

Haploinsufficiency of *KMT2B*, Encoding the Lysine-Specific Histone Methyltransferase 2B, Results in Early-Onset Generalized Dystonia

Michael Zech,^{1,2} Sylvia Boesch,³ Esther M. Maier,⁴ Ingo Borggraefe,⁴ Katharina Vill,⁴ Franco Laccone,⁵ Veronika Pilshofer,⁶ Andres Ceballos-Baumann,^{2,7} Bader Alhaddad,⁸ Riccardo Berutti,⁹ Werner Poewe,³ Tobias B. Haack,^{8,9,10} Bernhard Haslinger,² Tim M. Strom,^{8,9} and Juliane Winkelmann^{1,2,8,11,*}

Early-onset generalized dystonia represents the severest form of dystonia, a hyperkinetic movement disorder defined by involuntary twisting postures. Although frequently transmitted as a single-gene trait, the molecular basis of dystonia remains largely obscure. By whole-exome sequencing a parent-offspring trio in an Austrian kindred affected by non-familial early-onset generalized dystonia, we identified a dominant de novo frameshift mutation, c.6406delC (p.Leu2136Serfs*17), in *KMT2B*, encoding a lysine-specific methyltransferase involved in transcriptional regulation via post-translational modification of histones. Whole-exome-sequencing-based exploration of a further 30 German-Austrian individuals with early-onset generalized dystonia uncovered another three deleterious mutations in *KMT2B*—one de novo nonsense mutation (c.1633C>T [p.Arg545*]), one de novo essential splice-site mutation (c.7050–2A>G [p.Phe2321Serfs*93]), and one inherited nonsense mutation (c.2428C>T [p.Gln810*]) co-segregating with dystonia in a three-generation kindred. Each of the four mutations was predicted to mediate a loss-of-function effect by introducing a premature termination codon. Suggestive of haploinsufficiency, we found significantly decreased total mRNA levels of *KMT2B* in mutant fibroblasts. The phenotype of individuals with *KMT2B* loss-of-function mutations was dominated by childhood lower-limb-onset generalized dystonia, and the family harboring c.2428C>T (p.Gln810*) showed variable expressivity. In most cases, dystonic symptoms were accompanied by heterogeneous non-motor features. Independent support for pathogenicity of the mutations comes from the observation of high rates of dystonic presentations in *KMT2B*-involving microdeletion syndromes. Our findings thus establish generalized dystonia as the human phenotype associated with haploinsufficiency of *KMT2B*. Moreover, we provide evidence for a causative role of disordered histone modification, chromatin states, and transcriptional deregulation in dystonia pathogenesis.

Dystonia is a motor disorder characterized by uncontrolled hyperkinetic muscle contractions resulting in twisting movements and disabling postures.^{1,2} In generalized manifestations, dystonic symptoms affect the trunk and at least two additional body regions and can evolve either in isolation (referred to as generalized isolated dystonia) or in combination with other movement abnormalities (referred to as generalized combined dystonia).^{1,2} Still, no mechanism-based treatment options are available for the vast majority of dystonic syndromes, mostly because of the fragmentary knowledge about their underlying pathophysiology.^{1,2} For generalized dystonia, only a few genes harboring causative mutations have been described, explaining a small proportion of cases.^{1,3,4}

In this study, we employed whole-exome trio analysis in a family affected by early-onset generalized dystonia and subsequent whole-exome sequencing (WES)-based interrogation of candidate genes in a replication cohort. We reveal four unique loss-of-function (LoF) mutations in *KMT2B* (MIM: 606834; originally named *MLL2*, *MLL4*, or *WBP7*), encoding a lysine-specific histone methyltransferase in chromosomal region 19q13.12. By combining our WES re-

sults with *KMT2B* expression studies in LoF mutation carriers and manual review of phenotypic effects in 19q13.12 contiguous gene-deletion syndromes, we have identified *KMT2B* haploinsufficiency as a highly penetrant risk factor for early-onset generalized dystonia in humans and define a role for defective histone modification in dystonia pathogenesis.

Probands of the study were enrolled from multiple recruitment sites including Innsbruck (11 individuals, including the index subject of family F1), Munich (18 individuals, including the probands of families F2 and F3), and Vienna (four individuals, including the probands of family F4). Biological samples were collected after written informed consent was given, and the study was approved by the local ethics committees. We revisited a simplex Austrian index subject with severe early-onset generalized dystonia (F1-II-5 in family F1; Figure 1; Table 1), in whom previous WES had not identified mutations of known dystonia-causing genes.⁵ To allow the detection of recessive or de novo dominant variants in a yet-undiscovered etiologically involved gene, we chose to complement the initial proband-only exome approach by performing full

¹Institut für Neurogenomik, Helmholtz Zentrum München, 85764 Munich, Germany; ²Klinik und Poliklinik für Neurologie, Klinikum rechts der Isar, Technische Universität München, 81675 Munich, Germany; ³Department of Neurology, Medical University Innsbruck, 6020 Innsbruck, Austria; ⁴Dr. von Haunersches Kinderspital, Ludwig-Maximilians-Universität München, 80337 Munich, Germany; ⁵Institute of Medical Genetics, Medical School of Vienna, 1090 Vienna, Austria; ⁶Krankenhaus der Barmherzigen Schwestern Linz, 4020 Linz, Austria; ⁷Schön Klinik München Schwabing, 80804 Munich, Germany; ⁸Institut für Humangenetik, Technische Universität München, 81675 Munich, Germany; ⁹Institut für Humangenetik, Helmholtz Zentrum München, 85764 Munich, Germany; ¹⁰Division of Molecular Genetics, Universitätsklinikum Tübingen, 72076 Tübingen, Germany; ¹¹Munich Cluster for Systems Neurology, SyNergy, 81377 Munich, Germany

*Correspondence: winkelmann@lrz.tu-muenchen.de

<http://dx.doi.org/10.1016/j.ajhg.2016.10.010>

© 2016 American Society of Human Genetics.

proband-parent trio WES. As described before for the index subject,⁵ exomic sequences of both unaffected parents were captured and processed at the Helmholtz Center Munich in Germany according to fully validated protocols. In short, blood-cell-derived genomic DNA libraries were enriched with the SureSelect Human All Exon 50 Mb Kit v.5 (Agilent Technologies) and sequenced on a HiSeq 2500 machine (Illumina) to an average read depth of 181× (Table S1). Reads were mapped to the human reference genome (UCSC Genome Browser hg19) with the Burrows-Wheeler Aligner (v.0.6.2.). SAMtools (v.0.1.18), PINDEL (v.0.2.4t), ExomeDepth (v.1.0.0), and Custom Perl scripts were employed for variant detection and annotation on all sequenced family members simultaneously. For inclusion in downstream analyses, variant calls were required to display a minimum read depth of 10× and a minimum quality score of 30 (defined as high-confidence calls). Considering the single occurrence of disease in family F1, bioinformatics filtering of variants was based on recessive and de novo dominant inheritance patterns. All retained candidate variants were verified by Sanger sequencing and tested for co-segregation in F1. Prioritization of recessive protein-altering variants with a minor allele frequency ≤ 0.001 in the Exome Aggregation Consortium (ExAC) Browser (v.0.3.1) or an internal database containing 7,900 control exomes yielded two co-segregating alterations that were deemed unlikely to be causative for the observed dystonia phenotype (Table S2). Regarding a de novo dominant effect, we next searched the index subject's WES data for protein-altering sequence variations that were absent from (1) the ExAC Browser, (2) our internal control exome database, and (3) the exome variant profiles of each parent. After this process and confirmatory Sanger evaluation, three heterozygous de novo events were left: two missense substitutions (c.32G>T [p.Gly11Val] in *TDRD6* [MIM: 611200] and c.2483G>A [p.Gly828Glu] in *NYNRIN*) and one LoF variant (c.6406delC [p.Leu2136Serfs*17] in *KMT2B* [GenBank: NM_014727.2 and NP_055542.1]) (Table S3). By introducing a frameshift and a premature translation stop at amino acid position 2,152, the *KMT2B* single-nucleotide deletion was the only identified de novo change predicted to have a severe impact on protein structure (Figure 1). Further, although we found evidence of large numbers of singleton missense variants in *TDRD6* and *NYNRIN* among population-based control individuals (ExAC-derived missense Z scores of -3.35 and 0.43 , respectively),⁶ we noted a highly restricted repertoire of *KMT2B* LoF variants in the ExAC Browser (probability of being LoF intolerant [pLI] score of 1.0)⁶ (Table S4). Similarly, we observed no single high-confidence *KMT2B* LoF variant call in 7,900 in-house control exomes, rendering *KMT2B* the most promising candidate for further evaluation.

