1) with an LGD mutation showed evidence of CHD (see Supplemental Note), and subsequent clinical follow-up of seven individuals for this phenotype provided no evidence of CHD. Interestingly, individual FR1 represented one of two splice mutations in our cohort of individuals with POGZ mutations. Although pleiotropic effects remain a possibility, an alternate hypothesis is that POGZ associates with the neurodevelopmental disease aspect and that CHD is caused by some additional risk factor in the CHD exome-sequencing-study individuals, given that many suffered from both neurodevelopmental delay and CHD.

In this study, we report three instances of the same de novo LGD mutation found in two unrelated individuals (at amino acid 674 in two individuals, at amino acid 1001 in two individuals, and at amino acid 1154 in two individuals). These events have not been reported previously in control samples or other studies.^{29,30,53} In some cases (particularly at amino acid 1154), these individuals show similar phenotypes, including behavioral deficits, but at the other two sites, the pairs of individuals do not appear to be more clinically similar. These data suggest that variable expressivity overrides allelic heterogeneity with respect to phenotypic manifestation, possibly as a result of differences in the genetic background or environmental effects. Nevertheless, the identification of identical de novo events is rare among unrelated affected probands.¹⁰ The presence of recurrent events in this study suggests that there might be a shared mechanism by which these mutations occur which will require further study to elucidate. For example, the mutation at amino acid 674 represents a CpG dinucleotide, a known hotspot for de novo mutation.⁵⁴

We find that POGZ is constitutively expressed across most tissues and has significantly higher levels of expres-

sion in the cerebellum and the pituitary gland (Figure 3 and Tables S7 and S8). Also, in *Drosophila*, *row* mRNA expression is observed across most tissues, and the strongest signal is in the CNS. The cerebellum is known to regulate motor control and some cognitive functions and has been implicated in the biology of ASD as well as related conditions, such as attention deficit hyperactive disorder (ADHD [MIM: 143465]).⁵⁵ Among individuals with POGZ mutations, we observed marked language deficits, delayed motor development, and lack of coordination, as well as hyperactivity in some individuals, consistent with the intimate association of the cerebellum and cerebral cortex with respect to ASD pathology.⁵⁶ The higher level of expression in the pituitary also warrants further investigation in light of the findings of obesity, feeding problems, and vision deficits in these individuals. Hypothalamic-pituitary structural lesions in pediatric individuals, for example, are strongly associated with endocrine dysfunction, neuro-ophthalmic presentation, and abnormal BMI and growth velocity.⁵⁷ Notably, we did not observe any individual in whom both feeding problems and obesity were present concurrently (Tables 1 and 2). Considering the potential importance of the pituitary, a future area of investigation might be to measure growth hormone production in individuals with POGZ mutations.

It is interesting that many of the potentially pathogenic mutations fall within the second half of the protein, suggesting that the protein domains located here might be responsible for some of the shared phenotypic features. Although little is known about the function of POGZ in vivo, two functional domains have been reported in the literature to lie within this region. The HP1-binding zinc-finger-like motif (amino acids 791–850) is thought to be integral for the binding of POGZ to heterochromatin protein 1 (CBX5, or HP1) for proper mitotic progression.³¹ When POGZ is knocked down in human cell lines, proper cell growth and division is abrogated.³¹ A second protein domain, the transposase-derived DDE domain (amino

acids 1117–1323) near the C-terminal end of the protein, is known to interact directly with PSIP1 (a.k.a., LEDGF, or p75),⁵⁸ which is known to have human expression in the fetal brain and is thought to be a transcriptional co-activator involved in neurogenesis.⁵⁹ Seizures were also a comorbidity reported in three individuals with *POGZ* mutations. These cases represent three of the four earliest truncating events in the protein that would effectively lack all of the putative functional domains.

