European Journal of Medical Genetics xxx (xxxx) xxxx

Contents lists available at ScienceDirect

European Journal of Medical Genetics

journal homepage: www.elsevier.com/locate/ejmg

DNA methylation fingerprint of monozygotic twins and their singleton sibling with intellectual disability carrying a novel *KDM5C* mutation

João V.S. Guerra[a,b], José Oliveira-Santos[c], Danyllo F. Oliveira[c,d], Gabriela F. Leal[e], João Ricardo M. Oliveira[d], Silvia S. Costa[c], Ana C.V. Krepischi[c], Angela M. Vianna-Morgante[c], Mariana Maschietto[a,f,*]

[a] Brazilian Biosciences National Laboratory (LNBio), Brazilian Center for Research in Energy and Materials (CNPEM), Campinas, SP, Brazil
[b] Postgraduate Program in Biosciences and Technology of Bioactive Products, Institute of Biology, University of Campinas, Campinas, SP, Brazil
[c] Department of Genetics and Evolutionary Biology, Institute of Biosciences, University of São Paulo, SP, Brazil
[d] Department of Neuropsychiatry and Keizo Asami Laboratory, Federal University of Pernambuco, Recife, PE, Brazil
[e] University of Pernambuco and Professor Fernando Figueira Integral Medicine Institute, Recife, PE, Brazil
[f] Boldrini Children's Hospital, Campinas, SP, Brazil

ARTICLE INFO

Keywords:
KDM5C
DNA methylation
Intellectual disability
XLID

ABSTRACT

Mutations in *KDM5C* (*lysine (K)-specific demethylase 5C*) were causally associated with up to 3% of X-linked intellectual disability (ID) in males. By exome and Sanger sequencing, a novel frameshift *KDM5C* variant, predicted to eliminate the JmjC catalytic domain from the protein, was identified in two monozygotic twins and their older brother, which was inherited from their clinically normal mother, who had completely skewed X-inactivation. DNA methylation (DNAm) data were evaluated using the Illumina 450 K Methylation Beadchip arrays. Comparison of methylation levels between the three patients and male controls identified 399 differentially methylated CpG sites, which were enriched among those CpG sites modulated during brain development. Most of them were hypomethylated (72%), and located mainly in shores, whereas the hypermethylated CpGs were more represented in open sea regions. The DNAm changes did not differ between the monozygotic twins nor between them and their older sibling, all presenting a global hypomethylation, similar to other studies that associated DNA methylation changes to different *KDM5C* mutations. The 38 differentially methylated regions (DMRs) were enriched for H3K4me3 marks identified in developing brains. The remarkable similarity between the methylation changes in the monozygotic twins and their older brother is indicative that these epigenetic changes were mostly driven by the *KDM5C* mutation.

1. Introduction

At least 21 *KDM5C* (*Lysine Demethylase 5C*) pathogenic mutations were identified in individuals with XLID, representing around 3% of the XLID cases (Abidi et al., 2008; Jensen et al., 2005; Rujirabanjerd et al., 2010; Santos-Rebouças et al., 2011). There is substantial clinical heterogeneity, however, short stature, microcephaly, hyperreflexia and aggressive behavior are usually reported in males (Gonçalves et al., 2014; Jensen et al., 2005), whereas female carriers rarely exhibit mild ID or learning difficulties (Rujirabanjerd et al., 2010). KDM5C is an epigenetic regulator that removes di- and trimethyl groups from lysine 4 of Histone 3 (H3K4me2/3) (Brookes et al., 2015; Tahiliani et al., 2007). In neurons, KDM5C is recruited to CpG islands at gene promoters targeted by H3K4me3 (Iwase et al., 2016). All *KDM5C* mutations were associated with reduced enzymatic activity, suggesting that the pathophysiological mechanism is a consequence of the protein loss

Abbreviations: DMRs, Differentially methylated regions; DNAm, DNA methylation; GEO, Gene expression Omnibus; H3K4, lysine 4 in histone H3; HM450K, Human Methylation 450 BeadChip microarrays; ID, Intellectual disability; MDS, Multidimensional scaling; MSET, Modular Single-set Enrichment test; OFC, Occipitofrontal circumference; P1, Monozygous twin 1; P2, Monozygous twin 2; P3, Older brother; shelves, regions defined as the 2 kb outside of a shores; shores, regions with lower CpG density that lie within the 2 kb up- and down-stream of a CpG island; TSS1500, region from transcript start site to 1500 nt upstream of transcript start site; XLID, X-linked intellectual disability