To determine whether mutations in *TDRD6*, *NYNRIN*, or *KMT2B* were present in other individuals with early-onset generalized dystonia, we extended our genetic investiga-

tion to a German-Austrian replication cohort of 30 individuals with unsolved disease (67% female, mean age at onset = 9.9 ± 9.2 years), including 26 simplex cases (Table S5). According to the dystonia consensus definition,¹ 23 of these individuals were classified as having generalized isolated dystonia, and seven were classified as having generalized combined dystonia (Table S5). In eight individuals, dystonic symptoms were accompanied by variable non-motor features (Table S5). Cerebral MRI was unremarkable across the replication cohort, and medical interviews did not indicate any acquired etiologies. The entire replication cohort had undergone WES at the Helmholtz Center Munich in Germany via the same sequencing methodology and variant-annotation techniques as detailed above (Table S6). Complementary trio WES had been performed in six individuals. In all cases, the WES data were used to exclude pathogenic mutations of known dystonia-related genes, as previously described.⁵ Despite a high depth of coverage over the target regions (Table S7), interrogation of the 30 replication-cohort exomes uncovered no additional rare protein-altering variants in *TDRD6* or *NYNRIN*. By contrast, three separate dystonia-affected individuals were found to harbor heterozygous high-confidence variant calls in *KMT2B*, none of which were registered in the ExAC Browser or our in-house control exome database (collectively encompassing $> 135,000$ sequenced alleles). Strikingly, each of the three unique sequence alterations was predicted to engender a LoF effect (Figure 1; Table 1). In individual RC-16 (simplex case F2-II-1 in family F2; Figure 1; Table 1), we encountered c.1633C>T resulting in a nonsense mutation at amino acid position 545 (p.Arg545*). Individual F2-II-1 had undergone trio WES, directly indicating that c.1633C>T (p.Arg545*) had arisen de novo. Sanger sequencing validated the de novo status of the mutation (Figure 1). This mutation was the only observed de novo change in F2-II-1 with an expected deleterious impact on protein structure (Table S8). Individual RC-26 (simplex case F3-II-3 in family F3; Figure 1; Table 1) was found to possess an A-to-G transition at the conserved AG splice-acceptor dinucleotide of intron 29 (c.7050–2A>G). Direct sequencing of cDNA fragments amplified by RT-PCR from F3-II-3's whole-blood-derived RNA revealed that the mutation produced a complex pattern of incorrect splicing, including exon 29 skipping, cryptic intron 29 splice-acceptor utilization, and partial intron 29 retention (Figure 2). In terms of structure, the misspliced transcript was predicted to contain a shift in the reading frame and a premature stop codon within exon 32 (p.Phe2321Serfs*93) (Figure 2). The c.7050–2A>G mutation was confirmed and proven to be de novo by Sanger sequencing of DNA samples from the subject and his biological parents (Figure 1). Finally, in RC-29 (familial case F4-III-2 in family F4; Figure 1; Table 1), we discovered c.2428C>T leading to a premature termination of the protein product at amino acid position 810 (p.Gln810*). Sanger sequencing verified the variant and showed its co-segregation with dystonia in the

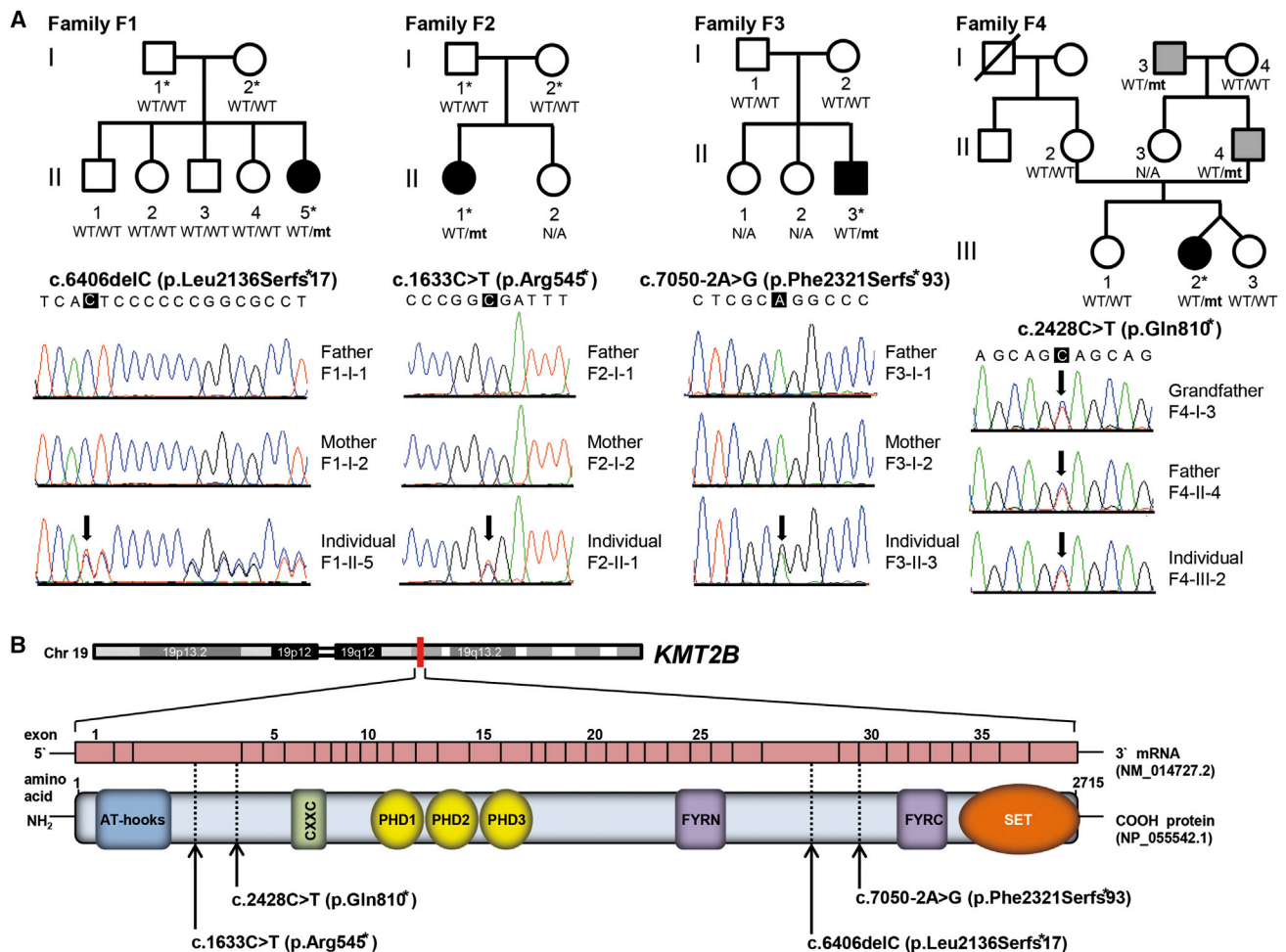


Figure 1. *KMT2B* LoF Mutations in Individuals with Early-Onset Generalized Dystonia

(A) Pedigree drawings and *KMT2B* LoF genotypes in families F1–F4. Filled black symbols represent individuals with early-onset generalized dystonia, filled gray symbols represent individuals with early-onset non-generalized dystonia, open symbols represent unaffected individuals, squares indicate males, and circles indicate females. Black asterisks indicate individuals subjected to WES. The mutation status is given below each individual. Mt indicates the mutant allele, WT indicates the wild-type allele, and N/A indicates an individual who has not been tested for the family mutation in *KMT2B*. Sanger-sequencing electropherograms confirm the de novo status of each LoF variant in families F1–F3 and co-segregation of the LoF variant in family F4. Mutated bases are boxed in black, and variants are indicated by an arrow.

(B) Graphical view of *KMT2B* chromosomal position, transcript structure, and protein product, along with the distribution of four identified LoF variants. *KMT2B* maps to chromosomal region 19q13.2 and encodes a 37-exon canonical transcript (8,148 bp). The translated protein (2,715 amino acids) encompasses an AT-hook DNA-binding domain (in blue), a CXXC zinc-finger domain involved in protein-protein interactions (in green), three plant homodomain (PHD) fingers involved in protein-protein interactions (in yellow), two FY-rich domains (N-terminal, FYRN; C-terminal, FYRC) involved in protein heterodimerization (in violet), and a SET (su[*var*]3-9 enhancer-of-zeste trithorax) domain responsible for histone lysine methylation (in orange). The LoF variants, indicated by arrows, are shown below the protein. The structure of the *KMT2B* transcript and organization of the protein domains are not drawn to exact scale.

proband's father (F4-II-4) and paternal grandfather (F4-I-3) (Figure 1). On the protein level, all identified LoF variants were expected to deleteriously affect the functionally important C-terminal SET domain, conferring methyltransferase activity (Figure 1).

