Journal of Genetics (2021)100:61
https://doi.org/10.1007/s12041-021-01308-z

Tissue-specific DNase I footprint analysis confirms the association of *GATAD2B* Q470* variant with intellectual disability

VIDYA NIKAM[1]* and NAUSHAD SHAIK MOHAMMAD[2]

[1]*Department of Chemistry, Savitribai Phule Pune University, Pune 411 007, India*
[2]*Biochemical Genetics, Sandor Speciality Diagnostics, Banjara Hills, Hyderabad 500 034, India*
*For correspondence. E-mail: vidyanikam63@gmail.com.

Received 8 March 2021; revised 19 April 2021; accepted 4 May 2021

Abstract. Intellectual disability (ID) is a neurodevelopmental disorder in which genetics play a key aetiological role. GATA zinc finger domain-containing 2B (*GATAD2B*) gene encodes a zinc-finger protein transcriptional repressor which is a part of the methyl-CpG binding protein-1 complex. Pathogenic variants in this gene are linked to ID, dysmorphic features, and cognitive disability. To date, only 18 cases are reported worldwide and only one case is reported from India. A 12-year-old girl presented with a heterozygous nonsense variation in exon 8 of the *GATAD2B* gene (chr1:153785737G>A). She has severe ID and significant delayed developmental milestones along with clinical features including broad arched eyebrows, low-set ears, a bulbous nose tip, thin upper lip, and wide mouth with downturned corners. This is the second report of a heterozygous mutation in the *GATAD2B* gene from India with a novel phenotype. To substantiate the association of *GATAD2B* mutation with ID, we performed DNase I footprint analysis of wild and mutant DNA sequences to establish k-mer binding profile and deduced GATA binding affinity using human ENCODE experimental data of foetal brain. We observed that in the presence of variation, GATA zinc finger domain was altered thus contributing to ID. Our findings support the importance of the *GATAD2B* gene in the study of neurodevelopmental disorders.

Keywords. intellectual disability; *GATAD2B* gene; next-generation sequencing; mutation analysis; genotype–phenotype correlation.

Introduction

Intellectual disability (ID) is a group of disorders characterized by significant impairment of cognitive functions with heterogeneous clinical presentation (Srivastava and Schwartz 2014). It has an estimated prevalence of 1–3%. The aetiology of ID is complex and variable, and the diagnosis poses a formidable challenge worldwide. Genetic factors are thought to cause ID in about 25–30% of cases (Kaufman *et al*. 2010). Since the emergence of next-generation sequencing technology (NGS), more than 800 genes for clinically established ID syndromes have been identified and many more await discovery. Exome sequencing has proven to be a powerful tool for recognizing the molecular basis of nonsyndromic ID.

Several ID genes encode proteins that are involved in chromatin modification (Van Bokhoven 2011). The human GATA zinc-finger domain containing 2B (*GATAD2B*) gene is located on chromosome 1q21.3 (OMIM: 614998). This gene encodes beta-subunit of the transcription repressor complex MeCP1-Mi2/nucleosome remodelling and deacetylase complex, which is involved in chromatin modification and regulation of transcription. Recent studies indicated that loss-of-function mutations in *GATAD2B* cause severe ID with significant phenotypic characteristics. The common features of these cases included childhood hypotonia, limited speech, and dysmorphic features, such as tubular-shaped nose with broad nasal tip, short philtrum, sparse hair, strabismus, a broad forehead, hypertelorism, periorbital fullness, deeply-set eyes, and narrow palpebral fissures, a short philtrum, and a large, broad mouth (Willemsen *et al.* 2013).

To our knowledge, 18 individuals with *de novo GATAD2B* mutation have been reported in the Western as well as Asian populations (De Ligt *et al.* 2012; Willemsen *et al.* 2013; Hamdan *et al.* 2014; Vanderver *et al.* 2016; Tim-Aroon *et al.*

Supplementary Information: The online version contains supplementary material available at https://doi.org/10.1007/s12041-021-01308-z.

