Introduction

Intellectual disability (ID) is defined by significant impairment of cognitive and adaptive functions with onset before 18 years of age. It has an estimated worldwide prevalence of 1–3%, with moderate or severe forms of ID (IQ <50) affecting up to 0.5% of the population in Western countries [1]. We and others have reported that de novo point mutations (including single nucleotide substitutions (SNVs) and small insertions/deletions, referred herein collectively as DNMs) play a significant role in the genetics of ID [2–5]. Similarly, DNMs were found to be implicated in the etiology of other neurodevelopmental disorders overlapping with ID, such as autism spectrum disorders (ASD), epileptic encephalopathy and schizophrenia [6–10]. DNMs represent the most extreme form of rare genetic variations; they are more deleterious, on average, than inherited variations because they have been subjected to less stringent evolutionary selection. Importantly, they provide a mechanism by which early-onset reproductively lethal diseases remain frequent in the population. This makes these mutations prime candidates for causing diseases that occur sporadically, and that decrease the reproductive fitness and incur a large degree of selection against phenotypes such as ID. Based on these considerations, we hypothesized that the contribution of DNMs is greater in more severe forms of ID. In order to explore this hypothesis, we performed high-depth exome sequencing in 41 trios consisting of individuals with moderate or severe ID and their healthy parents and assessed the contribution of DNMs to this condition.

Results/Discussion

We performed exome sequencing in 41 individuals with ID and their unaffected parents. We identified a total of 83 putative DNMs in as many genes within both coding and consensus splice site sequences. Sanger sequencing confirmed 81 of these as de novo and 2 as inherited from one of the parents (Table S1). All of these DNMs were represented by ≥25% of reads, suggesting that they are unlikely to be associated with somatic mosaicism. The fact that...
the mutant and wild-type peaks on Sanger chromatograms were comparable in size is consistent with this conclusion. The average DNM rate per trio was 1.98, with only 3 trios containing no detectable DNMs (Figure 1). The observed de novo SNV rate in the consensus coding sequences (CCDS) was 1.56 events per trio or 2.58 \times 10^{-8} per base per generation (64 SNVs in 2,477,702,175 CCDS bases sequenced at \(10^6\) in the 41 affected individuals), which is significantly higher than the expected population rate of 1.65 \times 10^{-8} (R binomial test, \(p = 0.0007\), or than the ones experimentally determined from exome sequencing studies in control trios (1.28 \times 10^{-8} and 1.51 \times 10^{-8}) [2,4]. Considering only de novo SNVs affecting the coding and the canonical splice sites (\(\text{AG, GT at intronic positions} = 1/-2\) and +1/+2 of the acceptor and donor splice sites, respectively), 73% were missense and 11% were nonsense and canonical splice site mutations. We found a significant excess of these de novo nonsense and splice site mutations in the probands of our cohort when compared to data from exome sequencing of 54 control trios with no family history of ID [4,11] or of 393 quartets, including unaffected siblings of individuals with ASD (R binomial test, \(p = 0.0015\) and \(p = 0.02\), respectively) (Table 1) [7,9,10]. Such an excess of deleterious DNMs suggest that at least a subset of them are pathogenic.

Twelve DNMs were found in as many probands in genes previously associated with ID based on the documentation of deleterious DNMs in at least 4 unrelated individuals with similar phenotypes. Nine of these DNMs are Loss-of-Function (LoF) variants (nonsense, frameshift and canonical splice variants) and affect the following genes: ARID1B [OMIM 614556] [12], CHD2 [OMIM 602119] [4,13], FOXG1 [OMIM 164874] [14], GATA2 [OMIM 614990] [2,15], MBD5 [OMIM 611472] [9,13,16–18], MED13L [OMIM 600771] [7,19–21], SETBP1 [OMIM 611060] [4,22], TCF4 [OMIM 602272] [4,23–25], and WDR45 [OMIM 300526] [26,27] (Tables 2 and 3). None of these 9 DNMs were found in public SNP databases. The phenotype of each of the probands is consistent with that of subjects previously described with mutations in these respective genes, with two exceptions (Text S1). Although truncating mutations in CHD2 have been reported in individuals with epileptic encephalopathy [4,6,13], the individual described herein with a CHD2 frameshift mutation has no history of epilepsy, suggesting that LoF mutations in CHD2 are associated with greater clinical heterogeneity than initially expected. Another example of a gene associated with clinical heterogeneity in our dataset is SETBP1. Missense mutations clustering in a conserved 11-bp coding region of SETBP1 have been reported to cause Schinzel-Giedion syndrome (OMIM 269150), a condition characterized by severe ID and specific craniofacial features [22]. In contrast, our case carried a de novo truncating mutation in SETBP1 and showed moderate non-syndromic ID without the typical craniofacial features of Schinzel-
Giedion syndrome. Recent studies reported a similar phenotype in individuals with a truncating mutation in SETBP1 or microdeletions encompassing SETBP1 [4,28]. Collectively, these observations suggest that SETBP1 haploinsufficiency results in a different phenotype than that induced by the missense mutations reported in Schinzel-Giedion syndrome, which presumably lead to a gain-of-function or a dominant negative effect [22]. We conclude that all of these 9 DNMs are likely to be pathogenic.