Review of the clinical findings revealed that the index subject from family F1 and the three replication-cohort individuals heterozygous for *KMT2B* LoF mutations shared a strikingly similar phenotype of childhood lower-limb-onset generalized dystonia (Table 1). Among affected relatives of individual F4-III-2, however, dystonia was less severe and non-generalized in distribution (Table 1), sug-

gesting that *KMT2B* LoF genotypes can be associated with variable expressivity. In individuals F2-II-1 and F3-II-3, as well as all affected members of family F4, additional non-motor features were evident, as specified below. Index subject F1-II-5, the fifth child of healthy Austrian parents, had normal psychomotor development. At age 7 years, dystonic inversion of her left foot, rapidly followed by the manifestation of writer's cramp, was noted. In subsequent years, dystonic symptoms progressed to involve other body areas including the face, neck, larynx, tongue, trunk, and all four extremities. At age 23 years, the proband underwent deep brain stimulation of globus pallidus

Table 1. Synopsis of *KMT2B* LoF Genotypes Identified in This Study and Associated Clinical Features

	Family F1	Family F2	Family F3	Family F4		
	F1-II-5 (Index Subject)	F2-II-1 (RC-16)	F3-II-3 (RC-26)	F4-III-2 (RC-29)	F4-II-4 (Father)	F4-I-3 (Grandfather)
Sex	female	female	male	female	male	male
Origin	Austrian	German	German	Austrian	Austrian	Austrian
Age at last examination (years)	31	11	15	6	36	61
<i>KMT2B</i> LoF variant						
Genomic position (hg19)	chr19: 36,223,856	chr19: 36,211,882	chr19: 36,224,662	chr19: 36,212,677	chr19: 36,212,677	chr19: 36,212,677
Exon ^a	28	3	IVS29 ^b	3	3	3
Variation nucleotide ^a	c.6406delC	c.1633C>T	c.7050–2A>G	c.2428C>T	c.2428C>T	c.2428C>T
Variation amino acid ^b	p.Leu2136Serfs*17	p.Arg545*	p.Phe2321Serfs*93	p.Gln810*	p.Gln810*	p.Gln810*
Variant type	frameshift	nonsense	canonical splice	nonsense	nonsense	nonsense
Inheritance	de novo	de novo	de novo	paternal	paternal	unknown
Pregnancy and delivery	normal	normal	normal	normal (gemini)	normal	normal
Postnatal development						
Head circumference	normal	microcephaly (<3 rd percentile)	microcephaly (<1 st percentile)	microcephaly (<1 st percentile)	microcephaly (<3 rd percentile)	normal
Height	normal	short stature (<3 rd percentile)	normal	short stature (<1 st percentile)	short stature (<3 rd percentile)	normal
Weight	normal	short weight (<3 rd percentile)	normal	short weight (<3 rd percentile)	normal	normal
Motor milestones	normal	moderately delayed	delayed	normal	normal	normal
Speech	normal	normal	delayed	delayed	delayed	delayed
Cognition and intellect	normal	normal	mild impairment	mild impairment	mild impairment	mild impairment
Dystonia characteristics						
Age at onset (years)	7	3	11	4	9	11
Site of onset	foot	foot	foot	foot	hand	hand
Distribution at final examination	generalized	generalized	generalized	generalized	focal	focal
Sites affected	lower face, neck, larynx, tongue, trunk, upper extremities, lower extremities	trunk, upper extremities, lower extremities	neck, tongue, trunk, upper extremities, lower extremities	neck, trunk, lower extremities	hand, forearm	hand, forearm
Tremor	no	no	no	no	no	yes (right arm, action induced)

(Continued on next page)

Table 1. Continued						
	Family F1	Family F2	Family F3	Family F4		
	F1-III-5 (Index Subject)	F2-II-1 (RC-16)	F3-II-3 (RC-26)	F4-III-2 (RC-29)	F4-II-4 (Father)	F4-I-3 (Grandfather)
Associated motor features	no	no	no	no	no	no
Additional neurological abnormalities	no	no	no	no	no	no
Levodopa response	no	no	no	pending	not tested	not tested
Treatment	bilateral globus pallidus internus deep brain stimulation	anticholinergics	anticholinergics	levodopa	no	no
Dystonia clinical classification	early-onset generalized dystonia	early-onset generalized dystonia	early-onset generalized dystonia	early-onset generalized dystonia	early-onset focal dystonia	early-onset focal dystonia
Brain MRI	normal	normal	normal	normal	not performed	not performed
Other anomalies	no	strabismus, vesicoureteral reflux	bilateral II-III syndactyly of toes, astigmatism	astigmatism	astigmatism	no
*Numbering is according to GenBank: NM_014727.2 and NP_055542.1 (Ensemble: ENST00000222270 and ENSP00000222270).						
‡VS, intervening sequence (intron).						

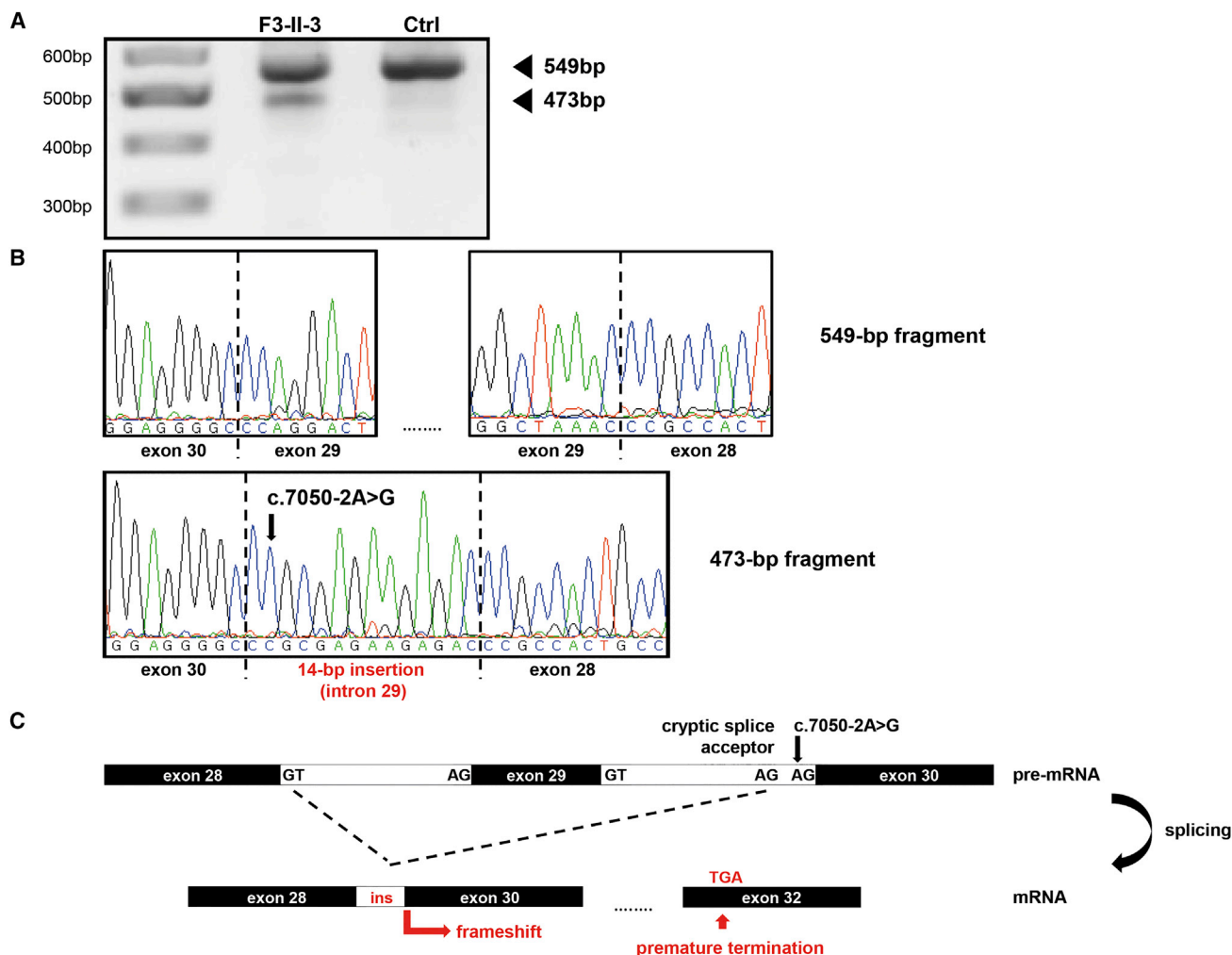


Figure 2. The *KMT2B* c.7050-2A>G Splice-Site Mutation Predictably Results in Protein Truncation

(A) Gel electrophoresis of *KMT2B* cDNA fragments. Total RNA was extracted from whole blood of individual F3-II-3 and a healthy control individual (Ctrl) (QIAGEN) and reverse transcribed into cDNA (Invitrogen). The relevant cDNA fragment was amplified with specific primers in exon 28 and the exon 32-33 junction. Individual F3-II-3 and the control subject produce the expected cDNA fragment of 549 bp. In F3-II-3, the heterozygous splice acceptor AG>GG substitution in intron 29 leads to the accumulation of small amounts of a second cDNA species (473 bp).

(B) Isolation and Sanger sequencing of the 549-bp cDNA fragment confirms a regular transcript structure with intact exon 28-29 and exon 29-30 boundaries. The 473-bp fragment represents a transcript that skips exon 29 and contains the terminal 14 nucleotides of intron 29. The mutant splice-acceptor allele is part of the misspliced transcript, as indicated by an arrow.