Published online: 23 August 2021

2017; Luo *et al.* 2017; Rabin *et al.* 2018; Ueda *et al.* 2019; Kaur *et al.* 2019). One 20 month-old-male child, heterozygous for *GATAD2B* c.346C>T (p.R116*) was reported from India with developmental delay, tonic-clonic seizures, hyperactivity, broad forehead, hypertelorism, deep-set eyes, depressed nasal bridge, tubular node, bulbous nasal tip, low-set ears, sparse scalp hair and eyebrows, long great toes and tapering fingers (Kaur *et al.* 2019). Here, we report the second case with a heterozygous mutation in the gene by NGS in a girl exhibiting severe ID with a novel clinical phenotype from India. Our study support previous research, adding *GATAD2B* to the list of ID-associated genes.

Materials and methods

Case presentation

We describe expanded phenotype including, broad arched eyebrows, low-set ears, a bulbous nose tip, thin upper lip, and a wide mouth with downturned corners in *GATAD2B* associated ID (figure 1). Tapered fingers and muscular hypotonia were also noticed. The case is a 12-year-old female, the second child of healthy but consanguineous Indian parents with no family history of ID. She has a healthy elder brother and a healthy younger sister. The girl was born at full-term after a normal pregnancy, at 39-week gestation, with normal birth weight and no complications upon delivery. She had childhood hypotonia and her developmental milestones were significantly delayed. She could support her head at 8 months of age; was able to sit independently at 17 months of age; did not crawl and was able to walk unaided at 3 years of age. She could understand simple instructions, although the language delay was obvious because she could only speak a few words and showed a tendency towards hyperactivity. Intellectual test asesed using Vineland Social Maturity Scale at the age of 6 years exhibited severe ID. She studied in a special school and did not present any social problems. At the age of 12, her physical examination revealed mild spasticity of the lower extremities and distinct facial characteristics under the guidance of a clinical geneticist. Her facial features included frontal bossing, narrow palpebral fissure with ocular hypertelorism and strabismus, flat nasal bridge, short philtrum, and thin lips. These features go along with the reported literature. However, the girl had thick black hair with no epicanthal folds. Written informed consent and clinical information were acquired from the parents. Ethical clearance for this study was obtained from the Savitribai Phule Pune University Ethics Committee. The limitation of the study was lack of parental status.

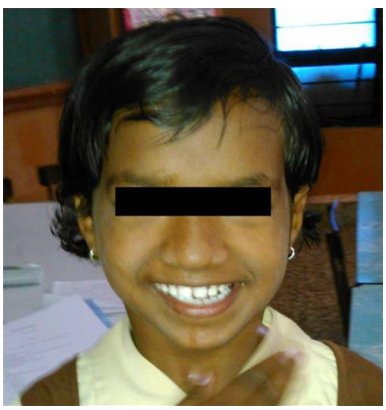

Figure 1. Clinical presentation of the case report.

Metabolic study

Amino acids and acylcarnitine profiles were analysed using tandem mass spectrometry at 12 years of age to detect metabolic abnormalities. These profiles were measured in dried blood spots by using TMS, AB Sciex API-3200 with Agilent 1200 series HPLC. The MS/MS data were processed using the Analyst Software 1.5.1 and Chemview Software 2.0.2. The sample preparation, processing procedures, and analysis were based on previously reported methods (Zytkovicz *et al.* 2001; Nikam *et al.* 2019).

Cytogenetic study

DNA was extracted from peripheral blood sample according to the standard protocol. The sample was further analysed using the technique of multiplex ligation-dependent probe amplification (MLPA) with different probesets (P064-MR1, P096-MR2, and P106-MRX) for the detection of known mutations in ∼80 ID reported genes.

