ORIGINAL ARTICLE

**GATAD2B loss-of-function mutations cause a recognisable syndrome with intellectual disability and are associated with learning deficits and synaptic undergrowth in Drosophila**

Marjolein H Willemsen,1 Bonnie Nijhof,1,2,3 Michaela Fenckova,1,2,3 Willy M Nillesen,¹ Ernie M H F Bongers,¹ Anna Castells-Nobau,1,2,3 Lenke Asztalos,4 Erika Viragh,5 Bregje W M van Bon,¹ Emre Tezel,1,2,3 Joris A Veltman,1,2,6 Han G Brunner,1,6 Ernie M H F Bongers,1 Anna Castells-Nobau,1,2,3 Lenke Asztalos,4 Erika Viragh,5 Bregje W M van Bon,¹ Emre Tezel,1,2,3 Joris A Veltman,1,2,6 Han G Brunner,1,6 Bert B A de Vries,1,6 Joep de Ligt,1,2,6 Helger G Yntema,¹ Hans van Bokhoven,1,2,3 Bertrand Isidor,7 Cédric Le Caignec,7 Elsa Lorino,8 Zoltan Asztalos,4,5,9 David A Koolen,¹ Lisenka E L M Vissers,¹,2,6 Annette Schenck,1,2,3 Tjitske Kleefstra¹,3,6

**ABSTRACT**

**Background** GATA zinc finger domain containing 28 (GATAD2B) encodes a subunit of the MeCP1-Mi-2/nucleosome remodelling and deacetylase complex involved in chromatin modification and regulation of transcription. We recently identified two de novo loss-of-function mutations in GATAD2B by whole exome sequencing in two unrelated individuals with severe intellectual disability.

**Methods** To identify additional individuals with GATAD2B aberrations, we searched for microdeletions overlapping with GATAD2B in inhouse and international databases, and performed targeted Sanger sequencing of the GATAD2B locus in a selected cohort of 80 individuals based on an overlap with the clinical features in the two index cases. To address whether GATAD2B is required directly in neurones for cognition and neuronal development, we investigated the role of Drosophila GATAD2B orthologue simjang (simj) in learning and synaptic connectivity.

**Results** We identified a third individual with a 240 kb microdeletion encompassing GATAD2B and a fourth unrelated individual with GATAD2B loss-of-function mutation. Detailed clinical description showed that all four individuals with a GATAD2B aberration had a distinctive phenotype with childhood hypotonia, severe intellectual disability, limited speech, tubular shaped nose with broad nasal tip, short philtrum, sparse hair and strabismus. Neuronal knockdown of Drosophila GATAD2B orthologue, simj, resulted in impaired learning and altered synapse morphology.

**Conclusions** We hereby define a novel clinically recognisable intellectual disability syndrome caused by loss-of-function of GATAD2B. Our results in Drosophila suggest that GATAD2B is required directly in neurones for normal cognitive performance and synapse development.

**INTRODUCTION**

Intellectual disability (ID) is a group of disorders with an extremely heterogeneous clinical and genetic presentation. More than 500 ID genes have been identified and many more await discovery. This large number of ID genes is believed to converge onto a limited number of common underlying pathways and processes.1 Several ID genes encode proteins that are involved in chromatin modification.2,3

Since the recent advent of next generation sequencing technology, whole exome sequencing (WES) has been successfully applied to the identification of genes for clinically established ID syndromes.3–6 In addition, family based WES was successful in elucidating causative de novo gene mutations in sporadic individuals who do not present with a recognisable syndrome.7–9 However, exploration of the pathogenicity of mutations in genes not previously associated with ID remains challenging. To establish a conclusive molecular diagnosis, it is therefore required to detect mutations in the same candidate genes in additional individuals with similar phenotype.10 Moreover, additional evidence and insights into functional properties of novel genes are desirable and can be obtained through studies in cell or animal models.