The three other DNMs in genes previously associated with ID include an in-frame insertion in GABRB3 [OMIM 137192], a missense in TBR1 [OMIM 604616] and a missense in GRIN2B [OMIM 138252] (Tables 2 and 3). All of these DNMs affect conserved residues and are predicted to be damaging. Moreover, none of them were found in public SNP databases. Damaging missense mutations in GABRB3 have been previously documented in cases with ID and intractable epilepsy with various types of seizures [6]. Individual 1843.647 also showed ID and intractable epilepsy with variable severity with or without ASD and epilepsy [2,6,31,32]. We conclude that these three DNMs are also likely to be pathogenic.

Among the remaining cases, 22 have predicted-damaging DNMs, including 7 LoF mutations (5 frameshifts, 1 nonsense, 1 consensus splice site), 13 missenses, 1 deletion, and 1 synonymous mutation whose predicted effect on splicing was confirmed by RT-PCR (Figure S1). Interestingly, deleterious DNMs in 6 of these genes (HNRNPU [OMIM 602869], WAC [OMIM 615049], RYR2 [OMIM 180902], MYH10 [OMIM 160776], EIF2CG1 [OMIM 606228], COLA3BP [OMIM 604677]) have previously been reported in at least one individual with ID. We discuss hereafter the DNMs that we identified in these genes (Tables 2 and 3).

HNRNPU [OMIM 602869] codes for a highly conserved protein that binds RNAs and mediates different aspects of their metabolism and transport. Chromosome 1q44 microdeletions have defined a critical region associated with ID and seizures that encompasses HNRNPU as well as two other genes [33,34]. Two truncating and one splice mutations in HNRNPU were subsequently identified in individuals with ID and seizures [6,13,24]. Two of these mutations occurred de novo whereas the origin of the other one was not elucidated. One of these individuals also showed ASD whereas the case with the splice mutation displayed syndromic features, including panhypopituitarism, bifid great toe and vertebral segmental defects. We identified an individual (1464.524) who carries a de novo truncating mutation (c.511C>T, p.Gln171*) in HNRNPU. This mutation is located in an upstream coding exon present in all isoforms, thus having the potential to induce nonsense mRNA mediated decay [35]. Moreover, inspection of the Exome Variant Server (EVS) database (6500 exomes) revealed no LoF variants in HNRNPU, indicating that haploinsufficiency of this gene is not tolerated. Our case displayed ID, epilepsy and ASD (Text S1), a phenotype that is similar to that of the other non-syndromic cases with DNMs in this gene, further supporting its involvement in ID.

WAC encodes a nuclear protein that interacts with RNF20/40 to regulate histone H2B ubiquitination, chromatin organization, and gene transcription [36]. De novo microdeletions encompassing WAC and a nonsense DNM in WAC in individuals with severe ID were recently reported [2,37]. Our subject (762.297) carries a truncating mutation in WAC (c.263_266del, p.Glu88Glyfs*103). This mutation is located in an upstream coding exon present in all isoforms. Inspection of the EVS database revealed no LoF variants in WAC. Individual 762.297 showed moderate ID without any distinguishing features on clinical examination and brain imaging.
Table 2. Top risk DNMs identified in this study.