* Corresponding author. Boldrini Children's Hospital, Rua Dr. Gabriel Porto, 1270 - Cidade Universitária, 13083-210, Campinas, SP, Brazil.
E-mail addresses: jvsguerra@gmail.com (J.V.S. Guerra), josesantos.biologiamolecular@hotmail.com (J. Oliveira-Santos), oliveira.danyllo@gmail.com (D.F. Oliveira), gabferraz@gmail.com (G.F. Leal), joao.ricardo@ufpe.br (J.R.M. Oliveira), silvinha.costa@gmail.com (S.S. Costa), ana.krepischi@ib.usp.br (A.C.V. Krepischi), avmorgan@ib.usp.br (A.M. Vianna-Morgante), marianamasc@gmail.com (M. Maschietto).

https://doi.org/10.1016/j.ejmg.2019.103737
Received 29 October 2018; Received in revised form 25 June 2019; Accepted 11 August 2019

of function (Johnson et al., 2007). Patients carrying *KDM5C* mutations present disruption of the methylation dynamics related to cellular growth, differentiation, and developmental processes, whereas changes in global histone methylation were not detected (Johnson et al., 2007; Smyth, 2004). Recently, DNAm signatures were associated with nine out of 14 syndromes caused by mutations in genes involved in the epigenetic machinery, including *KDM5C*. The *KDM5C* DNAm signature showed enough specificity and accuracy to be proposed as a diagnostic tool (Schenkel et al., 2018).

Here, we report a novel frameshift *KDM5C* pathogenic variant, and its impact on the epigenome.

2. Materials and methods

2.1. Subjects

The research protocol was approved by the institutional ethical committee (Institute of Biosciences, University of São Paulo, São Paulo, Brazil) and informed consent was undersigned by the parents. This study was carried out in accordance the Declaration of Helsinki.

Two monozygotic twins (P1 and P2) and their older brother (P3) were the only children born to a healthy non-consanguineous couple. After uneventful pregnancies, all three had pre-term vaginal delivery (about 37 weeks). The twins P1 and P2 weighed 1.530 g and 1720 g, respectively, with length of 44 cm. Birth weight of P3 was 2080 g, no other measurements being recorded. Conditions at birth were good, and the newborns were discharged on the third day after birth.

The twins were examined at 12 6/12 years of age; severe intellectual disability was apparent, and they could speak only a few isolated words. P1 was 134.5 cm tall (2nd centile), weighing 28 kg (10th centile) and his occipitofrontal circumference (OFC) was 55.5 cm (75th centile). The height of P2 was 133.5 cm (2nd centile), weigh 27 kg (10th centile) and OFC 54.5 cm (50–75th centiles). They had no dysmorphic signs. They presented a calm behavior and a constant smiling facial expression. Brain MRI at the age of 13 years revealed no abnormalities in both twins. There was no history of seizures. Their motor development was slightly delayed, and they walked without support at the ages of 1 6/12 years (P1) and 1 7/12 years (P2).

Evaluated at the age of 10 8/12 years, patient P3, as his brothers, presented severe ID, and his speech had not developed beyond a few isolated words. His height was 118.5 cm (below the 2nd centile), weight 22.5 kg (2–10th centiles), and OFC 52.0 cm (25–50th centiles). He had no dysmorphic signs. He was calm and cheerful. At the age of 3 years, he started presenting tonic seizures. He walked without support at the age of 1 4/12 years.

Both parents were clinically normal. The stature of the father, however, was slightly below the 2.5th centile (159 cm); the mother had a normal height: 156 cm (25–50th centiles).

The control group for methylation analysis was formed by 12 clinically normal unrelated males, with ages ranging from 1 to 18 years (10.5 ± 5.8 years). We also evaluated the two clinically normal maternal aunts of the affected boys, who did not carry the *KDM5C* pathogenic variant found to be causative of ID in the family, as a control for the mother.

The control group and the patients were from self-declared white Brazilian families, but ancestry was not determined. All experimental procedures were conducted using DNA samples extracted from peripheral blood collected at the time of the evaluation described here.

2.2. Identification of the pathogenic variant

Whole-exome library of P1 was prepared using the Nextera Rapid Capture Expanded Exome (Illumina, California USA) and was paired-end sequenced (2x77bp) on HiSeq 2500 (Illumina, California, USA). Reads were aligned to the reference genome (GRCh37), using the BWA algorithm, and VCF files were generated using GATK tools (Broad Institute, Cambridge, USA). ANNOVAR (Wang et al., 2010) was used for annotation, filtering, and prioritization. Variants were filtered per frequency (< 1%) against NHLBI ESP6500SI–V2 exome variant frequencies, ClinVar, dbSNP138, 1000 Genome Project and ExAC Browser databases. Sanger sequencing was used to validate the *KDM5C* pathogenic variant (KDM5C–7F CACCTGGGGAGAAAAGTCAG and KDM5C-7R AAGGGACAAGAAGCAGAGCA), and for segregation analysis in the family.