(C) Scheme illustrating the complex effect of the *KMT2B* splicing mutation. The mutation at the conserved AG dinucleotide activates a cryptic splice acceptor in intron 29, resulting in a mixture of intronic sequence retention and exon skipping. The missplicing event is predicted to lead to a frameshift and introduction of a premature termination codon in the reading frame of exon 32 (p.Phe2321Serfs*93). Sites joined by splicing are indicated.

was seen in >5,000 non-dystonia parent-offspring trios sequenced as part of diverse disease-related projects⁷⁻¹⁴ and (2) *KMT2B* is severely depleted of disruptive variation in the human population.^{6,15,16} *KMT2B* has a residual variation-intolerance score of -2.37 (1.12th percentile)¹⁶ and ranks among the top 1,003 constraint genes released by the de novo excess algorithm,¹⁵ indicating that it is extremely sensitive to mutational changes. Moreover, application of the ExAC-derived pLI score (1.0) defines *KMT2B* as a gene that is (1) particularly intolerant to LoF variants, (2) dosage sensitive, and (3) predicted to have a high probability of being relevant to haploinsufficient disease.⁶

In support of *KMT2B* haploinsufficiency, each of the identified dystonia-related mutations created a premature termination codon, predictably resulting in mRNA products that are vulnerable to nonsense-mediated decay.¹⁷ In agreement, cDNA from the aberrant transcript associated with c.7050-2A>G (F3-II-3) was detectable at significantly lower levels than the corresponding wild-type species (Figure 2). To test whether *KMT2B* LoF alleles are subject to nonsense-mediated decay, we investigated the c.6406delC (F1-II-5) and c.7050-2A>G (F3-II-3) variants by qRT-PCR in fibroblast cell lines acquired from proband skin biopsies. As diagrammed in Figure 3, the mutations

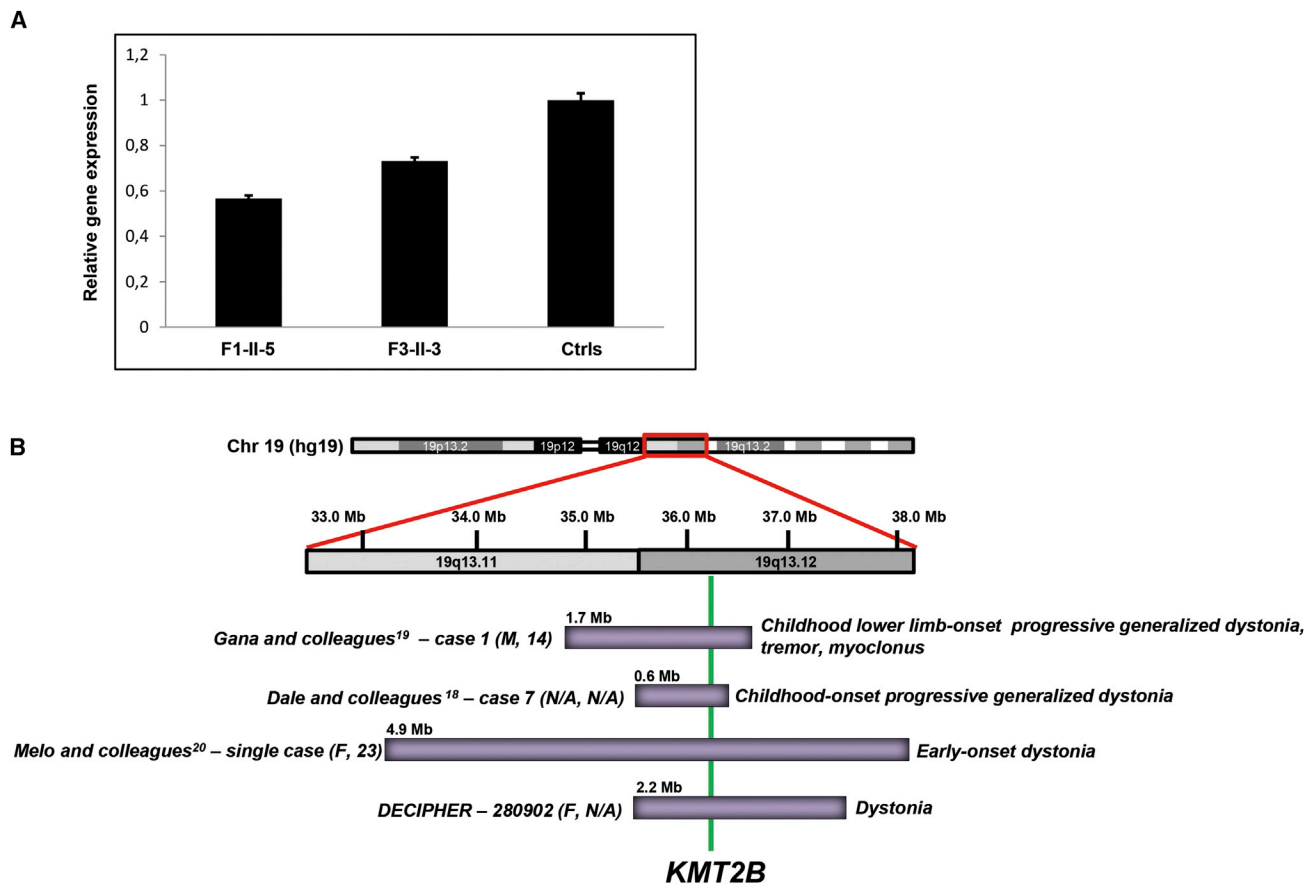


Figure 3. *KMT2B* Haploinsufficiency Causes Dystonia

(A) Total *KMT2B* mRNA levels are reduced in dermal fibroblasts from individuals F1-II-5 and F3-II-3. Total RNA was extracted from fibroblasts of individuals F1-II-5 and F3-II-3, as well as two healthy control individuals (Ctrls) (QIAGEN). Integrity of the RNA was confirmed by the Agilent 2100 Bioanalyzer with the use of RNA 6000 Nano chips (RNA integrity number [RIN] = 10 for each sample). Reverse transcription was conducted with 1,000 ng total RNA as a template (Invitrogen). qRT-PCR was performed with TaqMan gene-expression assays (ThermoFisher) for *KMT2B* and an endogenous reference gene (*GAPDH*). Each sample was run in quadruplicate, and expression levels were determined by the $\Delta\Delta CT$ method. Error bars indicate the SEM of four independent experiments.

(B) Published individuals with de novo *KMT2B*-involving 19q13.1 microdeletions manifest dystonia. A systematic search for original publications and individual clinical reports in PubMed and DECIPHER identified a total of four dystonia-affected individuals who harbor small interstitial de novo deletions of the 19q13.1 chromosomal region, including *KMT2B*. The dystonia-associated deletions (0.6–4.9 Mb in size) are depicted as violet horizontal bars in correlation to an ideogram of chromosome 19 and the relevant interval of 19q13.11–q13.12 (coordinates corresponding to UCSC Genome Browser build hg19). The solid green line extending through the diagram indicates the *KMT2B* locus (chr19: 36,208,921–36,229,781). Sources and individual case identifiers, along with the description of the associated dystonic manifestation, are given as provided. Sexes and ages at last examination (in years) of the individuals carrying the deletions are shown in parentheses. Abbreviations are as follows: M, male; F, female; N/A, not available.

showed a reduction of *KMT2B* mRNA expression to approximately 55%–70% of that of control individuals, compatible with relevant degradation of the mutant transcripts.

To gain independent evidence for a causal role of *KMT2B* mutations in dystonia and establish haploinsufficiency as the underlying pathogenic mechanism, we assessed the phenotypic consequences of DNA structural variations resulting in the loss of one *KMT2B* copy. Systematic analysis of the medical literature led to identification of nine unrelated individuals with chromosomal deletions compromising *KMT2B* in 19q13.12 and ranging between 0.6 and 11 Mb in size.^{18–24} Evaluation of the associated phenotypes revealed that three out of nine individuals with a 19q13.12 deletion (33%) had manifested dystonia,^{18–20}

which was further specified as childhood-onset generalized dystonia in two of them^{18,19} (Figure 3). Like three of the aforementioned LoF mutations, the dystonia-related 19q13.12 deletions had occurred de novo, and none was found to contain any haploinsufficient gene previously linked to dystonia.^{18–20} Interestingly, we noted that the remaining six deletion-affected individuals not reported to have overt dystonia shared several signs of motor dysfunction with the dystonia probands described here, such as motor delay ($n = 6$), progressive gait problems ($n = 3$), speech impairment ($n = 2$), and clumsiness ($n = 2$).^{19,21–24} Given that all of these subjects had been ascertained in early childhood (<10 years), one could speculate that at least some will develop dystonia over time. Together with another dystonia-manifesting individual

with a *KMT2B* deletion listed in the DECIPHER database (Figure 3), our findings strongly suggest that reduction of *KMT2B* dosage is a relevant risk factor for the evolution of dystonia. Moreover, it could well be hypothesized that *KMT2B* haploinsufficiency explains dystonic presentations seen in 19q13.12 microdeletion syndromes. The integration of the whole-gene deletion data into our WES results brings the number of individuals with dystonia and disruption of *KMT2B* to ten, thus confirming *KMT2B* as the gene causally implicated in the subjects' phenotype.