NGS study

DNA was used to perform targeted gene capture using a custom capture kit by NGS technique. The libraries were sequenced to mean >80–100× coverage on the Illumina sequencing platform. The sequences obtained were aligned to the human reference genome (GRCh37/hg19) using the BWA program and analysed using Picard and GATK v. 3.6 to identify variants relevant to the clinical indication. Gene annotation of the variants was performed using the VEP program against the Ensembl release 87 human gene model. Clinically relevant pathogenic variants were annotated using published variants in literature and a set of disease databases: ClinVar, OMIM, GWAS, HGMD, and SwissVar. Common variants were filtered based on allele frequency in 1000 Genome Phase 3, ExAC, EVS, dbSNP147, 1000 Japanese Genome, and our internal Indian population database. The nonsynonymous variants effect was calculated using multiple algorithms such as PolyPhen-2, SIFT, Mutation Taster2, Mutation Assessor, and LRT. Only nonsynonymous and splice site variants found in the clinical exome panel

consisting of 8332 genes were used for clinical interpretation. Candidate variants were validated by Sanger sequencing.

DNase footprint analysis to study the impact of pathogenic variant on GATA binding motif

We used 'Sasquatch' module for a comprehensive k-mer based analysis of DNase footprints in foetal brain using human ENCODE experimental data to determine the changes in GATA binding motif (Schwessinger *et al.* 2017). It is based on the principle that DNase I causes frequent double cuts in open chromatin regions. Deep sequencing allows high-density mapping of the ends of DNA fragments digested by DNase I and with this approach the positions protected by transcription factor binding can be identified. Sasquatch module piles up genomewide cut frequency profiles and identifies the k-mers associated with transcription factor binding. By clubbing this information with the human ENCODE experimental data, tissue-specific transcription factor binding in the presence or absence of the variant was deduced.

Two sequences of same length of 20-bp spanning wild/mutant alleles were used as templates. The organism was selected as human and Dnase fragmentation method was employed. K-mer length was adjusted to 7. ENCODE_UW_fetal brain was selected as the specific tissue. Propensity-based (erythroid) normalization was performed to identify tissue-specific DNase footprints.

Results

The tandem mass spectrometry test of our patient showed low levels of tetradecanoylcarnitine (C14) acylcarnitine. No copy number variations were observed for the targeted genes in any of the samples in the MLPA test, thereby excluding those genes as the causative factor for the condition. NGS test depicted a heterozygous nonsense variation in exon 8 of the *GATAD2B* gene (chr1:153785737G>A; depth: 45×) that results in a stop codon and premature truncation of the protein at codon 470 (p.Gln470Ter; ENST00000368655, rs587776931) was detected (figure 2). The observed variation has previously been reported twice in individuals affected with ID (De Ligt *et al.* 2012; Willemsen *et al.* 2013). No other candidate genes were identified in autosomal recessive or X-linked inheritance models.

DNase footprint analysis showed potential damage to GATA-binding motif due to pathogenic variant

The *GATAD2B* pathogenic variant changes the kmer sequence from CAGCAGG to TAGCAGG, which changes shoulder-to-footprint ratio (SFR) from 1.603 to 1.271 resulting in a total damage of 0.324 and absolute damage of 0.515 (table 1). These SFR patterns suggest that due to C>T substitution, GATA zinc-finger domain is perturbed (figure 3).

Discussion

To date, 18 individuals have been diagnosed with *GATAD2B* mutations from different parts of the world (table 1 in electronic supplementary material at http://www.ias.ac.in/jgenet/). Two cases have been reported with a recurrent *de novo GATAD2B* pathogenic variant (c.1408C>T; p.Gln470Ter) with typical clinical symptoms. Here, we describe the second report from India with the recurrent pathogenic variant and novel clinical phenotype. To date, nine pathogenic variants were reported in this gene and this is the third case with Q470* pathogenic variant (figure 4). To provide mechanistic basis for this association, we performed tissue-specific DNase footprint analysis using the human ENCODE experimental data on foetal brain. The *GATAD2B* pathogenic variant has resulted in the disruption of GATA zinc-finger domain.