In addition, we studied the effect of *POGZ* loss of function in a *Drosophila* model. Although human *POGZ* and *Drosophila row* share low overall sequence conservation at the protein level (10% identity), the central zinc finger motifs are well conserved (25% identity). Given the evolutionary divergence between vertebrates and invertebrates, such divergence is not unexpected among orthologs but does limit functional extrapolations given that these differences could have had effects on protein function. The proteins fulfill the reciprocal best BLAST hit criterion and are annotated orthologs on the Ensembl and Treefam databases. Moreover, human *POGZ* and *Drosophila row* show conserved protein interactions, such as CBX5/HP1,³¹ ZMYM4/Wok,⁶⁰ and NFX1/Nfx2⁶¹ (human/*Drosophila* protein symbol). Targeting *Drosophila row* revealed defects in habituation, a form of non-associative learning. This learning defect is not only relevant to the human ID phenotype of individuals with *POGZ* mutations, but also interesting in light of the co-occurring ASD. Habituation has been recently proposed as a potential mechanism underlying deficits in predictive coding in ASD, and the ability to suppress the response to known irrelevant sensory stimuli might be required for the prevention of overstimulation and salience in ASD individuals.⁶²

In summary, we conclude that *POGZ* adds to the large number of proteins implicated in ID and ASD etiology present in protein complexes that act by modifying chromatin structure and gene regulation.³² The phenotypic similarities of the individuals also predict a clinical subtype of ASD and/or ID distinct from that of individuals more generally diagnosed with autism. The enrichment of specific features but variability of phenotypic presentation is reminiscent of other mutations, especially large CNVs, associated with ASD and/or ID.⁴ We estimate that de novo LGD mutations of *POGZ* might describe up to 0.14% of individuals with undefined ASD and/or ID. Based on our ascertainment, we conclude that the frequency of de novo LGD mutations is 3- to 4-fold higher in ID and/or DD than in ASD, which might explain some of the atypical ASD cases observed in this study. Indeed, ID is a comorbidity often associated with ASD (present in 31% of the SSC individuals);³⁵ all of the individuals with *POGZ* mutations in this study showed borderline-moderate ID, even those recruited on a primary ASD diagnosis (Table 2). Further identification of individuals with disruptive de novo mutations in *POGZ* will be required to determine whether mutations in this gene also lead to ASD in the range of normal IQ.

Supplemental Data

Supplemental Data include Acknowledgments, a Supplemental Note, two figures, and nine tables and can be found with this article online at <http://dx.doi.org/10.1016/j.ajhg.2016.02.004>.

Conflicts of Interest

E.E.E. is on the scientific advisory board of DNAnexus and is a consultant for Kunming University of Science and Technology as part of the 1000 China Talent Program.

Received: December 11, 2015

Accepted: February 5, 2016

Published: March 3, 2016

Web Resources

The URLs for data presented herein are as follows:

BrainSpan Atlas of the Developing Human Brain, <http://www.brainspan.org/>
Ensembl, <http://uswest.ensembl.org/index.html>
ExAC Browser, <http://exac.broadinstitute.org/>
GTEx Portal, <http://www.gtexportal.org/home/>
Human Protein Reference Database, <http://www.hprd.org/>
NCBI Gene, <http://www.ncbi.nlm.nih.gov/gene>
NHLBI Exome Sequencing Project (ESP) Exome Variant Server, <http://evs.gs.washington.edu/EVS/>
OMIM, <http://www.omim.org/>
RefSeq, <http://www.ncbi.nlm.nih.gov/RefSeq>
The R Project for Statistical Computing, <http://www.R-project.org/>
UniProt, <http://www.uniprot.org/>