2.3. X-inactivation analysis

X-chromosome inactivation pattern was determined in the patients' mother and maternal aunts, based on the methylation status of the androgen receptor gene (*AR*) (Allen et al., 1992). PCR products were separated by capillary electrophoresis on an ABI Prism 3700 Genetic Analyzer (*Applied Biosystems*), and analysis was performed using the Gene Mapper Software v4.1 (*Applied Biosystems*).

2.4. Hybridization to 450 K methylation arrays

High-quality bisulfite-converted DNA samples (EZ DNA Methylation Kit, Zymo Research Corp) were hybridized to the Infinium HumanMethylation450 BeadChip microarray (HM450K, Illumina). Raw data were extracted by the iScan SQ Scanner (Illumina), using GenomeStudio software (v.2011.1), with the methylation module v.1.9.0 (Illumina) as IDAT files. Probes were annotated according to the Illumina annotation file, using the Human GRCh37/hg19.

Analyses were conducted with the minfi package (Aryee et al., 2014). Data were normalized using FunNorm package (Aref-Eshghi et al., 2018) with further removal of probes located on X and Y chromosomes, and SNPs as well as cross-reactive probes, unless otherwise stated. Following the batch effect correction (Johnson et al., 2007), we quantified methylation differences resultant from cellular composition of the blood by using a purified dataset containing six blood cell types. Based on the estimated percentages of each cell type, a projection of the methylation levels for each sample (Houseman et al., 2012) was deconvoluted. Methylation levels for CpG sites were given as beta-values (0: unmethylated, 1: methylated); M-values (loggit of beta-values) were used for statistical analyses because they presented homoscedastic behavior (Fig. S1), whilst beta-values were used to plot the graphs.

The HM450K dataset from this study is available at NCBI Gene Expression Omnibus (GEO), under accession number GSE108423.

2.5. Analysis of differentially methylated CpG sites and regions

Differential methylation analyses were performed by comparing the three patients carrying the *KDM5C* mutation (duplicates) versus twelve male controls. The methylation pattern in the patients' mother, a clinically normal carrier, was compared to the patterns of her two non-carrier sisters. An empirical Bayesian analysis (Smyth, 2004) was applied to normalized M-values (duplicates were averaged), using a linear regression model from limma package (Ritchie et al., 2015), to identify differentially methylated CpG sites We considered significant those CpG sites with adjusted p-value < 0.05 (Benjamini and Hochberg, 1995).

DMRcate (Peters et al., 2015) was used to identify differentially methylated regions (DMRs), considering at least three CpGs in 500 bases as parameters. We considered significant those regions with absolute mean beta change higher than 10% and adjusted p-value < 0.05 (Benjamini and Hochberg, 1995).

2.6. Comparison to other datasets

The differentially methylated CpG sites were compared to those reported by three studies with publicly available data (Grafodatskaya et al., 2013; Schenkel et al., 2018; Spiers et al., 2015), using all probes from HM450K as background, using the Modular Single-set Enrichment

Test (MSET) (Eisinger et al., 2013) to randomly assess the enrichment of our list of differentially methylated CpG sites within the published lists. These studies included two lists of CpG sites from patients presenting different mutations in *KDM5C* (Eisinger et al., 2013; Schenkel et al., 2018), and one list of CpG sites modulated during brain development (Spiers et al., 2015). Briefly, to verify this enrichment, first we generated 10,000 simulated sets of CpG sites, without replacement, randomly generated from the HM450K. The degree of overlap between our list and each CpG site list from the other studies was compared to the 10,000 simulated sets. Then, a *p*-value was calculated based on the number of simulated randomized sets that contained an equal or higher number of CpG sites belonging to our list. Comparisons with an *in silico* p-value < 0.05 were considered significant.

By using the using chromosomes coordinates, DMRs were compared to H3K4me3 marks generated by ChIPseq of male and female human brains downloaded from the ENCODE portal (Davis et al., 2018) (https://www.encodeproject.org/) with the following identifiers: ENCFF290TTE, ENCFF385DTB, ENCFF442WNS, ENCFF483PRQ, ENCFF412ZGA and ENCFF415WPZ.

3. Results

3.1. Whole-exome sequencing discloses a novel frameshift KDM5C pathogenic variant

Whole-exome sequencing revealed a novel *KDM5C* frameshift variant [NM_004187.3(*KDM5C*):c.807delC; p.(Thr270Glnfs × 2)] in the twin P1 (Fig. 1A), further confirmed by Sanger sequencing; this variant was also present in the twin P2 and his older brother (P3), as well as in their clinically normal mother, but not in their two maternal aunts (Fig. 1B). This mutation in *KDM5C* exon 7 results in a stop codon, and the predicted protein, if produced, is devoid of the JmjC catalytic domain, thus pointing to the absence of a functional protein. This variant can be found in ClinVar under the Submission ID SUB5849796.