Ligt *et al.* (2012) first reported this *de novo* mutation in a child with severe developmental delay, delayed motor milestones, limited speech, and overlapping facial features that are a consistent finding in our case study. Following Willemsen *et al.* (2013), who briefly described the particular facial features of the child as tubular nose with a broad tip, deeply set eyes, a broad forehead, a short philtrum, a broad mouth, a grimacing facial expression, strabismus, and long fingers. Our case did show some overlapping characteristics like a tubular nose, broad nasal tip, frontal bossing, and strabismus. However, proband showed some novel features different from those of previously reported individuals. Proband displayed thick black hair without any tented mouth appearance or dental misalignment against the data reported by Tim-Aroon *et al.* (2017). She had some behavioral abnormalities, including hyperactivity, aggression, tantrums, and inappropriate laughter with the absence of epilepsy. Vermeulen *et al.* (2017) reported that *GATAD2B*-related syndrome results in low levels of adaptive functioning, particularly in social functioning, and weaknesses in communication skills. Individuals with this condition have difficulty with expressive language, but their nonverbal communication skills are relatively well conserved. This goes consistent with our case study.

Acylcarnitine profiles are widely used to identify inborn errors of metabolism in humans (Millington *et al.* 1990). Carnitine was found to accumulate to a lower extent in the brain as compared to peripheral tissues (Bresolin *et al.* 1982). The enzymes needed for the synthesis of carnitine, and the acyltransferases necessary for the synthesis of acylcarnitines are present in brain tissues (Rebouche and Engel 1980; Bird *et al.* 1985). Further evidence suggests that carnitine level alteration is found in several neurological

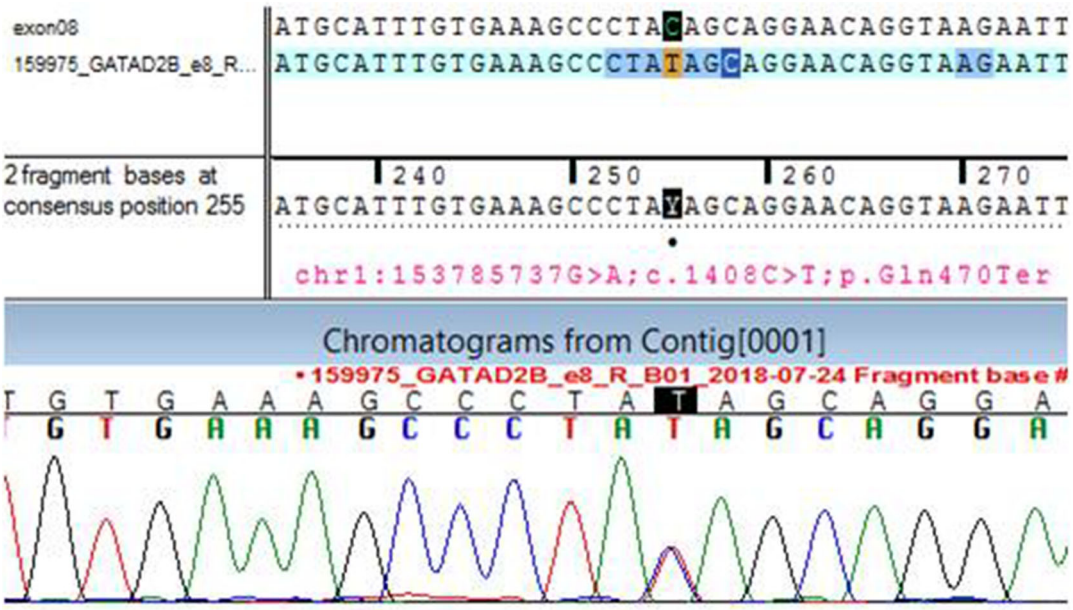

Figure 2. Sequence chromatogram of the case showing pathogenic variant in exon 8 of *GATAD2B* gene identifying MR18 syndrome and confirmed by Sanger sequencing.