References

1. Srivastava, A.K., and Schwartz, C.E. (2014). Intellectual disability and autism spectrum disorders: causal genes and molecular mechanisms. *Neurosci. Biobehav. Rev.* **46**, 161–174.
2. van Bokhoven, H. (2011). Genetic and epigenetic networks in intellectual disabilities. *Annu. Rev. Genet.* **45**, 81–104.
3. Matson, J.L., Rivet, T.T., Fodstad, J.C., Dempsey, T., and Boisjoli, J.A. (2009). Examination of adaptive behavior differences in adults with autism spectrum disorders and intellectual disability. *Res. Dev. Disabil.* **30**, 1317–1325.
4. Mefford, H.C., Batshaw, M.L., and Hoffman, E.P. (2012). Genomics, intellectual disability, and autism. *N. Engl. J. Med.* **366**, 733–743.
5. Lai, M.C., Lombardo, M.V., and Baron-Cohen, S. (2014). Autism. *Lancet* **383**, 896–910.
6. Kochinke, K., Zweier, C., Nijhof, B., Fenckova, M., Cizek, P., Honti, F., Keerthikumar, S., Oortveld, M.A., Kleefstra, T., Kramer, J.M., et al. (2016). Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am. J. Hum. Genet.* **98**, 149–164.
7. Xu, L.M., Li, J.R., Huang, Y., Zhao, M., Tang, X., and Wei, L. (2012). AutismKB: an evidence-based knowledgebase of autism genetics. *Nucleic Acids Res.* **40**, D1016–D1022.
8. de Ligt, J., Willemsen, M.H., van Bon, B.W., Kleefstra, T., Yntema, H.G., Kroes, T., Vulto-van Silfhout, A.T., Koolen, D.A., de Vries, P., Gilissen, C., et al. (2012). Diagnostic exome