X-chromosome inactivation analyses revealed complete skewing (100:0) in the carrier mother, but not in her non-carrier sisters (36:65 and 36:64) (Fig. 1C).

3.2. Male patients carrying the frameshift KDM5C pathogenic variant exhibit DNA hypomethylation

Methylation values from 456,513 CpG sites were derived from the affected brothers (n = 3), male controls (n = 12), the carrier mother (n = 1) and non-carrier maternal aunts (n = 2). A multidimensional scaling (MDS) analysis of the 1% most variable CpGs assessed the inter-sample variability within samples not revealing a systematic methylation change among groups (Fig. 2A).

Comparison between the three patients carrying the *KDM5C* pathogenic variant (duplicates) with the 12 male controls. identified 399 differentially methylated CpG sites located in or next to 244 genes (Table S1). Upon clustering of the 399 CpGs (Pearson correlation with complete linkage), two groups were identified: Group 1 (the three affected brothers) and Group 2, which further separated into a cluster containing all but two male controls, and a cluster containing two male controls and the non-carrier aunts, showing that the methylation levels of these CpGs could differentiate patients from controls (Fig. 2B). Methylation levels of the 399 CpG sites did not differ between the three patients, including between the twins.

The patients presented a hypomethylated pattern compared to male controls, with 288 out of 399 CpG sites (72%) presenting hypomethylation (p-value < 0.0001, two-tailed Fisher's exact test). Slight hypomethylation was detected in the carrier mother compared to her non-carrier sisters.

Because our study explored only siblings, including a pair of monozygotic twins, we validated the detected differentially methylated CpG sites, using independent sets of unrelated individuals. We compared our data to the top 53 and to the 1769 differentially methylated CpG sites previously identified in patients carrying *KDM5C* mutations (Grafodatskaya et al., 2013; Schenkel et al., 2018), and found an enrichment of our dataset relative to these studies, with eight and 176 shared CpG sites (p-value < 0.0001), respectively (Table S1).

Hypomethylated CpGs were mainly located in shores (43%) of CpG islands, whereas hypermethylated CpGs were more represented in open sea regions (48%; p-value < 0.05, Chi-squared test with Yates' correction; Table S2). With respect to genes, the hypermethylated CpGs were mainly located in gene body (39%), whereas the hypomethylated CpGs were mainly represented in TSS1500 (33%, p-value < 0.05, Chi-squared test with Yates' correction; Table S3). Fig. 2C summarizes these results.

3.3. Differentially methylated regions were mostly located next to protein-coding genes

We found 38 DMRs (adjusted p-value < 0.05) randomly distributed across chromosomes, most of them located next to protein-coding genes (26 DMRs; 68%). Methylation levels of all DMRs were similar for the three siblings. Twenty-four out of 38 DMRs were within CpG islands and three others, next to transcriptional start sites. By comparing the 38 DMRs to the peaks of contiguous H3K4me3 enrichment identified in brains from a 17-week female and two 122-days male embryos (from ENCODE), we identified 29 and five DMRs (p-value < 0.0001), respectively (Table 1). Nineteen DMRs were more methylated in patients compared to controls (11 located next to *BRDT*, *SOX30*, *TSKS*, *GABRG1*, *TNNI3*, *GSTM5*, *PIEZO2*, *ANKH*, *CTB*, *HOOK2*, *ALO35460.1*, *EBF4* and *RP11*, and seven DMRs not located close to known genes); 19 DMRs were hypomethylated in patients relative to controls (located next to *NCL*, *KLHL21*, *POMC, MYO1F*, *AC013402.5, AC005082.12*, *SOX18*, *ZNF512B*, *TCEA2*, *DLEU2*, *OR1R1P*, RP11, *POU5F1*, *CTBP2* and *TLDC1*).

Three of the detected DMRs have been previously reported in patients carrying *KDM5C* mutations (Schenkel et al., 2018): next to *NCL* and next to *POMC,* and at chr6:164092588-164093099, not close to known genes. Three other DMRs had been identified in patients with other neurodevelopmental disorders (Aref-Eshghi et al., 2018): located next to *MYO1F* (in individuals with Floating-Harbor or Kabuki syndromes), next to *GABRG1* (individuals with Floating-Harbor or Sotos syndromes), and at chr1:247694041-247694531, not closely located to known genes (individuals with Claes-Jensen or Sotos syndromes).

In order: H3K4me3 – female, 17 weeks: ENCFF290TTE, ENCFF385DTB, ENCFF442WNS and ENCFF483PRQ; H3K4me3 – female, 17 weeks: ENCFF412ZGA and ENCFF415WPZ.