By application of trio based WES, we recently reported the identification of 22 candidate genes for ID.9 Among these was the GATA zinc finger domain containing 2B (GATAD2B) gene (NM_020699.2). In this gene, we identified two loss-of-function mutations in two unrelated individuals (c.584dupT; p.(Asn195fs) and c.1408 C>T; p.(Gln470fs)). GATAD2B encodes p66beta, a subunit of the transcription repressor complex MeCP1-Mi-2/nucleosome remodelling and deacetylase (NuRD), responsible for silencing of methylated DNA by nucleosome remodelling and histone deacetylation.11,12

Here, we report an additional individual with a disruptive GATAD2B mutation (c.563 566del; p. (Gln190fs)), representing the third loss-of-function mutation in this gene. The mutation was revealed by direct Sanger sequencing in a selected cohort of 80 individuals with overlapping features comprising ID, childhood hypotonia and an abnormal shape of the nose (including a tubular shape, prominent and

---


broad base of the nose). In addition, we found a fourth individual with a 240 kb de novo microdeletion encompassing GATAD2B. Comparison of the phenotype of all four individuals revealed a remarkable overlap in clinical presentation.

Modelling of the Drosophila orthologue of GATAD2B loss-of-function confirms its role in cognition and shows a critical role of GATAD2B in synapse development. Together, our data establish a novel ID syndrome. It adds to the growing list of ID conditions that are caused by mutated genes involved in chromatin remodelling that can shed light onto the epigenetic control of cognition.

METHODS

Patients

Individuals 1 and 2 were ascertained through family based WES studies recently reported by our group7 8 (trio 4 in Vissers et al7 and trio 69 in Ligt et al8 respectively). The GATAD2B mutation in individual 1 was noticed in retrospect upon reanalysis of the sequencing data, after the detection of the mutation in individual 2.

Subsequently, we searched for individuals with small microdeletions overlapping with GATAD2B in our inhouse database and international databases, including the database of the European Cytogenetists Association Register of Unbalanced Chromosome Aberrations (ECARUCA) and the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (Decipher).

In addition, we selected a cohort of 80 individuals guided by phenotypic overlap with individuals 1 and 2. This cohort was selected based on the presence of severe ID (IQ ≤50), limited speech ability, childhood hypotonia and facial features including abnormal shaped nose (large/prominent nose, full nose tip, tubular shaped/pear-shaped nose, broad base of the nose). All individuals had been referred to the Department of Human Genetics of the Radboud University Medical Centre in Nijmegen, the Netherlands, for genetic diagnostic evaluation of unexplained ID/developmental delay. Their parents/legal representatives consented to this study. The study was approved by the local ethical committee.

Mutation screening

Targeted Sanger sequencing of GATAD2B (NM_020699.2) was performed using standardised methods. Primers are available upon request. The mutation in individual 4 and the low grade mosaic in her mother were confirmed with a second independent primer pair.

Fly stocks and maintenance

Fly stocks were kept on standard Drosophila diet (cornmeal/sugar/yeast) at 25°C and 45%-60% humidity at 12 h light-dark cycle. Flies were reared at 25°C, 70% humidity for habituation experiments and real-time PCR, and at 28°C, 60% humidity to evaluate synapse morphology. Simjjang (Simji) (CG32067) is the Drosophila orthologue of GATAD2B (Ensemble, see also Results section). An inducible RNA interference (RNAi) line against simji (vdrKK100285, a line with the highest specificity score, s19=1) and its corresponding control line (vdrKK60100) were obtained from Vienna Drosophila RNAi Center (http://www.vdrc.at). RNAi was induced by the UAS-Gal4 system using a panneuronal elav-Gal4 driver or a ubiquitous actin-Gal4 driver obtained from the Drosophila Bloomington Stock Center (http://flystocks.bio.indiana.edu/). Line w^{1118}; UAS-Dicer2; elav-Gal4 driver was used to study synapse morphology, and a w^{1118}; actin-Gal4/CyOGFP driver to generate homogeneous knockdown material for real-time PCR.