- sequencing in persons with severe intellectual disability. *N. Engl. J. Med.* 367, 1921–1929.
9. Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Ende, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., Di Donato, N., et al. (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet* 380, 1674–1682.
 10. Iossifov, I., O’Roak, B.J., Sanders, S.J., Ronemus, M., Krumm, N., Levy, D., Stessman, H.A., Witherspoon, K.T., Vives, L., Patterson, K.E., et al. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature* 515, 216–221.
 11. De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S., et al.; DDD Study; Homozygosity Mapping Collaborative for Autism; UK10K Consortium (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* 515, 209–215.
 12. Gilissen, C., Hehir-Kwa, J.Y., Thung, D.T., van de Vorst, M., van Bon, B.W., Willemsen, M.H., Kwint, M., Janssen, I.M., Hoischen, A., Schenck, A., et al. (2014). Genome sequencing identifies major causes of severe intellectual disability. *Nature* 511, 344–347.
 13. Bernier, R., Golzio, C., Xiong, B., Stessman, H.A., Coe, B.P., Penn, O., Witherspoon, K., Gerds, J., Baker, C., Vulto-van Silfhout, A.T., et al. (2014). Disruptive CHD8 mutations define a subtype of autism early in development. *Cell* 158, 263–276.
 14. Stessman, H.A., Bernier, R., and Eichler, E.E. (2014). A genotype-first approach to defining the subtypes of a complex disease. *Cell* 156, 872–877.
 15. van Bon, B.W., Coe, B.P., Bernier, R., Green, C., Gerds, J., Witherspoon, K., Kleefstra, T., Willemsen, M.H., Kumar, R., Bosco, P., et al. (2016). Disruptive de novo mutations of DYRK1A lead to a syndromic form of autism and ID. *Mol. Psychiatry* 21, 126–132.
 16. Helsmoortel, C., Vulto-van Silfhout, A.T., Coe, B.P., Vandeweyer, G., Rooms, L., van den Ende, J., Schuurs-Hoeijmakers, J.H., Marcelis, C.L., Willemsen, M.H., Vissers, L.E., et al. (2014). A SWI/SNF-related autism syndrome caused by de novo mutations in ADNP. *Nat. Genet.* 46, 380–384.
 17. Mullegama, S.V., Alaimo, J.T., Chen, L., and Elsea, S.H. (2015). Phenotypic and molecular convergence of 2q23.1 deletion syndrome with other neurodevelopmental syndromes associated with autism spectrum disorder. *Int. J. Mol. Sci.* 16, 7627–7643.
 18. Giannandrea, M., Bianchi, V., Mignogna, M.L., Sirri, A., Carrabino, S., D’Elia, E., Vecellio, M., Russo, S., Cogliati, F., Larizza, L., et al. (2010). Mutations in the small GTPase gene RAB39B are responsible for X-linked mental retardation associated with autism, epilepsy, and macrocephaly. *Am. J. Hum. Genet.* 86, 185–195.
 19. Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.H., Narzisi, G., Leotta, A., et al. (2012). De novo gene disruptions in children on the autistic spectrum. *Neuron* 74, 285–299.
 20. Neale, B.M., Kou, Y., Liu, L., Ma’ayan, A., Samocha, K.E., Sabo, A., Lin, C.F., Stevens, C., Wang, L.S., Makarov, V., et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* 485, 242–245.
 21. Fromer, M., Pocklington, A.J., Kavanagh, D.H., Williams, H.J., Dwyer, S., Gormley, P., Georgieva, L., Rees, E., Palta, P., Ruderfer, D.M., et al. (2014). De novo mutations in schizophrenia implicate synaptic networks. *Nature* 506, 179–184.
 22. Fukai, R., Hiraki, Y., Yofune, H., Tsurusaki, Y., Nakashima, M., Saitsu, H., Tanaka, F., Miyake, N., and Matsumoto, N. (2015). A case of autism spectrum disorder arising from a de novo missense mutation in POGZ. *J. Hum. Genet.* 60, 277–279.