3.4. Differentially methylated CpG sites associated with the novel KDM5C pathogenic variant are involved in brain development

DNA extracted from peripheral blood cells was used as a surrogate to assess the impact of the *KDM5C* variant on methylation. Thus, we submitted all differentially methylated CpG sites to BECon (Edgar et al., 2017), which provides the correlation of methylation levels between matched brain and blood samples from 16 individuals, enabling biological interpretation of blood-based human DNAm data in the context of brain tissues. Among the 399 differentially methylated CpGs, 353 could be evaluated: 220 (62%) presented a positive correlation of methylation levels between blood and brain (42% with > 50% correlation; data not shown), suggesting that peripheral blood cells could be used as a proxy to evaluate methylation in neurons.

Additionally, we compared the detected differentially methylated sites to all CpGs with methylation levels modulated during brain development (Spiers et al., 2015). Our dataset presented an enrichment of 68 CpGs (p-value < 0.0001, Table S1).

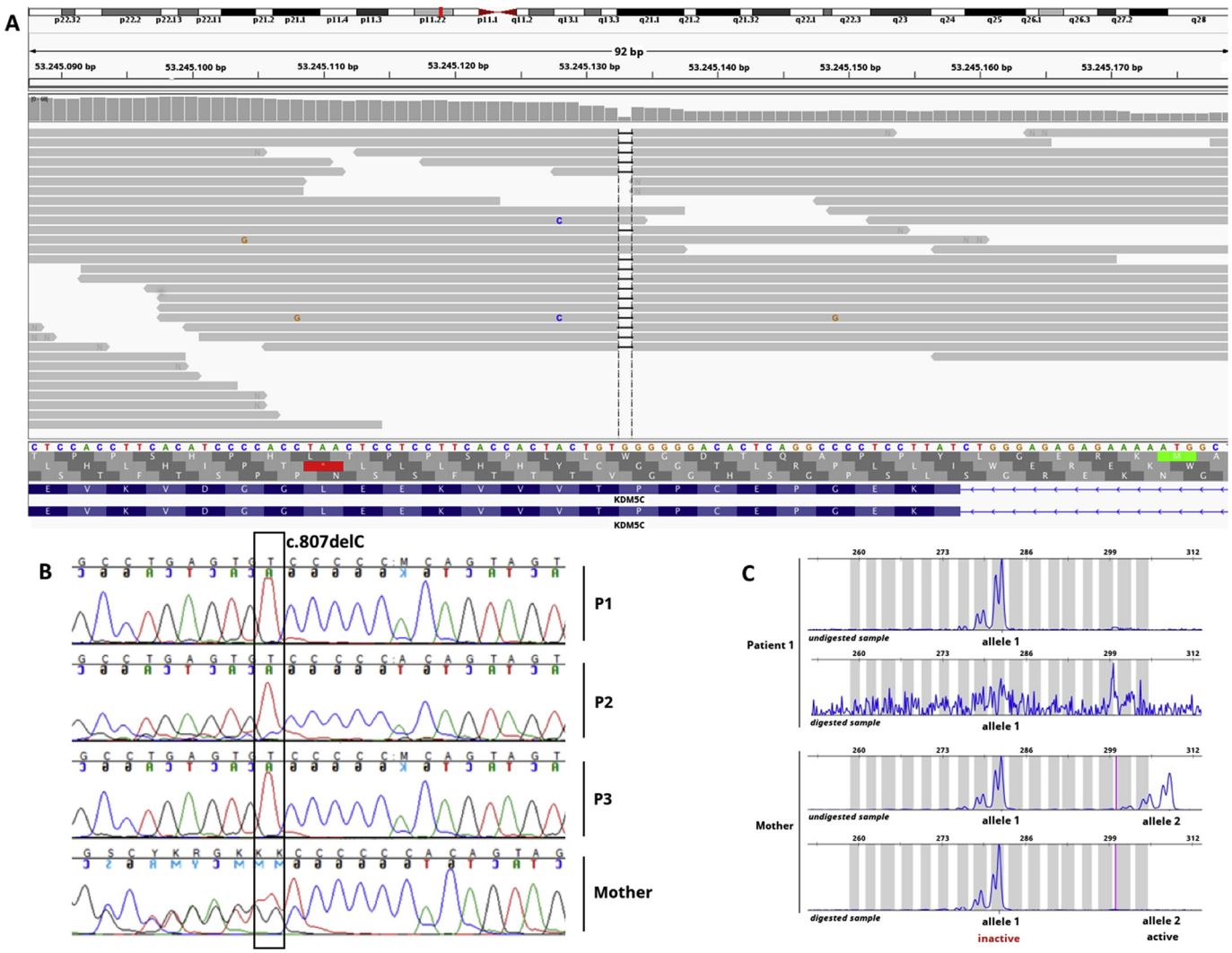

Fig. 1. A novel frameshift *KDM5C* pathogenic variant detected in three brothers presenting intellectual disability and short stature. A. *KDM5C* pathogenic variant [NM_004187.3(*KDM5C*):c.807delC; p.(Thr270Glnfs × 2)] identified by exome sequencing in one of the patients. BAM file image shows the hemizygous 1 bp deletion (*black bars*). B. Segregation analysis by Sanger sequencing validated the frameshift 1 bp deletion in exon 7 of the *KDM5C* gene, in the three patients and their clinically normal mother. C. X-chromosome inactivation analysis revealed complete skewing (100:0) in the carrier mother, pointing to the inactivation of the X-chromosome bearing the pathogenic variant.