| Individual | sex | Gene | Genomic change (hg19) | NCBI RefSeq. | AA | MutationType | Change; prediction (score) |
|------------|-----|------|----------------------|--------------|----|--------------|--------------------------|
| 289.143    | F   | ARID1B | chr6:157511198delC   | NM_020732.3   | 2236| frameshift del | c.3716delC (Pro1239Hisfs*5) |
| 1396.504   | F   | CHD2  | chr2:934705145>G     | NM_001271.3   | 1828| nonsense     | c.335C>G (p.Ser112*)     |
| 893.339    | F   | FOXG1 | chr14:2923691delG    | NM_005249.4   | 489 | frameshift del | c.506delG (p.Gly169Alafs*23) |
| 1907.666   | F   | GATA2D2 | chr1:153785930T>C   | NM_020699.2   | 593 | CSS          | c.1217-2A>G              |
| 79.65      | M   | MBDS3 | chr2:162274305T>C    | NM_018328.4   | 1494| frameshift del | c.340_347del (p.Lys114Glyfs*35) |
| 820.316    | F   | MED13L | chr12:116446509_116446510delCT | NM_015335.4 | 2210| frameshift del | c.1708_1709delCT (p.Ser570Phefs*27) |
| 893.339    | F   | SETBP1 | chr18:23531126delC   | NM_015559.2   | 1596| frameshift del | c.1821delC (p.Ser608Alafs*22) |
| 1045.400   | M   | TCF4  | chr18:52921925G>A    | NM_00183962.4 | 667 | nonsense     | c.1153C>T (p.Arg385*)    |
| 1883.659   | F   | WDR45 | chrX:48935736G>A     | NM_007075.3   | 361 | nonsense     | c.19T (p.Arg7*)           |
| 1843.647   | M   | GABRB3 | chr15:26866506_26866507insACC | NM_021912.4 | 473 | insertion    | c.413_415dupACC (p.Asn138_Arg139insHis); PVN (~12.3) |
| 121.83     | M   | TBR1  | chr2:16227430T>C     | NM_006593.2   | 682 | missense     | c.811T>C (p.Trp271Arg); SIFT (0.00); PFF2 (1.0); PVN (~11.5) |
| 838.321    | M   | GRN2B | chr12:13720098C>T    | NM_000834.3   | 1484| missense     | c.2459G>A (p.Gly820Glu); SIFT (0.00); PFF2 (1.0), PVN (~7.5) |
| 1464.524   | M   | HNRNPC | chr1:2824672_2824678delAGAG | NM_016628.4 | 647 | frameshift del | c.413G>A (p.Arg138Pro); SIFT (0.01); PFF2 (0.48); PVN (~24.9) |
| 341.162    | M   | RYR2  | chr1:237995909T>G    | NM_001035.2   | 4967| missense     | c.1838C>T (p.Arg612Cys); SIFT (0.00), PFF2 (1.00), PVN (~7.5) |
| 1871.656   | F   | MTHFD1 | chr17:9455445G>A     | NM_001256012.1 | 2007| missense     | c.595G>A (p.Gly199Ser); SIFT (0.00), PFF2 (1.00), PVN (~5.2) |
| 702.278    | F   | EIF2C1 | chr1:36359357G>A     | NM_012199.2   | 857 | missense     | c.1111G>A (p.Gly371Arg); SIFT (0.002), PFF2 (0.97), PVN (~7.3) |
| 1312.477   | M   | COL4A3BP | chr1:27412811C>T    | NM_00113010.5 | 752 | stop gain    | c.699_701delCT (p.Tyr233*) |
| 115.81     | M   | SET   | chr9:131456084_131456086delCTT | NM_001223821.1 | 290 | frameshift del | c.1347_1348insA (p.Tyr450leufs*92) |
| 670.267    | F   | EGR1  | chr5:137803485_137803485insA | NM_001964.2 | 543 | frameshift ins | c.1347_1348insA (p.Tyr450leufs*92) |
| 1439.518   | F   | PPP1CB | chr2:22902299dupA    | NM_0206871.8 | 327 | frameshift del | c.909dupA (p.Tyr304del*19) |
| 580.240    | M   | CHMP2A | chr19:59063688_59063688insG | NM_121826.2 | 222 | frameshift ins | c.286_287insC (p.Asn96Thrfs*35) |
| 1841.646   | M   | PPP2R2B | chr1:46070692C>G   | NM_018678.2   | 501 | missense     | c.413G>C (p.Arg138Pro); SIFT (0.01); PFF2 (0.48); PVN (~49) |
| 985.382    | M   | VPS4A | chr16:69353403_69353405delTCCC | NM_013245.2 | 437 | deletion     | c.577_579delTCCC (p.Ser193del); PVN (~12.3) |

AA, total amino acids. All predictions by SIFT (http://sift.jcvi.org/), PFF2 (PolyPhen-2; http://genetics.bwh.harvard.edu/pph2/) and PROVEAN (PVN; (http://provean.jcvi.org genome_submit.php) were damaging (scores indicated in parenthesis). CSS, Canonical splice site.

doi:10.1371/journal.pgen.1004772.t002
**Table 3. Genes affected by predicted-damaging DNMs identified herein and their implication in ID.**

| Mutation type               | ID-associated Genes with likely pathogenic DNMs | Candidate Genes with possibly pathogenic DNMs | Genes of unknown significance to ID* |
|-----------------------------|-----------------------------------------------|-----------------------------------------------|-------------------------------------|
| Missense                    | GRIN2B, TBR1                                  | MYH10, RYR2, EIF2C1, COL4A3BP, PPP2R2B        | BCORL1, VIPR1, MTUS1, WDR33, R3HDM1, |
| Nonsense                    | TCF4, CHD2, WDR45                              |                                               | FBXO28, MAPK8P1, KCNH1               |
| Canonical splice site*      | GATA2D2B                                      |                                               |                                     |
| Consensus splice site*      |                                               |                                               | GIT1                                |
| Frameshift deletion         | MBDS, ARID1B, MED13L, FOXG1, SETBP1           | WAC, SET                                      |                                     |
| Inframe insertion           | GABRB3                                        |                                               |                                     |
| Frameshift insertion        | EGR1, CHMP2A, PPP1CB                          |                                               |                                     |
| Inframe deletion            | VP54A                                         |                                               | NANS**                              |
| Synonymous – splicing*      |                                               |                                               |                                     |

All missense mutations were predicted damaging by SIFT and Polyphen-2. All in-frame deletions and insertions here were predicted damaging by PROVEAN (http://provean.jcvi.org genome_submit.php). *Predicted to affect splicing by both Human Splicing Finder (http://www.umd.be/HSF/) and Mutation Taster (http://www.mutationtaster.org/). **Splicing defect verified by RT-PCR (Figure S1).

doi:10.1371/journal.pgen.1004772.t003

a phenotype that is consistent with that observed in the previously reported patient with a truncating mutation in this gene (Text S1) [2]. Our finding, thus, further supports the involvement of WAC in ID.