 23. Buxbaum, J.D., Daly, M.J., Devlin, B., Lehner, T., Roeder, K., and State, M.W.; Autism Sequencing Consortium (2012). The autism sequencing consortium: large-scale, high-throughput sequencing in autism spectrum disorders. *Neuron* 76, 1052–1056.
 24. Gulsuner, S., Walsh, T., Watts, A.C., Lee, M.K., Thornton, A.M., Casadei, S., Rippey, C., Shahin, H., Nimgaonkar, V.L., Go, R.C., et al.; Consortium on the Genetics of Schizophrenia (COGS); PAARTNERS Study Group (2013). Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. *Cell* 154, 518–529.
 25. Deciphering Developmental Disorders, S.; Deciphering Developmental Disorders Study (2015). Large-scale discovery of novel genetic causes of developmental disorders. *Nature* 519, 223–228.
 26. Hashimoto, R., Nakazawa, T., Tsurusaki, Y., Yasuda, Y., Nagayasu, K., Matsumura, K., Kawashima, H., Yamamori, H., Fujimoto, M., Ohi, K., et al. (2015). Whole-exome sequencing and neurite outgrowth analysis in autism spectrum disorder. *J. Hum. Genet.* Published online November 19, 2015. <http://dx.doi.org/10.1038/jhg.2015.141>.
 27. Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.X., et al. (2015). Excess of rare, inherited truncating mutations in autism. *Nat. Genet.* 47, 582–588.
 28. Sanders, S.J., He, X., Willsey, A.J., Ercan-Sencicek, A.G., Samocha, K.E., Cicek, A.E., Murtha, M.T., Bal, V.H., Bishop, S.L., Dong, S., et al.; Autism Sequencing Consortium (2015). Insights into Autism Spectrum Disorder Genomic Architecture and Biology from 71 Risk Loci. *Neuron* 87, 1215–1233.
 29. White, J., Beck, C.R., Harel, T., Posey, J.E., Jhangiani, S.N., Tang, S., Farwell, K.D., Powis, Z., Mendelsohn, N.J., Baker, J.A., et al. (2016). POGZ truncating alleles cause syndromic intellectual disability. *Genome Med.* 8, 3.
 30. Ye, Y., Cho, M.T., Retterer, K., Alexander, N., Ben-Omran, T., Al-Mureikhi, M., Cristian, I., Wheeler, P., Crain, C., Zand, D., et al. (2015). De novo POGZ mutations are associated with neurodevelopmental disorders and microcephaly. *Cold Spring Harb. Mol. Case Stud.* 1, a000455.
 31. Nozawa, R.S., Nagao, K., Masuda, H.T., Iwasaki, O., Hirota, T., Nozaki, N., Kimura, H., and Obuse, C. (2010). Human POGZ modulates dissociation of HP1alpha from mitotic chromosome arms through Aurora B activation. *Nat. Cell Biol.* 12, 719–727.
 32. Hormozdiari, F., Penn, O., Borenstein, E., and Eichler, E.E. (2015). The discovery of integrated gene networks for autism and related disorders. *Genome Res.* 25, 142–154.
 33. Isidor, B., Kury, S., Rosenfeld, J.A., Besnard, T., Schmitt, S., Joss, S., Davies, S.J., Lebel, R.R., Henderson, A., Schaaf, C.P., et al. (2016). De novo truncating mutations in the kinetochore-microtubules attachment gene CHAMP1 cause syndromic intellectual disability. *Hum. Mutat.* Published online January 11, 2016. <http://dx.doi.org/10.1002/humu.22952>.
 34. O’Roak, B.J., Stessman, H.A., Boyle, E.A., Witherspoon, K.T., Martin, B., Lee, C., Vives, L., Baker, C., Hiatt, J.B., Nickerson, D.A., et al. (2014). Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat. Commun.* 5, 5595.
 35. Fischbach, G.D., and Lord, C. (2010). The Simons Simplex Collection: a resource for identification of autism genetic risk factors. *Neuron* 68, 192–195.