Table 1. Tissue-specific DNase footprint analysis showing the impact of *GATAD2B* mutation on GATA binding motif.

Seq. name	Sequence	SFR	ΔSFR
Ref.	GAAAGCCCTACAGCAGGAA	1.046	−0.023
Var.	GAAAGCCCTATAGCAGGAA	1.070	
Ref.	GAAAGCCCTACAGCAGGAA	1.061	0.017
Var.	GAAAGCCCTATAGCAGGAA	1.044	
Ref.	GAAAGCCCTACAGCAGGAA	1.124	0.050
Var.	GAAAGCCCTATAGCAGGAA	1.074	
Ref.	GAAAGCCCTACAGCAGGAA	1.112	0.022
Var.	GAAAGCCCTATAGCAGGAA	1.090	
Ref.	GAAAGCCCTACAGCAGGAA	1.226	−0.028
Var.	GAAAGCCCTATAGCAGGAA	1.255	
Ref.	GAAAGCCCTACAGCAGGAA	1.203	−0.044
Var.	GAAAGCCCTATAGCAGGAA	1.246	
Ref	GAAAGCCCTACAGCAGGAA	1.603	0.331
Var.	GAAAGCCCTATAGCAGGAA	1.271	
Total absolute damage			0.515

Seq., sequence; ref., reference; var., variant. The difference in the sequence between reference and variant are underlined.

disorders like autism spectrum disorder (Frye *et al.* 2013), schizophrenia (Cuturic *et al.* 2016), and Alzheimer's disease (Ciavardelli *et al.* 2016). These studies suggest that neurological defects often share similar pathologies in mitochondrial dysfunction regarding energetic metabolic abnormalities. Tetradecanoylcarnitine is involved in the β-oxidation of long-chain fatty acids (Bene *et al.* 2005). Depletion of the C14 long-chain acylcarnitine may be caused by the activation or inhibition of different transporters in the plasma membrane. There are no reports about the depletion of tetradecanoylcarnitine in blood samples. The normal range of C14 acylcarnitine is from 0.03 to 1.22 μM and elevated levels are reported to be associated with very long-chain acyl-CoA dehydrogenase deficiency. The features of this syndrome include hypotonia, language deficit, and motor retardation (Brown *et al.* 2014). We can say that the lower level of C14 acylcarnitine might be one of the reasons for the distinct symptoms including skeletal hypotonia, speech delay, and delayed developmental milestones shown by the child. However, further research is necessary to shed light on this fact.

GATAD2B encodes transcriptional repressor p66 beta-component of the MeCP1 complex which functions in chromatin modification. It is a 593-amino acid protein that has a serine-rich region near the N-terminus, followed by two conserved regions, CR1 and CR2, essential for gene silencing. CR2 region (aa340–aa480) contains a GATA-type zinc finger motif responsible for histone tail-binding (Feng *et al.* 2002). Both CR1 and CR2 regions are required for speckled nuclear localization. Methyl-CpG-binding domain proteins (MBD) mediate functional responses of methylated DNA. MBD2 and MBD3 are components of the MeCP1 protein complex. The transcriptional repressor p66 protein reacts with MBD2 and MBD3 and may play a role in synapse development (Brackertz *et al.* 2002). The pathogenic variant at p470 causes *GATAD2B* loss-of-function,