KMT2B encodes a widely²⁵ expressed epigenetic regulator that catalyzes the transfer of methyl groups to the fourth lysine residue of histone H3 (H3K4).^{26–28} According to the Allen Mouse Brain Atlas²⁹ and previous work,^{30,31} *Kmt2b* is abundantly expressed during brain development and detectable at intense levels throughout the adult brain, including areas responsible for motor control. Conveying key roles in embryogenesis, cell identity, and tissue homeostasis, methylation of H3K4 (H3K4me) is a chromatin-modifying event with fundamental roles in gene expression.^{26–28} In the central nervous system (CNS), H3K4me drives the activation of critical target genes during development and maintains proper neural function via modulating DNA accessibility in later periods.^{27,32} Deregulation of H3K4me signatures has been associated with a wide variety of CNS disease states, covering both the neurodevelopmental and neurodegenerative spectra.^{27,28,33,34} To date, germline mutations in at least eight genes encoding H3K4-specific methyltransferases (*KMT2A* [MIM: 159555], *KMT2C* [MIM: 606833], *KMT2D* [MIM: 602113], and *KMT2F* [MIM: 611052]) or demethylases (*KDM1A* [MIM: 609132], *KDM5A* [MIM: 180202], *KDM5B* [MIM: 605393], and *KDM5C* [314690]) have been related to human neuropsychiatric illness.^{27,28,33} Among the *KMT2* family, mutations in *KMT2F* have been linked to schizophrenia (MIM: 181500);³⁵ mutations in *KMT2A* have been linked to Wiedemann-Steiner syndrome (MIM: 605130),³⁶ a rare developmental disorder with neurological features including intellectual disability, microcephaly, and behavioral difficulties; mutations in *KMT2C* have been linked to Kleefstra syndrome (MIM: 610253),³⁷ a class of hereditary intellectual disability disorders; and mutations in *KMT2D* have been linked to Kabuki syndrome (MIM: 147920),³⁸ a congenital multiple-anomaly disease characterized neurologically by intellectual disability, autistic-type behaviors, and seizures. Notably, the majority of mutations in *KMT2* genes are thought to be pathogenic as heterozygous-null alleles, which is in good agreement with the mutational spectrum observed in our group of individuals with *KMT2B* variants.

In mice, global haploinsufficiency of *Kmt2a*, the closest homolog to *Kmt2b*, leads to aberrant synaptic plasticity and robust deficits in motor-task performance.³⁹ Furthermore, conditional deletion of *Kmt2a* from neural precursors of cerebellar granule cells, the hippocampus, and the subventricular zone substantially impairs neurogenesis

and produces ataxic-like movements.⁴⁰ Although they do not demonstrate overt motor phenotypes, mouse models of *Kmt2b* display an equally wide array of neurological complications. Embryos homozygous for *Kmt2b*-null alleles show widespread apoptosis, resulting in incomplete closure of the neural tube and early lethality.³⁰ Consistent with a defect in neural development, immature neurons of *Kmt2b*-null mice exhibit drastically decreased proliferative capacity.³¹ As an explanation, it has been proposed that loss of *Kmt2b* causes perturbation of core transcriptional programs and thereby hampers the establishment of differentiation processes.^{30,31} In an independent mouse model, *Kmt2b* has been selectively ablated from the adult forebrain to reveal transcriptional repression of key players in the control of synaptic function.⁴¹ Specifically, conditional mutagenesis of *Kmt2b* in hippocampal excitatory neurons led to aberrant H3K4me clusters and associated transcriptional downregulation across 169 distinct genes, mostly critical to neuroplasticity.⁴¹ Strikingly, removal of *Kmt2b* was also accompanied by reductions in the expression of genes directly linked to monogenic generalized dystonia, including *PRKRA* (MIM: 603424; associated with dystonia 16 [MIM: 612067]) and *ADCY5* (MIM: 600293; associated with *ADCY5*-related dyskinesia [MIM: 606703]).⁴¹

Dystonia is considered to be a sensorimotor circuit disorder with an important neurodevelopmental component.⁴ Defective establishment of synaptic contacts during embryonic and early postnatal stages might cause altered connectivity at critical nodes of the motor system, ultimately resulting in irregular motor-signal processing and dystonic states.⁴ Fitting well with this concept, active histone remodeling and associated regulation of gene expression have been implicated in a myriad of brain developmental events, including axonal migration and synaptogenesis.^{32,42} In particular, H3K4me modifications are part of an evolutionarily conserved cellular memory system that promotes integrity of the neural network.^{27,42} On the basis of these observations and the *Kmt2b* mouse-model data, it is tempting to speculate that aberrant *KMT2B* could broadly affect neuronal and/or glial transcriptomes and thereby contribute to the formation of erratic motor circuitry. Further, the greatly increased demand for synaptogenic components during development could explain a particularly strong requirement for appropriate *KMT2B* dosage, thus providing the basis for haploinsufficiency as the etiological disease mechanism. Beyond the neurodevelopmental perspective, it is remarkable to consider that reduced *KMT2B* dosage influences transcript levels of known generalized dystonia-causing genes.⁴¹ This finding implies that (1) disruption of these target genes might contribute to the development of dystonia in *KMT2B*-associated disease and (2) the identification of *KMT2B* dosage reduction constitutes a crucial step toward the elucidation of unifying molecular pathways in dystonia. Comprehensive investigation of *KMT2B*-sensitive genes could enhance our understanding of the biological

modules that underpin dystonia and potentially suggest molecular targets for specific treatments.

It remains to be explored to what extent *KMT2B* haploinsufficiency establishes stable epigenetic alterations during a developmental window or whether disordered histone modifications and chromatin states might be amenable to therapeutic interventions in adult life. H3K4me promotes chromatin accessibility, which is antagonized by histone deacetylation. In fact, it has been shown in the case of *Kmt2d* deficiency that postnatal treatment with histone deacetylase inhibitors restores aberrant epigenetic profiles and alleviates functional neural deficits.⁴³ This raises the intriguing possibility of pharmacological manipulation of epigenetic states as a strategy for the treatment of the as-yet-incurable movement disorder dystonia.

In conclusion, we have identified a causal role for dominant LoF mutations in *KMT2B* in individuals with early-onset generalized dystonia and implicate haploinsufficiency as the most likely pathomechanism. Our results add generalized dystonia to the catalog of Mendelian disorders caused by mutations in the epigenetic machinery and highlight the pivotal role of transcriptional regulation through histone modification in the genesis of dystonic movements. Future investigations will help to reveal how the core phenotype of *KMT2B*-associated disease comprises generalized dystonia without systemic signs or, more commonly, includes additional syndromic features. Furthermore, it remains to be seen whether variable expressivity implicates this condition in an even broader spectrum of dystonic manifestations.

Supplemental Data

Supplemental Data include eight tables and can be found with this article online at <http://dx.doi.org/10.1016/j.ajhg.2016.10.010>.

Acknowledgments

We thank all individuals with dystonia and their family members who participated in this study. This study was funded by a research grant from the Else Kröner-Fresenius-Stiftung (2015_A151) as well as in part by in-house institutional funding from Technische Universität München (Munich), Helmholtz Zentrum München (Munich), and Medizinische Universität Innsbruck (Austria). M.Z. receives intramural funding from the Langmatz-Stiftung. T.B.H. was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (grant FKZ 01ZX1405C).

Received: September 26, 2016

Accepted: October 25, 2016

Published: November 10, 2016

Web Resources

Allen Mouse Brain Atlas, <http://www.brain-map.org/>
ClinVar, <http://www.ncbi.nlm.nih.gov/clinvar/>
dbSNP, <http://www.ncbi.nlm.nih.gov/projects/SNP/>

DECIPHER, <https://decipher.sanger.ac.uk/>

Ensembl, <http://www.ensembl.org/index.html>

Exome Aggregation Consortium (ExAC) Browser, <http://exac.broadinstitute.org/>

Genic Intolerance, <http://genic-intolerance.org/>

OMIM, <http://www.omim.org/>

PubMed, <http://www.ncbi.nlm.nih.gov/pubmed>

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

The Human Protein Atlas, <http://www.proteinatlas.org/>

UCSC Genome Browser, <https://genome.ucsc.edu/>

References

1. Albanese, A., Bhatia, K., Bressman, S.B., Delong, M.R., Fahn, S., Fung, V.S., Hallett, M., Jankovic, J., Jinnah, H.A., Klein, C., et al. (2013). Phenomenology and classification of dystonia: a consensus update. *Mov. Disord.* 28, 863–873.
2. Balint, B., and Bhatia, K.P. (2014). Dystonia: an update on phenomenology, classification, pathogenesis and treatment. *Curr. Opin. Neurol.* 27, 468–476.
3. Klein, C. (2014). Genetics in dystonia. *Parkinsonism Relat. Disord.* 20 (Suppl 1), S137–S142.
4. Domingo, A., Erro, R., and Lohmann, K. (2016). Novel Dystonia Genes: Clues on Disease Mechanisms and the Complexities of High-Throughput Sequencing. *Mov. Disord.* 31, 471–477.
5. Zech, M., Boesch, S., Jochim, A., Weber, S., Meindl, T., Schormair, B., Wieland, T., Lunetta, C., Sansone, V., Messner, M., et al. (2016). Clinical exome sequencing in early-onset generalized dystonia and large-scale resequencing follow-up. *Mov. Disord.* <http://dx.doi.org/10.1002/mds.26808>.
6. Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al.; Exome Aggregation Consortium (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285–291.
7. 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.
8. 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.
9. Gulsuner, S., Walsh, 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.
10. Xu, B., Ionita-Laza, I., Roos, J.L., Boone, B., Woodruff, S., Sun, Y., Levy, S., Gogos, J.A., and Karayiorgou, M. (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. *Nat. Genet.* 44, 1365–1369.
11. Allen, A.S., Berkovic, S.F., Cossette, P., Delanty, N., Dlugos, D., Eichler, E.E., Epstein, M.P., Glauser, T., Goldstein, D.B., Han, Y., et al.; Epi4K Consortium; Epilepsy Phenome/Genome Project (2013). De novo mutations in epileptic encephalopathies. *Nature* 501, 217–221.