36. Buxbaum, J.D., Bolshakova, N., Brownfeld, J.M., Anney, R.J., Bender, P., Bernier, R., Cook, E.H., Coon, H., Cuccaro, M., Freitag, C.M., et al. (2014). The Autism Simplex Collection: an international, expertly phenotyped autism sample for genetic and phenotypic analyses. *Mol. Autism* 5, 34.
37. Samocha, K.E., Robinson, E.B., Sanders, S.J., Stevens, C., Sabo, A., McGrath, L.M., Kosmicki, J.A., Rehnström, K., Mallick, S., Kirby, A., et al. (2014). A framework for the interpretation of de novo mutation in human disease. *Nat. Genet.* 46, 944–950.
38. Rosenfeld, J.A., Coe, B.P., Eichler, E.E., Cuckle, H., and Shaffer, L.G. (2013). Estimates of penetrance for recurrent pathogenic copy-number variations. *Genet. Med.* 15, 478–481.
39. Dietzl, G., Chen, D., Schnorrer, F., Su, K.C., Barinova, Y., Fellner, M., Gasser, B., Kinsey, K., Oettel, S., Scheiblauer, S., et al. (2007). A genome-wide transgenic RNAi library for conditional gene inactivation in *Drosophila*. *Nature* 448, 151–156.
40. Willemsen, M.H., Nijhof, B., Fenckova, M., Nillesen, W.M., Bongers, E.M., Castells-Nobau, A., Asztalos, L., Viragh, E., van Bon, B.W., Tezel, E., et al. (2013). GATAD2B loss-of-function mutations cause a recognisable syndrome with intellectual disability and are associated with learning deficits and synaptic undergrowth in *Drosophila*. *J. Med. Genet.* 50, 507–514.
41. van Bon, B.W., Oortveld, M.A., Nijtmans, L.G., Fenckova, M., Nijhof, B., Besseling, J., Vos, M., Kramer, J.M., de Leeuw, N., Castells-Nobau, A., et al. (2013). CEP89 is required for mitochondrial metabolism and neuronal function in man and fly. *Hum. Mol. Genet.* 22, 3138–3151.
42. Kramer, J.M., Kochinke, K., Oortveld, M.A., Marks, H., Kramer, D., de Jong, E.K., Asztalos, Z., Westwood, J.T., Stunnenberg, H.G., Sokolowski, M.B., et al. (2011). Epigenetic regulation of learning and memory by *Drosophila* EHMT/G9a. *PLoS Biol.* 9, e1000569.
43. Coe, B.P., Witherspoon, K., Rosenfeld, J.A., van Bon, B.W., Vulto-van Silfhout, A.T., Bosco, P., Friend, K.L., Baker, C., Buono, S., Vissers, L.E., et al. (2014). Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat. Genet.* 46, 1063–1071.
44. Achenbach, T.M., Howell, C.T., Quay, H.C., and Conners, C.K. (1991). National survey of problems and competencies among four- to sixteen-year-olds: parents' reports for normative and clinical samples. *Monogr. Soc. Res. Child Dev.* 56, 1–131.
45. GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* 348, 648–660.
46. van der Voet, M., Nijhof, B., Oortveld, M.A., and Schenck, A. (2014). *Drosophila* models of early onset cognitive disorders and their clinical applications. *Neurosci. Biobehav. Rev.* 46, 326–342.
47. Flicek, P., Ahmed, I., Amode, M.R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fairley, S., et al. (2013). Ensembl 2013. *Nucleic Acids Res.* 41, D48–D55.
48. Ruan, J., Li, H., Chen, Z., Coghlan, A., Coin, L.J., Guo, Y., Hériché, J.K., Hu, Y., Kristiansen, K., Li, R., et al. (2008). TreeFam: 2008 Update. *Nucleic Acids Res.* 36, D735–D740.
49. Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landonlin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2011). The developmental transcriptome of *Drosophila melanogaster*. *Nature* 471, 473–479.
50. Chintapalli, V.R., Wang, J., and Dow, J.A. (2007). Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat. Genet.* 39, 715–720.
51. Brand, A.H., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* 118, 401–415.
52. Engel, J.E., and Wu, C.F. (1996). Altered habituation of an identified escape circuit in *Drosophila* memory mutants. *J. Neurosci.* 16, 3486–3499.
53. Homsy, J., Zaidi, S., Shen, Y., Ware, J.S., Samocha, K.E., Karczewski, K.J., DePalma, S.R., McKean, D., Wakimoto, H., Gorham, J., et al. (2015). De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science* 350, 1262–1266.
54. Francioli, L.C., Polak, P.P., Koren, A., Menelaou, A., Chun, S., Renkens, I., van Duijn, C.M., Swertz, M., Wijmenga, C., van Ommen, G., et al.; Genome of the Netherlands Consortium (2015). Genome-wide patterns and properties of de novo mutations in humans. *Nat. Genet.* 47, 822–826.
55. Gowen, E., and Miall, R.C. (2007). The cerebellum and motor dysfunction in neuropsychiatric disorders. *Cerebellum* 6, 268–279.
56. D'Mello, A.M., and Stoodley, C.J. (2015). Cerebro-cerebellar circuits in autism spectrum disorder. *Front. Neurosci.* 9, 408.
57. Taylor, M., Couto-Silva, A.C., Adan, L., Trivin, C., Sainte-Rose, C., Zerah, M., Valteau-Couanet, D., Doz, F., Chalumeau, M., and Brauner, R. (2012). Hypothalamic-pituitary lesions in pediatric patients: endocrine symptoms often precede neuro-ophthalmic presenting symptoms. *J. Pediatr.* 161, 855–863.
58. Bartholomeeusen, K., Christ, F., Hendrix, J., Rain, J.C., Emiliani, S., Benarous, R., Debyser, Z., Gijsbers, R., and De Rijck, J. (2009). Lens epithelium-derived growth factor/p75 interacts with the transposase-derived DDE domain of PogZ. *J. Biol. Chem.* 284, 11467–11477.
59. Chylack, L.T., Jr., Fu, L., Mancini, R., Martin-Rehrmann, M.D., Saunders, A.J., Konopka, G., Tian, D., Hedley-Whyte, E.T., Folkner, R.D., and Goldstein, L.E. (2004). Lens epithelium-derived growth factor (LEDGF/p75) expression in fetal and adult human brain. *Exp. Eye Res.* 79, 941–948.
60. Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G., and Mann, M. (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. *Cell* 142, 967–980.
61. Castello, A., Fischer, B., Eichelbaum, K., Horos, R., Beckmann, B.M., Strein, C., Davey, N.E., Humphreys, D.T., Preiss, T., Steinmetz, L.M., et al. (2012). Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 149, 1393–1406.
62. Ramaswami, M. (2014). Network plasticity in adaptive filtering and behavioral habituation. *Neuron* 82, 1216–1229.