Light-off jump reflex habituation

The light-off jump reflex habituation assay was performed as previously described13 with minor adaptations to the protocol. Briefly, 3–7-day-old individual male flies were tested for jump response in two independent 16-unit light-off jump reflex habituation systems. A total of 32 flies (16-flies/system) were simultaneously exposed to series of 100 short (15 ms) light-off pulses with a 1 s interval between the pulses. The noise amplitude of wing vibration following every jump response was recorded for 500 ms after the start of pulse and a carefully chosen threshold was applied to distinguish the jump response. Data were collected and analysed by custom-made Labview Software (National Instruments). High initial jumping response to light-off pulse decreased with growing number of trials and flies were considered habituated when they failed to jump in five consecutive trials (non-jump criterion). Habituation was scored as the number of trials required to reach the non-jump criterion (Trials To Criterion (TTC)). Mean TTC values of eight independent groups of 16 flies (tested on four different days) were compared with mean TTC values of control flies using one-way Analysis of variance (ANOVA) with correction for different experimental day and system.

Drosophila synapse morphology

Type 1b neuromuscular junctions (NMJs) at muscle 4 were analysed after dissection of L3 larvae and fixation in 3.7% paraformaldehyde (PFA) for 30 min. Preparations were colabelled for bruchpilot (brp) and discs large 1 (dlg1). Brp was visualised using the primary antibody n582 (1:125) (Developmental Studies Hybridoma Bank, University of Iowa) applied overnight at 4°C, and a secondary Alexa 488-labelled goat-antimouse antibody (1:500) (Invitrogen). Discs large was visualised using the primary antibody anti-dlg1 (1:25) (Developmental Studies Hybridoma Bank) in combination with the Zenon Alexa Fluor 568 Mouse IgG1 labelling kit (Invitrogen). NMJ pictures were obtained using a Leica automated brightfield multi-colour epifluorescent microscope. Individual synapses were imaged and the muscle area, NMJ area, perimeter, length, branching pattern and amount of active zones were quantitatively assessed using an inhouse developed macro.

Analysis of simji mRNA levels by real-time PCR

Total RNA from 3rd instar larvae was isolated using RNeasy Lipid Tissue Mini Kit (Qiagen). RNA was treated with DNase (DNAfree Kit, Ambion). First strand cDNA synthesis was performed using the Moloney Murine Leukemia Virus (M-MLV) Reverse Transcriptase (Life Technologies) and oligo(dT) primer. Gene expression was analysed by real-time PCR (7900HT Fast Real-Time PCR system, Applied Biosystems). PCR reactions were performed in a volume of 25 μl containing 150 nM primers and GoTag Green Mastermix (Promega). Primer sequences used for amplification of simji were 5’-CAGCACCATTCCGGTTGTG-3’ (forward primer) and 3’-GCCTTAGTGGCTTTCA-3’ (reverse primer). A PoI1 primer sequences were 5’-TCAGAGTCCGCGTAA-3’ (forward primer) and 5’-TGTCACAAGTGGCTTTCA-3’ (reverse primer).
RESULTS

Clinical phenotype description

We here report the detailed clinical description of four affected individuals. Individual 1 (figure 1A–D) was born after 36 weeks and 5 days of pregnancy with a normal birth weight and no major complications upon delivery. At 2 years of age, her parents consulted a paediatrician because of global developmental delay. She could sit and crawl at the age of 16 months and stand at the age of 18 months. Hypotonia was noticed. After the age of 2 years she spoke her first words. An ophthalmologist was consulted because of strabismus. At the age of 3 years and 8 months she was able to speak about 30 single words, but her comprehension of language was at a higher level. She had a normal height (100 cm, 30th centile) and head circumference (47.5 cm, 10th centile). Her facial features included thin blond hair, narrow palpebral fissures, periorbital fullness, a tubular shape of the nose with full nasal tip, a short philtrum, thin upper lip, broad mouth and grimacing facial expression. Three years later, at the age of 6 years and 7 months she had severe developmental delay with verbal skills more severely impaired than motor skills. Her height was normal (118.5 cm, 16th centile) and she had a low to normal head circumference (48 cm, 2nd centile). Long fingers and toes with broad distal phalanges were observed.

Individual 2 (figure 1E–H) was born with intrauterine growth retardation. Her birth weight was 2780 g at 42 weeks of pregnancy (<2nd centile). At 3 months of age a developmental delay was noticed. She could walk independently at the age of 3 years. From 8 years on, she could speak single words. Behaviour was characterised by tics and wandering during the night. Her pain threshold was elevated. Vision was impaired due to hypoplasia of the optic nerve. Upon clinical evaluation at the age of 34 years she was diagnosed with severe ID. At that time, height and head circumference were normal (both >16th centile). She had thin blond hair, deeply set eyes, narrow and upward slanting palpebral fissures, strabismus, a tubular shape of the nose with broad nasal tip, and a large mouth with short philtrum and thin upper lip. She grimaced. Her fingers were long, thin and slightly tapering.