**RYR2** encodes the cardiac and brain-expressed calcium release channel ryanodine receptor 2. Mutations in **RYR2** are typically associated with exercise-induced ventricular and atrial arrhythmias. Virtually all reported mutations in **RYR2** are missenses or in-frame deletions that are believed to confer a gain of function, resulting in an increase of Ca+ release [38,39]. We identified an individual (341.162) with ID, seizures, short stature and severe atrial arrhythmias (Text S1) who carries a predicted-damaging *de novo* missense mutation in **RYR2** (c.14864G>A, p.Gly4955Glu). Interestingly, 3 patients with seizures have previously been reported with DNMs in **RYR2**: 1) an individual with epileptic encephalopathy but presumably without a history of arrhythmia was recently found to carry a nonsense mutation (c.9586C>T, p.Arg3190*) in **RYR2** [6]; this DNM might not be disease-causing considering that the pathogenic impact of truncating mutations in **RYR2** remains unclear and that inspection of EVS revealed 5 different heterozygous LoF mutations in **RYR2**: 2) an individual with cognitive impairment, intractable seizures, short stature and subclinical ventricular tachycardia was found to carry a missense mutation (c.12563T>G, p.Leu4188Pro) [40]; and 3) an individual with intractable seizures but without cognitive impairment and arrhythmia was described with a missense mutation (c.14803G>A, p.Gly4935Arg) [41]. It is noteworthy that the mutation found in this last individual is in close proximity to that of our subject, affecting a highly conserved C-terminal region of the protein. Interestingly, mice heterozygous for the missense mutation p.R2474S in Ryr2 display generalized seizures and arrhythmias [42]. More recently, two brothers with ID, seizures and atrial arrhythmias were found to carry a missense mutation in **CLIC2** (OMIM 300138), which maps to the X chromosome [43]. **CLIC2** is a negative regulator of **RYR2**. The mutation was shown to stimulate the release of Ca2+ by keeping the **RYR2** channel in an open state, possibly due to a higher binding affinity for the **RYR2** protein. The specificity of the phenotype observed in our subject and its similarity with that of other individuals with DNMs in **RYR2** or with the mutation in **CLIC2** suggest that the mutation identified herein may be causal.

**MYH10** encodes the non-muscle myosin heavy chain IIB that is critical for heart and brain development [44,45]. Loss of **Myh10** function in mice results in embryonic lethality, hydrocephalus and neuronal migration defects but the cognitive and behavioural phenotype of heterozygous mice has not yet been reported. We identified a predicted-damaging *de novo* missense mutation (c.838C>T, p.Arg280Cys; individual 1871.656) in **MYH10**, affecting its conserved motor domain, whereas another group recently reported a *de novo* truncating mutation (c.2722G>T, p.Glu908*) in the same gene [46]. Both individuals displayed severe ID, microcephaly, and feeding difficulties as well as cerebral atrophy with increased intensities in bilateral basal ganglia and thalami on brain MRI (Text S1). The similarities between the phenotypes of these individuals raise the possibility that these mutations in **MYH10** are pathogenic. O’Roak et al. (2012) also reported a predicted-damaging *de novo* missense mutation (c.794A>G, p.Y265C; NM_001256012.1) in the motor domain of **MYH10**, in close proximity to the mutation identified herein, in a patient with ASD and moderate to severe ID. However no additional phenotypic data was available. Interestingly our patient with the **MYH10** mutation also displayed autistic features. Inspection of EVS for potential LoF mutations in **MYH10** showed the presence of a heterozygous frameshift deletion and a heterozygous splice site mutation. It is important to note, however, that these EVS variants were seen in single individuals and were not validated.

DNMs in **EIF2C1** and **COL4A3BP** have also been previously reported in single individuals with severe ID [2,4]. For each of these genes, the phenotype of the affected individuals appears similar to that of our subjects (Text S1). However, because of the lack of specific clinical features in these individuals, the occurrence of DNMs in unrelated subjects does not readily indicate pathogenicity, especially in the case of missense mutations whose functional consequences are not validated.

Among the remaining cases, we also identified 6 predicted-damaging DNMs in genes (**SET** [OMIM 600960], **EGRI** [OMIM 128990], **PPP1CB** [OMIM 600590], **CHMP2A**...
[OMIM 610893], PPP2R2B [OMIM 604325], and VPS4A [OMIM 609982]) that play biological functions relevant to ID (Table 2). Inspection of the EVS database revealed no LoF variants in these genes, with the exception of a single heterozygous variant in PPP1C1B (MAF = 1/12518) with a potential effect on splicing. In addition, some of these genes were found in proteomic studies to physically interact with the product of at least one ID-associated gene, further increasing the probability of their involvement in this disorder (see below and Figure 2). Each of these DNMs is discussed hereafter.