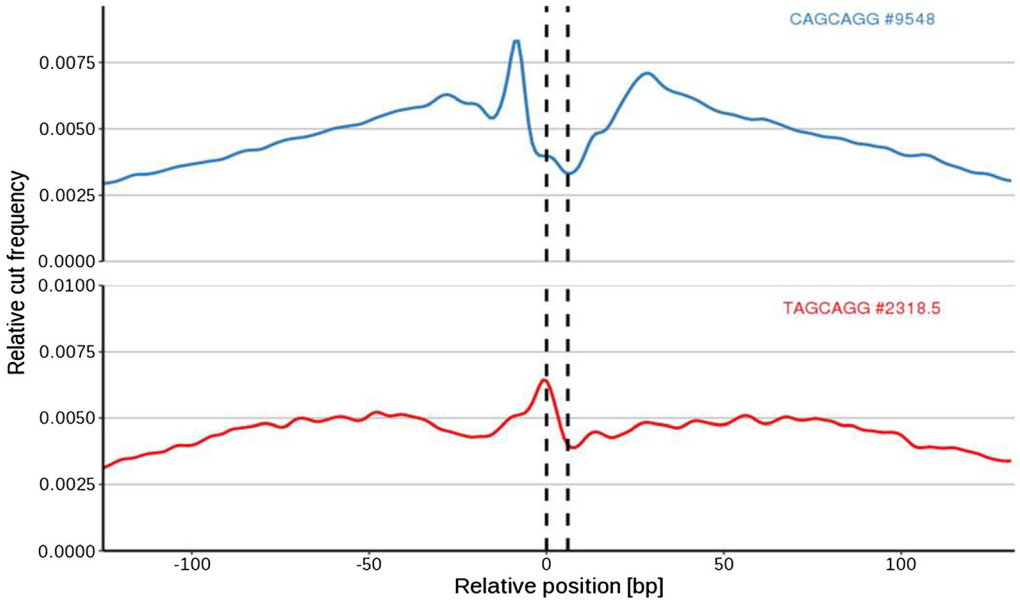

Figure 3. The impact of *GATAD2B* pathogenic variant on GATA zinc-finger domain. The blue and red graphs represent wild and variant k-mer sequences corresponding to GATA-motif. The wild sequence showed a strong signal corresponding to binding affinity of GATA while that signal is disrupted in the mutant sequence.

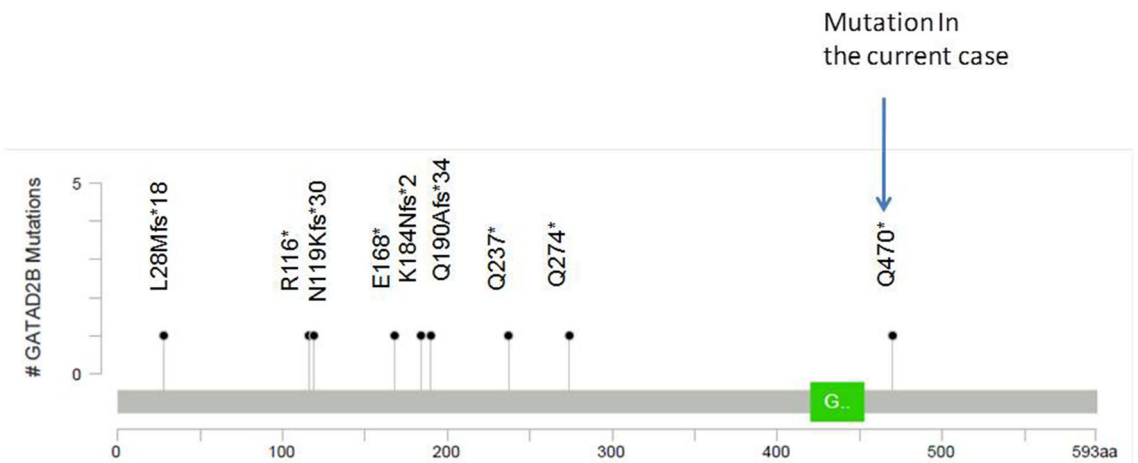

Figure 4. Spectrum of *GATAD2B* pathogenic variants identified worldwide and their respective location on the linear protein structure.

hence causing neuronal knockdown of the gene, which indicates its contribution to the ID phenotype. Loss of GATA-binding motif in *GATAD2B* might disrupt the interaction with NuRD complex binding partners, which might influence synaptic transmission.