Individual 3 (figure 1I,J) was born after an uncomplicated pregnancy of 40 weeks duration, with a low-normal birth weight...
and a normal head circumference. In the neonatal period she presented with hypotonia and feeding difficulties. Her psychomotor development was delayed. She learnt to walk without support at the age of 2 years and 9 months. At the age of 3 years speech was severely delayed with only two to three disyllabic words. Her behaviour was characterised by low frustration tolerance. Medical problems included hypermetropia and strabismus and she had shown one episode of absence epilepsy. At the age of 3 years she had a height of 90 cm (5th centile), and a head circumference of 51 cm (84th centile). Facial dysmorphism included hypertelorism, a broad forehead, a broad and flat nasal bridge and a full square tip of the nose. She had thin, blond hair. Medical problems included hypermetropia and strabismus and she had shown one episode of absence epilepsy. At the age of 3 years she had a height of 90 cm (5th centile), and a head circumference of 51 cm (84th centile). Facial dysmorphism included hypertelorism, a broad forehead, a broad and flat nasal bridge and a full square tip of the nose. She had thin, blond hair. Her fingers were long and she had fleshy hands.

**Mutation detection**

Individual 1 showed no abnormalities on 250K Singly Nucleotide Polymorphism (SNP) array analysis, screening for Fragile X syndrome, a metabolic screen in blood and urine, and MRI imaging of the brain. Family based WES revealed a de novo frameshift mutation in GATAD2B (c.584dup; p. (Asn195fs); figure 2A). This mutation is located in the central part of the gene close to the conserved region 1 (CR1) domain (figure 2D). Likewise, individual 2 did not show any abnormality on 250K SNP array analysis and a metabolic screen. Methylation tests of the Angelman syndrome related region were normal. Family based exome sequencing revealed a de novo mutation in GATAD2B (c.1408 C>T; p.(Gln470*); figure 2B). This mutation is located in the conserved region 2 (CR2) domain (figure 2D).
Individual 4 was previously investigated for mutations in the genes FMR1, MECP2, TCF4, RA1 and UBE3A, all showing normal results. Methylation tests of the Angelman syndrome related region and a metabolic screen gave also normal results. Conventional G-banded karyotyping yielded a maternally-inherited apparently balanced translocation between chromosomes 1q21.3 and 9q13. Further 250K SNP array analysis did not show any imbalances of these regions. Targeted sequencing of GATAD2B based on phenotypic overlap with the first two individuals led to the identification of a frameshift mutation (c.565_566del; p.(Gln190fs); figure 2C). This mutation is located at the border of the CR1 domain (figure 2D). The healthy mother carried the GATAD2B mutation in a low mosaic level, which was estimated to be below 10% in peripheral blood lymphocytes (figure 2C). The mother is highly educated and works as a social worker. She has no facial dysmorphism.

Copy number variation detection

Array comparative genomic hybridisation (CGH) analysis in individual 3 by a 60 K Agilent array, using the International Standards for Cytogenomic Arrays (ISCA) design, revealed a 240 kb de novo deletion of chromosome 1q21.3 encompassing 10 genes, including GATAD2B (chr1: 153 893 110–154 132 780, Hg19). The distal breakpoint disrupts GATAD2B. The deleted region encompasses no known disease associated genes. The deletion was confirmed by Fluorescent in situ Hybridisation (FISH) with the RP11-422P24 probe. Segregation analysis in the parents was done by FISH analysis and showed that the deletion had occurred de novo.