**SET** encodes a widely expressed multifunctional nuclear protein that affects pathways involved in ID, such as chromatin remodelling and gene transcription [47]. SET physically binds SETBP1 [48], whose disruption is known to cause severe ID (see above). In addition, recent studies indicate that SET directly interacts with MCPH1 (OMIM 607117) to ensure the proper temporal activation of chromosome condensation during mitosis [49]. Cells with SET knockdown exhibited abnormal condensed chromosomes similar to those observed in MCPH1-deficient fibroblasts. In addition, mutations that impair binding of MCPH1 to SET affect the ability of the former to rescue the abnormal chromosome condensation phenotype in fibroblasts from Mcph1 mutant mice. Recessive mutations in MCPH1 cause primary microcephaly, which is characterized by reduced brain size, without major structural abnormalities, and mild-to-moderate ID [50]. We identified a **de novo** deletion resulting in the creation of a premature stop codon in SET (c.699_701del, p.Tyr233*) in an individual (115.81) with congenital microcephaly, normal brain MRI, and moderate ID without any other distinguishing feature (Text S1). The functional relationship between MCPH1 and SET and the phenotypical similarities between cases with mutations in MCPH1 and our subject suggest that the truncating DNM in SET may be pathogenic.

**EGR1** encodes a transcription factor that plays a key role in learning and memory [51]. We identified a **de novo** truncating mutation (c.1347_1348insA, p.Tyr450Ilefs*92) in EGR1 in an individual (670.267) with severe non-syndromic ID and acquired microcephaly (Text S1). Mice harbouring a heterozygous deletion...
of Egr1 showed synaptic plasticity, learning and memory impairments [52,53]. Due to the prominent role of EGR1 in learning and memory and the impact of its haploinsufficiency on cognition in mice, we postulate that the truncating DNM identified herein in EGR1 may be pathogenic.

**PPP1CB**, which encodes a brain-enriched beta catalytic subunit of protein phosphatase 1 (PPP1), and **PPP2R2B**, which encodes a neuron-specific B regulatory subunit of protein phosphatase 2 (PPP2A), have been shown to regulate synaptic plasticity pathways [54,55]. Individual 1439.518 carries a truncating mutation (c.909delA, p.Tyr303Hle*19) in **PPP1CB**. This individual displayed severe ID, growth retardation and some dysmorphic features (Text S1). Individual 1841.646 carries a predicted-damaging missense mutation (c.413G>C, p.Arg138Pro) in **PPP2R2B**. This individual showed ID, intractable seizures and autistic features (Text S1). The pathogenic impact of these mutations remains uncertain at this point.

Among these candidate genes, CHMP2A and VPS4A are of special interest, as the proteins encoded by each are interacting partners. VPS4 ATPases play a critical role in the ESCRT pathway by recognizing membrane-associated ESCRT-III complexes and catalyzing their disassembly, a process that involves a direct interaction between CHMP2A and VPS4A [56]. The ESCRT-III pathway is involved in key cellular processes, including formation of endocytic multivesicular bodies, the abscission stage of cytokinesis, as well as centrosome and spindle maintenance [57]. Specific depletion of either CHMP2A or VPS4A proteins in cultured cells disrupts mitosis by inhibiting abscission and altering centrosome and spindle pole numbers [58].

We identified an individual (580.240) with a de novo frameshift insertion (c.286_287insC, p.Asn96Thrfs*35) in CHMP2A. We identified only one case (692.274) that could potentially be recessive or X-linked forms of ID, epilepsy or ASD (see Table S3 for the complete list of inherited rare variants in each proband).

The contribution of inherited autosomal or X-linked recessive mutations remains uncertain at this point. In summary, our trio exome sequencing study identified deleterious DNMs in genes previously causally linked to ID in 12 cases out of the 41 studied herein, resulting in a molecular diagnostic yield of 29%. Recently, de Ligt et al. (2012) and Rauch et al. (2012) performed trio exome sequencing in individuals with severe ID and obtained a diagnostic yield, based on the presence of predicted-damaging point mutations in currently known ID genes, of 20% and 35%, respectively [2,4,21]. Overall, the contribution of inherited autosomal or X-linked recessive mutations appears limited in the three cohorts. The study of Rauch et al. (2012) and ours were intentionally centered on sporadic cases, which might have created a bias against inherited mutations. However, it is important to emphasize that most cases with moderate or severe ID are sporadic, at least in Western societies. de Ligt et al. (2012) observed a proportionally smaller number of DNMs in their cohort when compared to that of Rauch et al. (2012) and ours. This difference may be related to the use of a different sequencing technology, which is associated with a lower depth, possibly accounting for the lower diagnostic yield observed in this study. Indeed, exploration of a subset of unexplained cases from this cohort using whole-genome sequencing revealed additional pathogenic DNMs in known ID genes, bringing the point mutation molecular diagnostic yield in this cohort to 34% [21].

Our study also provides evidence for the potential pathogenicity of 12 additional DNMs in as many genes. Some of these genes represent strong candidates. For instance, both HNRNPU and WAC map to small critical regions associated with ID, which were defined by a series of microdeletions. De novo truncating mutations in each of these genes were previously described in cases with severe ID. We now report additional truncating DNMs in these genes in cases with similar phenotypes as those already published, further supporting their involvement in ID. Similarly, we and others have identified damaging DNMs in RYR2 and MYH10 in patients with similar features. Finally, we discovered a truncating DNM in EGR1, the haploinsufficiency of which affects learning and memory in mice. Although the characterization of additional cases will be needed to confirm the involvement of these candidate genes in ID, these results indicate that the contribution of DNMs to the pathogenesis of moderate or severe ID could be even greater than that suggested by the diagnostic rate observed in this study.