12. Chesi, A., Staahl, B.T., Jovičić, A., Couthouis, J., Fasolino, M., Raphael, A.R., Yamazaki, T., Elias, L., Polak, M., Kelly, C., et al. (2013). Exome sequencing to identify de novo mutations in sporadic ALS trios. *Nat. Neurosci.* 16, 851–855.
13. 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.
14. 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.
15. 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.
16. Petrovski, S., Wang, Q., Heinzen, E.L., Allen, A.S., and Goldstein, D.B. (2013). Genic intolerance to functional variation and the interpretation of personal genomes. *PLoS Genet.* 9, e1003709.
17. Kervestin, S., and Jacobson, A. (2012). NMD: a multifaceted response to premature translational termination. *Nat. Rev. Mol. Cell Biol.* 13, 700–712.
18. Dale, R.C., Grattan-Smith, P., Nicholson, M., and Peters, G.B. (2012). Microdeletions detected using chromosome microarray in children with suspected genetic movement disorders: a single-centre study. *Dev. Med. Child Neurol.* 54, 618–623.
19. Gana, S., Veggiotti, P., Sciacca, G., Fedeli, C., Bersano, A., Miccieli, G., Maghnie, M., Ciccone, R., Rossi, E., Plunkett, K., et al. (2012). 19q13.11 cryptic deletion: description of two new cases and indication for a role of WTIP haploinsufficiency in hypospadias. *Eur. J. Hum. Genet.* 20, 852–856.
20. Melo, J.B., Estevinho, A., Saraiva, J., Ramos, L., and Carreira, I.M. (2015). Cutis Aplasia as a clinical hallmark for the syndrome associated with 19q13.11 deletion: the possible role for UBA2 gene. *Mol. Cytogenet.* 8, 21.
21. Forzano, F., Napoli, F., Uliana, V., Malacarne, M., Viaggi, C., Bloise, R., Coviello, D., Di Maria, E., Olivieri, I., Di Iorgi, N., and Faravelli, F. (2012). 19q13 microdeletion syndrome: Further refining the critical region. *Eur. J. Med. Genet.* 55, 429–432.
22. Schuurs-Hoeijmakers, J.H., Vermeer, S., van Bon, B.W., Pfundt, R., Marcelis, C., de Brouwer, A.P., de Leeuw, N., and de Vries, B.B. (2009). Refining the critical region of the novel 19q13.11 microdeletion syndrome to 750 Kb. *J. Med. Genet.* 46, 421–423.
23. Malan, V., Raoul, O., Firth, H.V., Royer, G., Turleau, C., Bernheim, A., Willatt, L., Munnich, A., Vekemans, M., Lyonnet, S., et al. (2009). 19q13.11 deletion syndrome: a novel clinically recognisable genetic condition identified by array comparative genomic hybridisation. *J. Med. Genet.* 46, 635–640.
24. Kulharya, A.S., Michaelis, R.C., Norris, K.S., Taylor, H.A., and Garcia-Heras, J. (1998). Constitutional del(19)(q12q13.1) in a three-year-old girl with severe phenotypic abnormalities affecting multiple organ systems. *Am. J. Med. Genet.* 77, 391–394.
25. Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A., Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* 347, 1260419.
26. Ansari, K.I., and Mandal, S.S. (2010). Mixed lineage leukemia: roles in gene expression, hormone signaling and mRNA processing. *FEBS J.* 277, 1790–1804.
27. Shen, E., Shulha, H., Weng, Z., and Akbarian, S. (2014). Regulation of histone H3K4 methylation in brain development and disease. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369. <http://dx.doi.org/10.1098/rstb.2013.0514>.
28. Vallianatos, C.N., and Iwase, S. (2015). Disrupted intricacy of histone H3K4 methylation in neurodevelopmental disorders. *Epigenomics* 7, 503–519.
29. Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J., et al. (2007). Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445, 168–176.
30. Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufeland, K.R., Aasland, R., Anastasiadis, K., Ang, S.L., and Stewart, A.F. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. *Development* 133, 1423–1432.
31. Lubitz, S., Glaser, S., Schaft, J., Stewart, A.F., and Anastasiadis, K. (2007). Increased apoptosis and skewed differentiation in mouse embryonic stem cells lacking the histone methyltransferase Mll2. *Mol. Biol. Cell* 18, 2356–2366.
32. Jakovcevski, M., and Akbarian, S. (2012). Epigenetic mechanisms in neurological disease. *Nat. Med.* 18, 1194–1204.
33. Berdasco, M., and Esteller, M. (2013). Genetic syndromes caused by mutations in epigenetic genes. *Hum. Genet.* 132, 359–383.
34. Bai, G., Cheung, I., Shulha, H.P., Coelho, J.E., Li, P., Dong, X., Jakovcevski, M., Wang, Y., Grigorenko, A., Jiang, Y., et al. (2015). Epigenetic dysregulation of hairy and enhancer of split 4 (HES4) is associated with striatal degeneration in post-mortem Huntington brains. *Hum. Mol. Genet.* 24, 1441–1456.
35. Takata, A., Xu, B., Ionita-Laza, I., Roos, J.L., Gogos, J.A., and Karayiorgou, M. (2014). Loss-of-function variants in schizophrenia risk and SETD1A as a candidate susceptibility gene. *Neuron* 82, 773–780.
36. Jones, W.D., Dafou, D., McEntagart, M., Woollard, W.J., Elmslie, F.V., Holder-Espinasse, M., Irving, M., Saggart, A.K., Smithson, S., Trembath, R.C., et al. (2012). De novo mutations in MLL cause Wiedemann-Steiner syndrome. *Am. J. Hum. Genet.* 91, 358–364.
37. Kleefstra, T., Kramer, J.M., Neveling, K., Willemsen, M.H., Koemans, T.S., Vissers, L.E., Wissink-Lindhout, W., Fenckova, M., van den Akker, W.M., Kasri, N.N., et al. (2012). Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. *Am. J. Hum. Genet.* 91, 73–82.
38. Ng, S.B., Bigham, A.W., Buckingham, K.J., Hannibal, M.C., McMillin, M.J., Gildersleeve, H.I., Beck, A.E., Tabor, H.K., Cooper, G.M., Mefford, H.C., et al. (2010). Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.* 42, 790–793.
39. Gupta, S., Kim, S.Y., Artis, S., Molfese, D.L., Schumacher, A., Sweatt, J.D., Paylor, R.E., and Lubin, F.D. (2010). Histone methylation regulates memory formation. *J. Neurosci.* 30, 3589–3599.
40. Lim, D.A., Huang, Y.C., Swigut, T., Mirick, A.L., Garcia-Verdugo, J.M., Wysocka, J., Ernst, P., and Alvarez-Buylla, A.

- (2009). Chromatin remodelling factor Mll1 is essential for neurogenesis from postnatal neural stem cells. *Nature* 458, 529–533.
41. Kerimoglu, C., Agis-Balboa, R.C., Kranz, A., Stilling, R., Bahari-Javan, S., Benito-Garagorri, E., Halder, R., Burkhardt, S., Stewart, A.F., and Fischer, A. (2013). Histone-methyltransferase MLL2 (KMT2B) is required for memory formation in mice. *J. Neurosci.* 33, 3452–3464.
42. Ronan, J.L., Wu, W., and Crabtree, G.R. (2013). From neural development to cognition: unexpected roles for chromatin. *Nat. Rev. Genet.* 14, 347–359.
43. Bjornsson, H.T., Benjamin, J.S., Zhang, L., Weissman, J., Gerber, E.E., Chen, Y.C., Vaurio, R.G., Potter, M.C., Hansen, K.D., and Dietz, H.C. (2014). Histone deacetylase inhibition rescues structural and functional brain deficits in a mouse model of Kabuki syndrome. *Sci. Transl. Med.* 6, 256ra135.