Analysis of GATAD2B orthologue in Drosophila

To obtain independent evidence for the involvement of GATAD2B in the ID phenotype of the described individuals, we decided to study its function in an intact nervous system, using Drosophila as a model. The Drosophila genome contains a single orthologue of the two closely related human GATAD2A and GATAD2B genes, named simjang (simj). Drosophila simj has been shown to modify Wnt signalling and, like both its human counterparts, to associate with the MeCP1-Mi-2–NuRD repressor complex.14

Neuronal knockdown of Drosophila GATAD2B orthologue results in impaired learning

Null mutations of Drosophila simj have previously been reported to be lethal,14 which precluded assessment of cognitive function in adult flies. We therefore targeted expression of simj specifically in Drosophila neurones using the UAS-Gal4 system in combination with inducible RNAi15 and examined the role of simj in habituation. Habituation is a form of non-associative learning where the probability of a behavioural response decreases with repeated presentations of a stimulus.16 Habituation was previously shown to be defective in classic learning and memory mutants and in a Drosophila model of ID.2 17 18

To induce a panneuronal knockdown of simj, flies carrying an inducible UAS-RNAi insertion against simj (vdrcKK100285) were mated to flies carrying the neuronal promoter element elav-Gal4 driver. The promoter line further carried one copy of UAS-dicer2 to enhance the RNAi-mediated knockdown and two insertions of GMR-wIR to reduce the eye colour as required in our assay. Male progeny of knockdown (simj-RNAi) and control flies was exposed to 100 short light-off stimuli (trials) at 1 s inter-trial intervals and scored for a jump response. Flies were considered to have habituated once they failed to jump in five consecutive trials (no-jump criterion). Habituation was scored as the number of trials required to reach the no-jump criterion (TTC). Both genotypes showed wt-like initial jump responses. Control flies quickly responded to the light-off stimuli and reduced jumping. We found that simj-RNAi flies habituate slower and maintain a higher jump response throughout the entire course of the experiment (figure 3A). The mean TTC of simj-RNAi flies was 1.7-fold higher compared with their genetic background controls (figure 3B, p=0.007), validating a significant deficit in habituation.

GATAD2B is required for synaptic development in Drosophila

Synaptic connectivity is essential for learning and for other cognitive processes.1 We therefore addressed a possible function of simj in synaptic development of the larval NMJ, a well-established synaptic model system that shares major features with central excitatory synapses in the mammalian brain19 and has been successfully used to investigate human ID disorders.20–22

As in our habituation experiment, we investigated the Drosophila NMJ architecture upon panneuronal knockdown of simj. The synaptic and subsynaptic organisation was visualised by coimmunolabelling against dlg1 (α-dlg1), a major scaffolding component of larval NMJs and member of the membrane-associated guanylate kinase subfamily, and anti-brp (α-brp, nc82), an integral part of active zones.22

Neurone-specific simj knockdown resulted in an NMJ undergrowth phenotype (figure 4A,B), with a consistent decrease in synaptic area (p=0.0028), perimeter (p=9.5×10^{-5}) and length...
Figure 4 Synapse morphology of neuromuscular junctions (NMJs) was studied in Simjneg (simj) knockdown flies (UAS-simjRNAi100285/UAS-Dicer2; elav-Gal4/+) (A). NMJs were visualised with coimmunolabelling against discs large 1 (A, upper panel) and brp (A, middle panel) and quantitatively assessed by computer-assisted analysis (A, lower panel and B). SimjRNA interference (RNAi) NMJs show a decrease in NMJ area (n=27, p=0.0028), perimeter (n=27, p=0.005), length (n=28, p=5.7e-008), the amount of branches (n=28, p=0.041) and branching points (n=28, p=0.021) and the amount of active zones (n=25, p=0.0010). The area of the muscle remained normal (n=28, p=0.077). For this plot, red bars represent the mean of each parameter of SimjRNAi NMJs normalised by the mean of each parameter of the control set (black bars). Error bars indicate the normalised standard error of the mean, p, p values (two-sided t test); n, number of quantified NMJs. (*) indicates a significant difference (p<0.05), (**) indicates a significant difference (p<0.01), (***) indicates a significant difference (p<0.001). Access the article online to view this figure in colour.

Therefore performed additional FISH analysis with the region 1q23.1 specific probe RP11-216N14 that covers the GATAD2B gene. This showed in all 29 evaluated cells a normal pattern, indicating that the translocation breakpoint is located distally from the GATAD2B region. Therefore, we concluded that this does not support a correlation between the mutation in GATAD2B and the balanced translocation.