In conclusion, we describe a girl owing to *GATAD2B* heterozygous pathogenic variant from India with a novel phenotype. Our research supports the importance of *GATAD2B* in neurodevelopment and indicates that the haploinsufficiency of *GATAD2B* causes severe ID, language disorder, and dysmorphic features. Our findings support the reported *GATAD2B* pathogenic variants and facilitate genetic diagnosis and counselling. Mechanistic rationale for this association was demonstrated through tissue-specific DNase I footprint analysis that showed disruption of GATA zinc finger motif due to this pathogenic variant.

Acknowledgements

We are grateful to the patient and her parents for participating in the study. We thank the Kamayani School for Mentally Challenged, Pune who provided and cared for the study individual. We acknowledge MedGenome Labs, Bengaluru for the genetic testing and Sandor Lifesciences, Hyderabad for the metabolic testing. Special thanks to Prof. M. V. Kulkarni, and the Department of Chemistry, Savitribai Phule Pune University for providing infrastructure for carrying out the research.

References

Bene J., Komlósi K., Gasztonyi B., Juhász M., Tulassay Z. and Melegh B. 2005 Plasma carnitine ester profile in adult celiac disease patients maintained on long-term gluten free diet. *World J. Gastroenterol.* **11**, 6671.

Bird M. I., Munday L. A., Saggerson E. D. and Clark J. B. 1985 Carnitine acyltransferase activities in rat brain mitochondria. Bimodal distribution, kinetic constants, regulation by malonyl-CoA and developmental pattern. *Biochem. J.* **226**, 323–330.

Brackertz M., Boeke J., Zhang R. and Renkawitz R. 2002 Two highly related p66 proteins comprise a new family of potent transcriptional repressors interacting with MBD2 and MBD3. *J. Biol. Chem.* **277**, 40958–40966.

Bresolin N., Freddo L., Vergani L. and Angelini C. 1982 Carnitine, carnitine acyltransferases, and rat brain function. *Exp. Neurol.* **78**, 285–292.

Brown A., Crowe L., Andresen B. S., Anderson V. and Boneh A. 2014 Neurodevelopmental profiles of children with very long chain acyl-CoA dehydrogenase deficiency diagnosed by newborn screening. *Mol. Genet. Metab.* **113**, 278–282.

Ciavardelli D., Piras F., Consalvo A., Rossi C., Zucchelli M., Di Ilio C. *et al*. 2016 Medium-chain plasma acylcarnitines, ketone levels, cognition, and gray matter volumes in healthy elderly, mildly cognitively impaired, or Alzheimer's disease subjects. *Neurobiol. Aging* **43**, 1–2.

Cuturic M., Abramson R. K., Breen R. J., Edwards A. C. and Levy E. E. 2016 Comparison of serum carnitine levels and clinical correlates between outpatients and acutely hospitalised individuals with bipolar disorder and schizophrenia: A cross-sectional study. *World J. Biol. Psychia.* **17**, 475–479.

De Ligt J., Willemsen M. H., Van Bon B. W., Kleefstra T., Yntema H. G., Kroes T. *et al*. 2012 Diagnostic exome sequencing in persons with severe intellectual disability. *New Eng. J. Med.* **367**, 1921–1929.

Feng Q., Cao R., Xia L. and Erdjument-Bromage H. 2002 Identification and functional characterization of the p66/p68 components of the MeCP1 complex. *Mol. Cell Biol.* **22**, 536–546.

Frye R. E., Melnyk S. and MacFabe D. F. 2013 Unique acyl-carnitine profiles are potential biomarkers for acquired mitochondrial disease in autism spectrum disorder. *Transl. Psychiat.* **3**, 220.

Hamdan F. F., Srour M., Capo-Chichi J. M., Daoud H., Nassif C. *et al*. 2014 De novo mutations in moderate or severe intellectual disability. *PLoS Genet.* **10**, 1004772.

Kaufman L., Ayub M. and Vincent J. B. 2010 The genetic basis of non-syndromic intellectual disability: a review. *J. Neurodev. Disord.* **2**, 182–209.