Supplemental Data

**Haploinsufficiency of *KMT2B*, Encoding
the Lysine-Specific Histone Methyltransferase 2B,
Results in Early-Onset Generalized Dystonia**

Michael Zech, Sylvia Boesch, Esther M. Maier, Ingo Borggraefe, Katharina Vill, Franco Laccone, Veronika Pilshofer, Andres Ceballos-Baumann, Bader Alhaddad, Riccardo Berutti, Werner Poewe, Tobias B. Haack, Bernhard Haslinger, Tim M. Strom, and Juliane Winkelmann

Table S1 Trio-whole-exome sequencing statistics for index family F1

Sample #	Reads	Mapped reads	Percent	Mapped sequence (Gb)	Target bases				Average coverage
					> 1x	> 4x	> 8x	> 20x	
F1-II-5 (index subject)	123,825,063	123,550,183	99.78	12.51	99.78	99.66	99.52	98.47	147.22
F1-I-1 (father)	171,161,611	170,402,003	99.56	17.29	99.92	99.85	99.8	99.2	193.94
F1-I-2 (mother)	145,562,251	145,140,930	99.71	14.7	99.83	99.72	99.65	98.94	168.71

Table S2 Biallelic, disease-segregating variants identified in index subject F1-II-5

Gene	OMIM disease-association	Genomic position (hg 19)	RefSeq transcript	Variation nucleotide	Variation amino acid	Variant type	Known disease-causing mutation (ClinVar)	dbSNP142	CADD prediction	Frequency in-house exomes ^a (allele count/ total allele number)	Frequency ExAC (allele count/ total allele number)	Mode	Segregation (Sanger sequencing validation)
<i>ANKRD1</i>	no	chr10:92,679,979	NM_014391.2	c.154C>G	p.Pro52Ala	missense	no	rs397517248	0.003	3/15800	13/121400	hom	yes
<i>ABCC2</i>	yes (MIM: 237500)	chr10:101,565,180	NM_000392.4	c.1506G>C	p.Met502Ile	missense	no	not found	29.4	not found	1/121366	hom	yes

Parent-child trio whole-exome sequencing data were used to identify rare (minor allele frequency ≤ 0.001) protein-altering (including missense, nonsense, splice-site, stop-loss, in-frame insertion and deletion, and frameshift) homozygous and compound heterozygous variants in index subject F1-II-5. After co-segregation testing in all available members of family F1, homozygous missense variants in *ANKRD1* and *ABCC2* were retained. Single heterozygous variants in *ANKRD1* have been linked to cardiomyopathies, without any definitive disease-association reported in the OMIM database. Given the predominant relevance of *ANKRD1* in cardiac and skeletal muscle and its association with cardiac disease, we considered this gene unlikely to be causative for the observed dystonia phenotype. Biallelic mutations of *ABCC2* have been causally related to Dubin-Johnson syndrome (MIM: 237500), a benign congenital liver disease characterized by conjugated hyperbilirubinemia. Given the well-established and exclusive association of *ABCC2* mutations with Dubin-Johnson syndrome, we considered this gene unlikely to be causative for the observed dystonia phenotype. Hom = homozygous. ^aconsisting of 7900 non-dystonia exomes.

Table S3 De novo monoallelic variants identified in index subject F1-II-5

Gene	OMIM disease-association	Genomic position (hg 19)	RefSeq transcript	Variation nucleotide	Variation amino acid	Variant type	Known disease-causing mutation (ClinVar)	dbSNP142	CADD prediction	Frequency in-house exomes ^a (allele count/ total allele number)	Frequency ExAC (allele count/ total allele number)	Read depth (WES)	Variant quality (WES)	Sanger sequencing confirmation
<i>TDRD6</i>	no	chr6:46,655,897	NM_001010870.2	c.32G>T	p.Gly11Val	missense	no	not found	23	not found	not found	39	136	yes
<i>NYNRIN</i>	no	chr14:24,880,350	NM_025081.2	c.2483G>A	p.Gly828Glu	missense	no	not found	24.9	not found	not found	194	225	yes
<i>KMT2B</i>	no	chr19:36,223,856	NM_014727.2	c.6406delC	p.Leu2136Serfs*17	frameshift	no	not found	25.8	not found	not found	90	217	yes

Parent-child trio whole-exome sequencing data were used to identify non-annotated, protein-altering (including missense, nonsense, splice-site, stop-loss, in-frame insertion and deletion, and frameshift) de novo variants in index subject F1-II-5. Sanger sequencing verified the variants and their de novo status. None of the three genes affected by the de novo mutational events were listed as disease-associated in the OMIM database. A single-nucleotide deletion in *KMT2B* was the only variant predicted to have a severe impact on protein structure. WES = whole-exome sequencing. ^aconsisting of 7900 non-dystonia exomes.

Table S4 *KMT2B* loss-of-function (LoF) variants annotated in the ExAC dataset

Genomic position (hg 19)	Reference	Alternate	cDNA change	Protein change	Variant type	Allele count/ total allele number	Predicted protein truncation	QualBy Depth (QD)	Flags
chr19:36,213,260	G	C	c.2458-1G>C	N/A	canonical splice	1/78836	no	12.4	none
chr19:36,214,909	G	T	c.3334+1G>T	N/A	canonical splice	1/106404	no	12.37	none
chr19:36,223,613	G	GC	c.6164dupC	p.Arg2057Profs*3	frameshift	1/92432	yes	4.44	homopolymer run
chr:19:36,223,857	T	TC	c.6408dupC	p.Ala2139Glyfs*6	frameshift	2/70086	yes	1.32	homopolymer run

The ExAC server was queried for high-confidence LoF variant calls (including nonsense, frameshift, and essential splice-site variants) affecting the canonical transcript of *KMT2B* (RefSeq: NM_014727.2, Ensemble: ENST00000222270). The c.2458-1G>C and c.3334+1G>T splice-site variants are expected not to cause a protein truncation but to maintain the translational reading frame (in-frame skipping of exons 4 and 8, respectively). Annotation for the two listed frameshift variants appears to be dubious and they possibly represent sequencing artefacts as they each map to a homopolymer run of six C nucleotides and show very low QualByDepth (QD) scores.¹ The ExAC-derived probability of being LoF intolerant (pLI) metric for *KMT2B* is 1.0, indicating that it belongs to the set of genes that is highly intolerant to LoF variation.²

¹Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* 43, 11 10 11-33.

²Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536, 285-291.

Table S5 Phenotypic characteristics of 30 individuals with early-onset generalized dystonia in the replication cohort

Replication cohort (RC) sample #	Sex	Age at sampling (y)	Age at onset (y)	Site of onset	Leg involvement	Generalized dystonia classification	Associated features		Developmental delay	Family history	Trio-WES
							Additional movement abnormalities	Other neurological or systemic manifestations			
RC-1 ^a	F	51	7	hand	+	isolated	none	none		negative	
RC-2 ^a	F	36	8	hand	+	isolated	none	none		negative	
RC-3 ^a	M	29	7	neck	-	isolated	none	none		negative	
RC-4 ^a	F	40	7	neck	+	isolated	none	none		negative	
RC-5 ^a	F	47	5	neck	+	isolated	none	none		negative	
RC-6 ^a	F	63	3	neck	+	isolated	none	none		negative	
RC-7 ^a	M	59	10	hand	+	isolated	none	none		positive	
RC-8 ^a	F	28	2	foot	+	combined	myoclonus	pyramidal signs, hyperreflexia		negative	+
RC-9 ^a	F	43	20	neck	+	isolated	none	none		negative	
RC-10	F	42	29	foot	+	combined	parkinsonism	none		positive	
RC-11	F	57	2	neck	+	isolated	none	none		negative	+
RC-12	M	41	1	hand	+	combined	myoclonus	microcephaly, intellectual disability	+	negative	
RC-13	M	19	0.5	upper limbs	+	isolated	none	none		negative	+
RC-14	M	47	1	neck	+	isolated	none	none		negative	
RC-15	M	20	6	foot	+	combined	myoclonus	microcephaly, cognitive impairment, short stature	+	negative	
RC-16 (F2, F2-II-1)	F	11	3	foot	+	isolated	none	microcephaly, short stature, short weight, strabismus	+	negative	+
RC-17	M	64	28	neck	+	isolated	none	none		negative	
RC-18	M	26	7	hand	+	isolated	none	none		negative	
RC-19	F	33	29	hand	+	isolated	none	none		negative	+
RC-20	F	48	29	foot	+	isolated	none	none		negative	
RC-21	F	50	10	neck	+	isolated	none	none		negative	
RC-22	F	27	6	neck	-	combined	myoclonus	none		negative	
RC-23	F	19	4	neck	+	combined	myoclonus	none		negative	+
RC-24	F	34	4	hand	+	isolated	none	none		positive	
RC-25	F	18	3	hand	+	isolated	none	seizures	+	negative	

RC-26 (F3, F3-II-3)	M	15	11	foot	+	isolated	none	mild intellectual disability, microcephaly, syndactyly, astigmatism	+	negative
RC-27	F	60	17	neck	+	isolated	none	seizures		negative
RC-28	F	29	20	neck	-	isolated	none	none		negative
RC-29 (F4, F4-III-2)	F	6	3	foot	+	isolated	none	mild intellectual disability, microcephaly, short stature, short weight, astigmatism	+	positive
RC-30	M	6	4	hand	+	combined	myoclonus	none	+	negative

^apreviously reported in Zech et al., Mov Disord 2016.¹ The following are the corresponding identifiers from Zech et al., Mov Disord 2016¹ for the replication cohort samples: RC-1=patient #2, RC-2=patient #5, RC-3=patient #6, RC-4=patient #8, RC-5=patient #9, RC-6=patient #10, RC-7=patient #11, RC-8=patient #12, RC-9=patient #15.