4. Discussion

A novel *KDM5C* frameshift variant was identified in three brothers presenting intellectual disability and short stature, and in their clinically normal mother, who had completely skewed X-inactivation, as also observed in other women heterozygous for *KDM5C* mutations (Õunap et al., 2012). This mutation resulted in a premature stop codon, and a predicted non-functional protein. The effect of the *KDM5C* mutation was observed in the monozygotic twins and their sibling, all presenting a very similar pattern in the DNAm levels despite of the 50% of differences in the genetic background, suggesting that this mutation produced a specific impact in the epigenome.

Reduced activity of genes encoding histone demethylases is predicted to disrupt DNAm dynamics and cause aberrations in the chromatin state. ChIP-seq experiments showed that Kdm5c peaks often coincide with CpG islands, with 94% of the Kdm5c bound promoters containing a CpG island in neurons. Together with *Kdm5c* absence, increased levels of H3K4me3 were observed specifically at the promoter of *Kdm5c* target genes that presented low expression, but H3K4me1 levels were higher in *Kdm5c* target genes independent of their level of expression. H3K4me3 and H3K4me1 data suggest that low-expressed genes are more sensitive to the loss of *Kdm5c,* pointing to a role in the fine-tuning regulation of expression in neurons, besides some functional redundancy that could overcome *Kdm5c* loss (Iwase et al., 2016). Consistent with the hypothesis that *KDM5C* have a functional role for correct brain differentiation and function, we found an enrichment of the H3K4me3 marks in brains from human male and female embryos (Davis et al., 2018) among the 38 DMRs. This finding suggest that *KDM5C* loss might have a function impact for human neurons differentiation and/or function through DNA methylation perturbation. While these alterations may be detrimental to neurodevelopment, as evidenced by the associated ID phenotypes, all cells carrying the mutations may have some alterations that could contribute to the overall phenotype (Aref-Eshghi et al., 2018; Brookes et al., 2015). The positive correlation between methylation levels from peripheral blood and brain supports the use of blood as a proxy in the study of neurodevelopmental diseases, although with caution (Aref-Eshghi et al., 2018; Grafodatskaya et al., 2013; Maschietto et al., 2012; Schenkel et al., 2018; Walton et al., 2016). DNAm epi-signatures in the peripheral blood of individuals presenting ID as a consequence of mutation in genes related to the regulation of the epigenome were found to be highly sensitive and accurate to identify at least nine syndromes,