Complementary functional studies of the Drosophila GATAD2B orthologue simj demonstrated a role for the evolutionarily conserved GATAD2B gene family in neurodevelopmental processes. The identified habituation defect in simj-RNAi flies demonstrates that the GATAD2B orthologue is required for non-associative learning. Furthermore, evaluation of synaptic morphology revealed an NMJ undergrowth phenotype with reduced number of active zones, suggesting a critical role for simj and GATAD2B in synaptic growth and function. Whether these may contribute to or cause the observed learning defects in flies and cognitive deficit in the humans remains to be determined.

Mutations in several recently identified genes involved in chromatin modification give rise to ID syndromes, such as the 17q21.31 microdeletion/Koolen-De Vries syndrome (KDVS (MIM 610443)),23 24 Coffin-Sirius syndrome13 25 (CSS (MIM135900)), Nicolaides-Baraitser syndrome26 27 (NCBRS (MIM 601358)), Kleefstra syndrome28 (MIM 610253), Wiedemann-Steiner syndrome29 (WDSTS (MIM 605130)) and Ohdo syndrome Say-Barber-Biesecker variant30 (SBBYSS (MIM 603736)).

GATAD2B encodes p66beta, which is a subunit of the transcription repressor complex McCP1-Mi2-NuRD that silences methylated DNA by nucleosome remodelling and histone deacetylation. This enzyme complex also comprises its close parologue p66alpha (encoded by GATAD2A), the histone deacetylases HDAC1 and HDAC2, two histone binding proteins RhAp46 and RhAp48, the methyl binding domain protein 3 (MBD3), two histone modifier proteins, MTA1 and MTA2, and nucleosome remodelling factor Mi-2.11 It is thus conceivable that additional genes from the McCP1-Mi2-NuRD complex, such as GATAD2A, are involved in phenotypes overlapping with the here defined novel syndrome caused by GATAD2B mutations. P66beta and its parologue, p66alpha, function synergistically and recruit the Mi-2–NuRD complex to its target sites. They interact with methyl-CpG bound MBD2 and with

New disease loci
non-acetylated histones to assemble in the so-called McCP1 repressor complex.\textsuperscript{1,1} It has been shown that p66alpha and p66beta, and particularly their highly conserved CR1 and CR2 domains, are crucial for complex formation and mediated gene silencing.\textsuperscript{12} Furthermore, homozygous loss-of-function of p66alpha in mice resulted in an embryonic lethal phenotype with severe global malformations, growth retardation and necrosis. Heterozygous mice were viable and appeared normal, but with severe global malformations, growth retardation and necrosis.\textsuperscript{2}

We recently used the powerful strategy of combining human genetic studies and Drosophila modelling to provide a novel chromatin remodelling module that underlies Kleefstra syndrome spectrum.\textsuperscript{3} Similar to that, we find a strikingly high conservation between human and Drosophila NuRD complexes, with all complex components being present in fly.\textsuperscript{3,2} As additional variants in human NuRD complex genes will arise, it will be straightforward to validate their significance. Our data add GoFAD2B, and possibly the whole McCP1-M12-NuRD complex, to the growing list of ID genes involved in chromatin remodelling.

**Author affiliations**

1Department of Human Genetics, Radboud University Medical Centre, Nijmegen, The Netherlands
2Nijmegen Center for Molecular Life Sciences, Radboud University Medical Centre, Nijmegen, The Netherlands
3Donders Institute for Brain, Cognition and Behavior, Radboud University Medical Centre, Nijmegen, The Netherlands
4Department of Genetics, Aktogen, University of Cambridge, Cambridge, UK
5Institute of Biochemistry, Biological Research Center, Hungarian Academy of Sciences, Szeged, Hungary
6Service de Génétique Médicale, CHU Nantes, Nantes, France
7Service de Pédiatric, CHU Nantes, Nantes, France
8AEG Interdisciplinary Research Institute of Genetics and Molecular Medicine, Szeged, Hungary

**Acknowledgements** We thank the participating individuals and their families. We also thank Suzanne Vloet-Keijzer and Martine Ruiterkamp for their technical assistance with the targeted Sanger sequencing.