¹Zech, M., Boesch, S., Jochim, A., Weber, S., Meindl, T., Schormair, B., Wieland, T., Lunetta, C., Sansone, V., Messner, M., et al. (2016). Clinical exome sequencing in early-onset generalized dystonia and large-scale resequencing follow-up. Mov Disord. <http://dx.doi.org/10.1002/mds.26808>.

Table S6 Summary statistics for 30 exomes in the replication cohort

Sample #	Reads	Mapped reads	Percent	Mapped sequence (Gb)	Target bases				Average coverage
					> 1x	> 4x	> 8x	> 20x	
RC-1	102,160,790	102,058,482	99.9	10.32	99.77	99.57	99.25	97.22	128.96
RC-2	112,816,224	112,682,303	99.88	11.39	99.77	99.56	99.27	97.55	144.42
RC-3	136,658,667	136,481,664	99.87	13.8	99.89	99.76	99.58	98.37	165.43
RC-4	132,113,474	131,969,813	99.89	13.34	99.78	99.64	99.6	98.25	161.24
RC-5	97,158,400	97,051,400	99.89	9.81	99.76	99.61	99.4	97.84	123.86
RC-6	96,281,590	96,146,663	99.86	9.72	99.76	99.6	99.36	97.69	122.44
RC-7	134,597,597	134,440,694	99.88	13.59	99.89	99.78	99.67	98.77	164.41
RC-8	116,714,305	116,519,263	99.83	11.79	99.77	99.65	99.51	98.41	140.28
RC-9	111,545,754	111,313,621	99.79	11.27	99.89	99.76	99.58	98.25	136
RC-10	134,803,142	134,230,442	99.58	13.62	99.83	99.7	99.61	98.8	155.89
RC-11	151,491,437	151,059,176	99.71	15.3	99.82	99.71	99.65	98.96	172.98
RC-12	137,054,877	136,157,539	99.35	10.14	99.89	99.79	99.67	98.77	156.84
RC-13	149,145,841	148,597,025	99.63	15.06	99.9	99.81	99.68	99.07	182.13
RC-14	153,037,501	151,847,154	99.22	15.46	99.95	99.82	99.63	98.77	188.67
RC-15	141,958,074	140,692,700	99.11	14.34	99.95	99.81	99.6	98.65	175.66
RC-16 (F2, F2-II-1)	106,695,812	106,088,678	99.43	10.78	99.82	99.63	99.31	97.7	133.78
RC-17	153,647,953	152,409,617	99.19	15.52	99.96	99.84	99.66	98.89	187.43
RC-18	143,675,890	142,482,773	99.17	14.51	99.96	99.84	99.68	98.9	175.98
RC-19	171,499,116	170,136,140	99.21	17.32	99.88	99.75	99.61	99.03	206.32
RC-20	153,587,798	152,801,538	99.49	15.36	99.82	99.71	99.64	98.96	175.96
RC-21	111,612,148	111,612,148	99.88	11.27	99.77	99.64	99.5	98.32	140.71
RC-22	124,013,553	123,157,935	99.31	12.53	99.9	99.75	99.57	98.74	155.62
RC-23	123,741,890	123,583,574	99.87	12.5	99.77	99.66	99.52	98.49	154.64
RC-24	131,020,076	130,861,509	99.88	13.23	99.78	99.67	99.57	98.7	162.02
RC-25	124,368,112	124,002,790	99.71	12.56	99.8	99.67	99.5	98.67	150.05
RC-26 (F3, F3-II-3)	128,583,965	128,408,566	99.86	12.99	99.89	99.78	99.66	98.71	161.41
RC-27	145,640,525	145,374,692	99.82	14.71	99.8	99.65	99.52	98.67	179.75
RC-28	120,865,528	120,721,846	99.88	12.21	99.77	99.66	99.54	98.53	150.01

RC-29 (F4, F4-III-2)	170,878,962	169,272,874	99.06	17.26	99.89	99.76	99.61	99.04	210.64
RC-30	114,051,928	113,905,597	99.87	11.52	99.88	99.75	99.59	98.33	142.17

Table S7 Coverage summary statistics for the target regions of *TDRD6*, *NYNRIN*, and *KMT2B* across 30 exomes in the replication cohort

Gene ^a	Exons ^b									
<i>TDRD6</i> (NM_001010870.2)	1	2	3	4						
	225.2±34.93 (163.42-292.71)	140.45±25.4 (97.9-186.41)	151.05±29.13 (99.75-217.9)	245.88±31.14 (186.68-315.74)						
<i>NYNRIN</i> (NM_025081.2)	1	2	3	4	5	6	7	8		
	131.23±38.35 (71.04-211.12)	243.93±58.66 (160.5-373.01)	231.35±41.14 (156.21-308.34)	189.34±45.15 (127.22-298.89)	253.31±55.9 (148.74-411.82)	98.66±17.36 (70.16-133.98)	218.38±40.82 (147.26-308.66)	195.17±38.42 (130.44-273.04)		
<i>KMT2B</i> (NM_014727.2)	1	2	3	4	5	6	7	8	9	10
	13.85±4.0 (5.07-21.55)	218.19±48.67 (147.41-316.32)	174.3±31.22 (117.55-229.94)	175.4±39.65 (109.77-250.9)	146.42±35.82 (79.53-222.12)	139.45±42.64 (82.32-235.39)	127.11±31.41 (73.56-197.12)	176.1±33.8 (100.21-237.5)	112.42±25.31 (70.44-172.62)	282.88±45.32 (192.47-377.94)
	11	12	13	14	15	16	17	18	19	20
	164.02±42.07 (105.67-237.22)	398.03±87.28 (230.72-530.49)	286.29±58.54 (157.66-374.77)	66.36±16.78 (30.61-95.16)	117.14±31.66 (71.92-179.2)	273.14±56.65 (177.67-404.05)	204.55±42.6 (114.26-289.67)	176.34±33.78 (110.4-231.1)	66.69±17.53 (39.29-100.44)	207.04±46.86 (114.45-282.26)
	21	22	23	24	25	26	27	28	29	30
	306.41±64.14 (188.4-398.22)	276.32±63.84 (161.36-427.48)	119.86±26.06 (74.46-180.41)	207.71±42.31 (137.43-305.51)	216.97±54.33 (124.95-315.4)	111.77±23.61 (69.99-171.22)	197.33±41.11 (116.66-269.47)	132.45±32.21 (71.96-183.51)	129.72±22.9 (77.53-172.32)	90.44±20.59 (56.21-133.21)
	31	32	33	34	35	36	37			
	230.67±41.09 (148.54-313.32)	296.8±56.39 (186.59-379.58)	133.35±33.16 (78.82-183.07)	157.53±30.42 (102.52-218.27)	204.78±47.22 (109.64-309.13)	283.45±62.7 (162.23-408.91)	252.26±70.24 (165.84-410.88)			

^acandidate gene prioritized in index subject F1-II-5 (NCBI accession number).

^bfor each exon and the adjacent splice junctions, average coverage ± standard deviation and range of coverage (minimum coverage – maximum coverage) across the 30 replication cohort exomes are given.

Table S8 De novo monoallelic variants identified in individual F2-II-1

Gene	OMIM disease- association	Genomic position (hg 19)	RefSeq transcript	Variation nucleotide	Variation amino acid	Variant type	Known disease- causing mutation (ClinVar)	dbSNP142	CADD prediction	Frequency in- house exomes ^a (allele count/ total allele number)	Frequency ExAC (allele count/ total allele number)	Read depth (WES)	Variant quality (WES)	Sanger sequencing confirmation
<i>HCN2</i>	no	chr19:605,113	NM_001194.3	c.1109T>A	p.Leu370His	missense	no	not found	26.2	not found	not found	19	182	yes
<i>KMT2B</i>	no	chr19:36,211,882	NM_014727.2	c.1633C>T	p.Arg545*	nonsense	no	not found	36	not found	not found	361	225	yes

Parent-child trio whole-exome sequencing data were used to identify non-annotated, protein-altering (including missense, nonsense, splice-site, stop-loss, in-frame insertion and deletion, and frameshift) de novo variants in individual F2-II-1. Sanger sequencing verified the variants and their de novo status. None of the two genes affected by de novo mutational events were listed as disease-associated in the OMIM database. A nonsense substitution in *KMT2B* was the only variant predicted to have a severe impact on protein structure. WES = whole-exome sequencing. ^aconsisting of 7900 non-dystonia exomes.