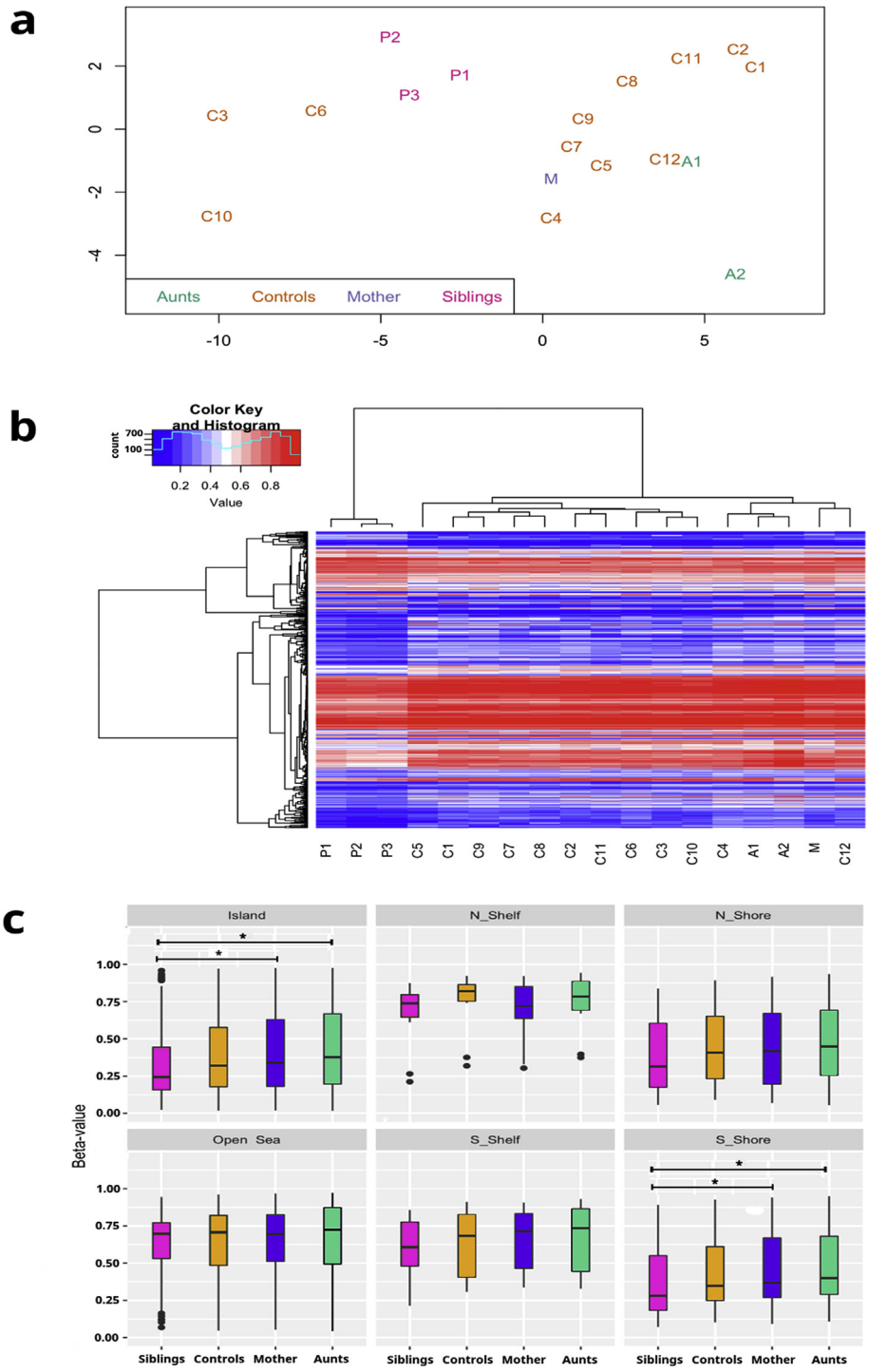

Fig. 2. Methylation analyses of individuals carrying the *KDM5C* pathogenic variant compared to controls. A. Multidimensional scaling of the top 1% most variable positions showed low intra-group variability. B. Hierarchical clustering (Pearson correlation and complete linkage method) of the 18 samples (columns), based on methylation levels of the 399 differentially methylated CpG sites. Heatmap colors refer to methylation levels: unmethylated (blue), partially methylated (white) and methylated (red). C. Boxplot of beta-values from the 399 differentially methylated CpG sites across the groups (affected siblings, male controls, and female relatives) per probe location in relation to CpG island. Y-axis represents the mean methylation values. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

including that caused by *KDM5C* mutations (Aref-Eshghi et al., 2018). Accordingly, most of the differentially methylated CpG sites identified in the present study (62%) are positively correlated between blood and brain.

The enrichment of the differentially methylated CpGs among the 7% of those modulated during brain development (Spiers et al., 2015) suggests that the disrupted methylation patterns during neocortex formation contribute to the ID phenotype of our patients. Our patients carrying the novel *KDM5C* pathogenic variant present a hypomethylated pattern compared to controls, with a significant proportion of CpGs being located at shores of CpG islands, corroborating previous findings (Schenkel et al., 2018). These data suggest that the disturbance of the methylation process, particularly in sites at or next to CpG islands is a result of the *KDM5C* mutation.

Comparing our data to DNAm changes reported for other patients carrying *KDM5C* variants (Aref-Eshghi et al., 2018; Grafodatskaya et al., 2013; Schenkel et al., 2018), we observed a significant overlap even though different methylation platforms were used, again suggesting that *KDM5C* mutations have an impact in the epigenome. From the 399 differentially methylated CpG sites detected by us, only 23 probes were represented in the previous study with the Infinium HumanMethylation27 Beadchip (Illumina) platform (Grafodatskaya et al., 2013), eight common CpG sites being modulated during brain development. These CpG sites were located at gene shores (*FBXL5*, *C2orf3*,

Table 1
Differentially methylated regions between patients and male controls. The overlap with CpG islands was verified in the UCSC Genome Browser (hg19). Distance to TSS (up to 1500 bp) was calculated considering the closest gene.