In conclusion, our study suggests that DNMs represent a predominant cause of moderate or severe ID. High-depth trio-based exome sequencing is an effective method to establish molecular diagnosis in such cases.

**Materials and Methods**

**Study subjects and ethics statement**

The cases reported here (18 males, 23 females) with moderate (n = 12) or severe (n = 29) ID were recruited at the Sainte-Justine Hospital (Montreal, Canada), after the approval of the ethics committee, and informed consent was obtained from each participant or legal guardian. Inclusion criteria for the probands were: 1) absence of a history of ID, epilepsy or ASD in first or second-degree relatives; 2) moderate or severe ID with or without
epilepsy or autistic features; 3) absence of pathogenic copy number variants as revealed by array comparative genome hybridization performed on a clinical basis (using a 135k-feature whole-genome microarray (SignatureChip OS2.0 manufactured for Signature Genomic Laboratories, Spokane, WA, USA) by Roche NimbleGen, Madison, WI, USA); 4) absence of specific changes on brain imaging. The clinical description of the 41 affected individuals is summarized in Table S4. For cases with likely or possibly pathogenic variants, a more detailed clinical description can be found in Text S1.

**Exome capture and sequencing**

Genomic DNA (3 μg) extracted from blood samples were used for exome capture and sequencing at the McGill University and Genome Quebec Innovation Center (Montreal, Quebec, Canada) using the Agilent SureSelect v4 exome capture kit, according to the manufacturer’s recommendations, followed by 100 bp paired-end sequencing of each trio exomes on a single lane of the Illumina HiSeq2000.

**Data analysis**

Exome sequence data processing, alignment (using a Burrows-Wheeler algorithm, BWA-mem), and variant calling were done according to the Broad Institute Genome Analysis Tool Kit (GATK v4) best practices (http://www.broadinstitute.org/gatk/guide/topic?name=best-practices), and variant annotation was done using Annovar [61]. The median coverage of the target bases was 135× with 95% of the target bases being covered ≥10×. We focused on variants affecting the exonic regions and consensus splice site sequences (defined herein as intronic bases up to positions +3 and +6 from the exon boundaries). Only variants whose positions were covered at ≥10× and supported by at least 4 variant reads constituting ≥20% of the total reads for each called position were retained. This typically yielded an average of ~22,000 variants. This variant list was subsequently reduced to an average of ~500 rare variants by filtering out those that are present in ≥0.5% of in-house exome data sets (n = 600) from unrelated projects, as well as variants present in the 1000 Genome or in the Exome Variant Server (EV: http://evs.gs.washington.edu/EVS/) with minor allele frequencies (MAF) ≥0.5%. Putative DNMs (typically <10/exome) were then extracted from the rare variant list by further excluding those that were present in the exomes of the parents. The sequencing reads carrying putative DNMs were inspected visually in each trio, using the Integrative Genomics Viewer (IGV) [62], to exclude obvious false positives. All putative DNMs were validated by bidirectional Sanger sequencing in the corresponding trio.

### Supporting Information

**Figure S1** Impact of the NANS synonymous de novo mutation in exon 4, c.603G>A (p. =) (NM_018946.3), identified in patient 143.91 on exon splicing. (DOC)

**Table S1** Confirmed DNMs identified in this study. (XLSX)

**Table S2** Genes included in the physical protein-protein interaction network analysis. (XLSX)

**Table S3** Inherited rare variations identified in the 41 probands of this study. (XLSX)

**Table S4** Clinical phenotypes of the 41 affected individuals. (XLSX)

**Text S1** Detailed clinical description of the patients with likely and possibly pathogenic DNMs identified in this study. (DOCX)

### Acknowledgments

We are grateful to the Génome Québec and McGill Innovation Center Massive Parallel Sequencing team for the exome sequencing service and to the bioinformatics analysis team of Réseau de Médecine Génétique Appliquée du Québec (RMGA) for the primary analysis of the sequence data. We thank Philippe Lemay for useful discussions. We are also grateful to all the families that participated in this study.

### Author Contributions

Conceived and designed the experiments: FFH GAR JLM. Performed the experiments: JMCC LP HD CN AF. Analyzed the data: FFH MS AA DS OD EH ADL AVP. Contributed reagents/materials/analysis tools: CM OD EH ADL AVP. Contributed to experiments: JMCC LP HD CN AF. Analyzed the data: FFH MS AA DS OD EH ADL AVP. Contributed reagents/materials/analysis tools: CM OD EH ADL AVP. Contributed to the interpretation: FFH MS OD EH ADL AVP. Wrote the paper: FFH JLM. Recruited patients and provided clinical information: CM MS OD EH ADL AVP. Critically reviewed the paper: FFH JLM. Recruited patients and provided clinical information: CM MS OD EH ADL AVP. Contributed reagents/materials/analysis tools: CM OD EH ADL AVP.