Kaur P., Mishra S., Rajesh S. M., Girisha K. M. and Shukla A. 2019 *GATAD2B*-related intellectual disability due to parental mosaicism and review of literature. *Clin. Dysmorphol.* **28**, 190–194.

Luo X., Zou Y., Tan B., Zhang Y., Guo J., Zeng L. *et al*. 2017 Novel *GATAD2B* loss-of-function mutations cause intellectual disability in two unrelated cases. *J. Hum. Genet.* **62**, 513–516.

Millington D. S., Kodo N., Norwood D. L. and Roe C. R. 1990 Tandem mass spectrometry: a new method for acylcarnitine profiling with potential for neonatal screening for inborn errors of metabolism. *J. Inherit. Metab. Dis.* **13**, 321–324.

Nikam V., Ranade S., Shaik Mohammad N. and Kulkarni M. 2019 A pilot study on machine learning approach to delineate metabolic signatures in intellectual disability. *Int. J. Dev. Disabil.* **12**, 1–7.

Rabin R., Millan F., Cabrera-Luque J. and Pappas J. 2018 Intellectual disability due to monoallelic variant in *GATAD2B* and mosaicism in unaffected parent. *Am. J. Med. Gen. A.* **176**, 2907–2910.

Rebouche C. J. and Engel A. G. 1980 Tissue distribution of carnitine biosynthetic enzymes in man. *Biochim. Biophys. Acta BBA Gen. Sub.* **630**, 22–29.

Schwessinger R., Suciu M. C., McGowan S. J., Telenius J., Taylor S., Higgs D. R. and Hughes J. R. 2017 Sasquatch: predicting the impact of regulatory SNPs on transcription factor binding from cell- and tissue-specific DNase footprints. *Genome Res.* **27**, 1730–1742.

Srivastava A. K. and Schwartz C. E. 2014 Intellectual disability and autism spectrum disorders: causal genes and molecular mechanisms. *Neurosci. Biobehav. Rev.* **46**, 161–174.

Tim-Aroon T., Jinawath N., Thammachote W., Sinpitak P., Limrungsikul A., Khongkhatithum C. *et al*. 2017 1q21. 3 deletion involving *GATAD2B*: An emerging recurrent microdeletion syndrome. *Am. J. Med. Gen. A.* **173**, 766–770.

Ueda K., Yanagi K., Kaname T. and Okamoto N. 2019 A novel mutation in the *GATAD2B* gene associated with severe intellectual disability. *Brain Dev.* **41**, 276–279.

Van Bokhoven H. 2011 Genetic and epigenetic networks in intellectual disabilities. *Annu. Rev. Genet.* **45**, 81–104.

Vanderver A., Simons C., Helman G., Crawford J., Wolf N. I., Bernard G. *et al*. 2016 Whole exome sequencing in patients with white matter abnormalities. *Ann. Neurol.* **79**, 1031–1037.

Vermeulen K., de Boer A., Janzing J. G., Koolen D. A., Ockeloen C. W. *et al*. 2017 Adaptive and maladaptive functioning in Kleefstra syndrome compared to other rare genetic disorders with intellectual disabilities. *Am. J. Med. Gen. A.* **173**, 1821–1830.

Willemsen M. H., Nijhof B., Fenckova M., Nillesen W. M., Bongers E. M. *et al*. 2013 *GATAD2B* loss-of-function mutations cause a recognisable syndrome with intellectual disability and are associated with learning deficits and synaptic undergrowth in Drosophila. *J. Med. Genet.* **50**, 507–514.

Zytkovicz T. H., Fitzgerald E. F., Marsden D., Larson C. A., Shih V. E. *et al*. 2001 Tandem mass spectrometric analysis for amino, organic, and fatty acid disorders in newborn dried blood spots: a two-year summary from the New England Newborn Screening Program. *Clin. Chem.* **47**, 1945–1955.

Corresponding editor: Ashwin Dalal