### References

1. Ropers HH (2010) Genetics of early onset cognitive impairment. Annu Rev Genomics Hum Genet 11: 161–187.
2. de Ligt J, Willemsen MH, van Bon BW, Kleefstra T, Yntema HG, et al. (2012) Diagnostic de novo exome sequencing in persons with severe intellectual disability. N Engl J Med 367: 1921–1929.
3. Hamdan FF, Gautier J, Araki Y, Lin DT, Yoshizawa Y, et al. (2011) Excess of de novo deleterious mutations in genes associated with glutamatergic systems in nonsyndromic intellectual disability. Am J Hum Genet 80: 306–316.
4. Rausch A, Waeckerle D, Graf E, Wieland T, Endele S, et al. (2012) Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet 380: 1674–1682.
5. Visser LE, de Ligt J, Gilissen C, Janssen I, Stroehouwer M, et al. (2010) A de novo paradigm for mental retardation. Nat Genet 42: 1109–1112.
6. Epigen Consortium, Epilepsy Phenome-Genome, Allen AS, Berkovic SF, Coster P, et al. (2013) De novo mutations in epileptic encephalopathies. Nature 501: 217–221.
7. InnesIF, Rosenm M, Levy D, Wang Z, Hakker I, et al. (2012) De novo gene disruptions in children on the autistic spectrum. Neuron 74: 285–299.
8. Neale BM, Kou Y, Liu L, Mal’ayan A, Samocha KE, et al. (2012) Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature 485: 242–245.
9. O’Roak BJ, Vives I, Girirajan S, Karak D, Krumm N, et al. (2012) Sporadic autism exomes reveal a highly interconnected protein network of de novo_pathogenic variants. Nature 485: 246–250.
10. Sanders SJ, Marthas MT, Gupta AR, Murdoch JD, Raubeson MJ, et al. (2012) De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485: 237–241.
11. Xu B, Ionita-Laza I, Roo J, Boone B, Woodrick S, et al. (2012) De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. Nat Genet 44: 1365–1369.
12. Hoyer J, Eicki AB, Endele P, Popp B, Wett C, et al. (2012) Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. Am J Hum Genet 90: 565–572.
13. Carville GL, Heavin SB, Yerid SC, McMahon JM, O’Roak BJ, et al. (2013) Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. Nat Genet 45: 825–830.
14. Kortum F, Das S, Flath M, Morris-Rosendahl DJ, Stefanovski F, et al. (2011) The core FOXG1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. J Med Genet 48: 396–406.
15. Willemsen MH, Nijhof B, Fenckova M, Nillesen WM, Bongers EM, et al. (2013) The core phenotype of osteogenesis imperfecta type V is caused by GYPC mutations. Am J Hum Genet 93: 565–572.
16. Sanders SJ, Marthas MT, Gupta AR, Murdoch JD, Raubeson MJ, et al. (2012) De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485: 237–241.
17. Xu B, Ionita-Laza I, Roo J, Boone B, Woodrick S, et al. (2012) De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. Nat Genet 44: 1365–1369.
18. Hoyer J, Eicki AB, Endele P, Popp B, Wett C, et al. (2012) Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. Am J Hum Genet 90: 565–572.
19. Carville GL, Heavin SB, Yerid SC, McMahon JM, O’Roak BJ, et al. (2013) Targeted resequencing in epileptic encephalopathies identifies de novo mutations in CHD2 and SYNGAP1. Nat Genet 45: 825–830.
38. Venetucci L, Denegri M, Napolitano C, Priori SG (2012) Inherited calcium
37. Wentzel C, Rajcan-Separovic E, Ruivenkamp CA, Chantot-Bastaraud S, Metay
36. Zhang F, Yu X (2011) WAC, a functional partner of RNF20/40, regulates
35. Isken O, Maquat LE (2007) Quality control of eukaryotic mRNA: safeguarding
34. Ballif BC, Rosenfeld JA, Traylor R, Theisen A, Bader PI, et al. (2012) High-
33. Lemke JR, Hendrickx R, Geider K, Laube B, Schwake M, et al. (2014) GRIN2B
32. Endele S, Rosenberger G, Geider K, Popp B, Tamer C, et al. (2010) Mutations
31. Palumbo O, Fichera M, Palumbo P, Rizzo R, Mazzolla E, et al. (2014) TBR1 is
30. Saitsu H, Nishimura T, Muramatsu K, Kodera H, Kumada S, et al. (2013) De
29. Haack TB, Hogarth P, Kruer MC, Gregory A, Wieland T, et al. (2012) Exome
28. Need AC, Shashi V, Hitomi Y, Schoch K, Shianna KV, et al. (2012) Clinical
27. Hoischen A, van Bon BW, Gilissen C, Arts P, van Lier B, et al. (2010) De novo
26. Hoischen A, van Bon BW, Gilissen C, Arts P, van Lier B, et al. (2010) De novo
25. van Haelst MM, Monroe GR, Duran K, van Binsbergen E, Breur JM, et al.
24. Asadollahi R, Oneda B, Sheth F, Azzarello-Burri S, Baldinger R, et al. (2013)
23. Kleefstra T, Kramer JM, Nevelting K, Willemse MH, Koemans TS, et al. (2012) Disruption of an EHMT1-associated chromatin-modification module causes intellectual disability. Am J Hum Genet 91: 73–82.
22. Takano K, Liu D, Tarpey P, Gallant E, Lam A, et al. (2012) An X-linked channelopathy with cardiomegaly due to a CLCN2 mutation enhancing ryanodine receptor channel activity. Hum Mol Genet 21: 4897–4907.
21. Ma X, Bao J, Adelstein RS (2007) Loss of cell adhesion causes hydrocephalus in nonmuscle myosin II-B-ablated and mutated mice. Mol Biol Cell 18: 2305–2312.
20. Ma X, Kascamento S, Hara Y, Adelstein RS (2004) A point mutation in the motor domain of nonmuscle myosin II-B impairs migration of distinct groups of neurons. Mol Cell Biol 15: 2568–2579.
19. Tuzovic I, Yu L, Zeng W, Li X, Lu H, et al. (2013) A human de novo mutation in MYH16 phenocopies the loss of function mutation in mice. Rare Dis 1: e2014.
18. Muto S, Senda M, Akiy Y, Sato L, Suzuki T, et al. (2007) Relationship between the structure of SET/TAF-beta/ENHAT and its histone chaperone activity. Proc Natl Acad Sci U S A 104: 4285–4290.
17. Lehnart SE, Mongillo M, Bellinger A, Lindegger N, Chen BX, et al. (2008) Leaky Ca2+ release channel/Ryanodine receptor 2 causes seizures and sudden cardiac death in mice. J Clin Invest 118: 2230–2245.
16. Johnson JT, Tester DJ, Bass NE, Ackerman MJ (2010) Cardiac channel molecular autophagy for sudden unexpected death in epilepsy. J Child Neurol 25: 916–921.
15. Lee NH, Mengillo M, Bellingier A, Lindegger N, Chen BX, et al. (2006) Leaky Ca2+ release channel/Ryanodine receptor 2 causes seizures and sudden cardiac death in mice. J Clin Invest 118: 2230–2245.
14. Takano K, Liu D, Tarpey P, Gallant E, Lam A, et al. (2012) An X-linked channelopathy with cardiomegaly due to a CLCN2 mutation enhancing ryanodine receptor channel activity. Hum Mol Genet 21: 4897–4907.
13. Morita E, Colf LA, Karren MA, Sandrin V, Rodesch CK, et al. (2010) Human
12. McCullough J, Colf LA, Sundquist WI (2013) Membrane fission reactions of the mammalian ESCRT pathway. Annu Rev Biochem 82: 663–692.
11. Tuzovic I, Yu L, Zeng W, Li X, Lu H, et al. (2013) A human de novo mutation in MYH16 phenocopies the loss of function mutation in mice. Rare Dis 1: e2014.
10. Muto S, Senda M, Akiy Y, Sato L, Suzuki T, et al. (2007) Relationship between the structure of SET/TAF-beta/ENHAT and its histone chaperone activity. Proc Natl Acad Sci U S A 104: 4285–4290.
9. Haack TB, Hogarth P, Kruer MC, Gregory A, Wieland T, et al. (2012) Exome sequencing reveals de novo WDR45 mutations causing a phenotypically distinct, X-linked dominant form of NBIA. Am J Hum Genet 91: 1144–1149.
8. Saito H, Nishimura T, Muramatsu K, Kodera H, Kumada S, et al. (2013) De novo mutations in the autophagy gene WDR45 cause static encephalopathy of childhood with neurodegeneration in adulthood. Nat Genet 45: 445–449, e441.
7. Froyen G, Corbett M, Vandewalle J, Jarvela I, Lawrence O, et al. (2008) GeneMANIA prediction server 2013 update. Nucleic Acids Res 41: W115–122.
6. Ballif BC, Rosenfeld JA, Traylor R, Theisen A, Bader PI, et al. (2012) High-
5. Jones MW, Errington ML, French PJ, Fine A, Blas TV, et al. (2001) Identification of the immediate early gene Zif268: mechanisms and gene dosage-dependent function in synaptic plasticity and memory formation. Hippocampus 12: 570–577.
4. Munton RP, Vizi S, Mannuy, DM (2004) The role of protein phosphatase-1 in the modulation of synaptic and structural plasticity. FEBS Lett 567: 121–128.
3. Stuchell-Brereton MD, Skalicky JJ, Kieffer C, Karren MA, Abe T, Copeland TD, et al. (2001) Identification and characterization of SEB, a novel protein that binds to the acetyltransferase GCN5 in human cells. Mol Cell 41: 384–397.
2. Takano K, Liu D, Tarpey P, Gallant E, Lam A, et al. (2012) An X-linked channelopathy with cardiomegaly due to a CLCN2 mutation enhancing ryanodine receptor channel activity. Hum Mol Genet 21: 4897–4907.
1. van Haelst MM, Monroe GR, Duran K, van Binsbergen E, Breur JM, et al. (2010) De novo mutations in the autophagy gene WDR45 cause static encephalopathy of childhood with neurodegeneration in adulthood. Nat Genet 45: 445–449, e441.