DMR	Genomic position (GRCh37)	DMR length (base pair)	Number of CpG sites	Genes related to DMR	Mean beta-value difference	Observation	TSS1500	CpG island overlap	H3K4me3 – female, 17 weeks	H3K4me3 – male, 122 days
Hypermethylated in individuals carrying KDM5C mutation										
DMR1	chr1:110254662-110255096	435	10	*GSTM5*	0.1068			–	Peak_16969, Peak_17214, Peak_16678, Peak_16969	
DMR2	chr1:147781827-147782558	732	7	-	0.1002			yes		
DMR3	chr1:19110734-19111089	356	4	-	0.2345			yes	Peak_23171, Peak_22839, Peak_26677, Peak_23171	
DMR4	chr1:247537159-247537369	211	3	-	0.1135			–	Peak_22368, Peak_24888, Peak_20810, Peak_22368	
DMR5	chr1:3269252-3269478	227	3	*PRDM16*	0.1021		(96 pb to TSS)	–	Peak_20827, Peak_25190, Peak_18952, Peak_20827	
DMR6	chr1:92414221-92414520	300	4	*BRDT*	0.1003		(1428 to TSS)	–		
DMR7	chr12:131118426–131118654	229	3	*RP11*	0.2488			–		
DMR8	chr17:184833–185263	431	4	*RPH3AL*	0.1165	Also identified in the DMP analysis		yes		
DMR9	chr18:11147146–11147785	640	3	*PIEZO2*	0.1073			–	Peak_13372, Peak_13595, Peak_24365, Peak_13372	Peak_17150, Peak_25079
DMR10	chr19:12876947–12877188	242	3	*HOOK2*	0.1747			yes	Peak_17348, Peak_17753, Peak_16853, Peak_17348	
DMR11	chr19:50249464–50249776	313	3	*TSKS*	0.1006			yes	Peak_30033 (only for ENCFF385DTB)	
DMR12	chr19:55667446–55668232	787	5	*TNNI3*	0.1048			yes	Peak_16745, Peak_15877, Peak_18581, Peak_16745	
DMR13	chr20:2674697–2675418	722	3	*AL035460.1; EBF4*	0.1910	Also identified in the DMP analysis		yes	Peak_14467, Peak_15310, Peak_14217, Peak_14467	
DMR14	chr4:46126066-46126448	383	6	*GABRG1*	0.1040	Similar to Aref-Eshghi et al. (2018)		–	Peak_9585, Peak_10267, Peak_8183, Peak_9585	Peak_26133, Peak_28660
DMR15	chr5:14870594-14870749	156	3	*ANKH; CTB*	0.1180			–	Peak_8476, Peak_7678, Peak_9419, Peak_8476	Peak_18262, Peak_18262
DMR16	chr5:157079404-157079520	117	4	*SOX30*	0.1005			yes	Peak_15038, NA, Peak_15640, Peak_15038	
DMR17	chr6:27185676-27186076	401	3	-	0.1030			–		
DMR18	chr7:158263204-158263277	74	3	-	0.1431			–		
DMR19	chr7:158789503-158790115	613	4	-	0.1264			yes		
Hypomethylated in individuals carrying KDM5C mutation										
DMR20	chr1:247694041-247694531	491	5	-	−0.1478	Similar to Aref-Eshghi et al. (2018)		yes		
DMR21	chr1:6664103-6664268	166	3	*KLHL21*	−0.1453	Also identified in the DMP analysis	(144 pb to TSS)	–	Peak_10034, Peak_10114, Peak_10386, Peak_10034	
DMR22	chr10:126850823–126851326	504	4	*CTBP2*	−0.1074			yes	Peak_9601, Peak_8160, Peak_9167, Peak_9601	
DMR23	chr10:63657059–63657363	305	3	-	−0.1079			–	Peak_13038, Peak_13175, Peak_12744, Peak_13038	
DMR24	chr13:50700606–50701256	651	4	*DLEU2*	−0.1251			yes	Peak_1584, Peak_2959, Peak_1845, Peak_1584	Peak_7454, Peak_7454
DMR25	chr15:31115839–31115871	33	3	-	−0.1568			–	Peak_29041, NA, Peak_24282, Peak_29041	
DMR26	chr16:53407594–53407808	215	4	*RP11*	−0.1041			yes	Peak_18107, Peak_17813, Peak_18686, Peak_18107	
DMR27	chr16:84538885–84539110	226	3	*TLDC1*	−0.1051			yes	Peak_5077, Peak_2502, Peak_7629, Peak_5077	Peak_45835, Peak_45835

(*continued on next page*)

Table 1 (*continued*)

DMR	Genomic position (GRCh37)	DMR length (base pair)	Number of CpG sites	Genes related to DMR	Mean beta-value difference	Observation	TSS1500	CpG island overlap	H3K4me3 – female, 17 weeks	H3K4me3 – male, 122 days
DMR28	chr17:3289363–3290164	802	6	*OR1R1P; RP11*	−0.1178			yes	Peak_20373, Peak_19533, Peak_22608, Peak_20373	
DMR29	chr19:8591364–8591776	413	3	*MYO1F*	−0.1430	Similar to Aref-Eshghi et al. (2018)		yes	Peak_7940, Peak_7797, Peak_8487, Peak_7940	
DMR30	chr2:105275613-105276153	541	5	*AC013402.5*	−0.1390			yes	Peak_14305, Peak_14859, Peak_14116, Peak_14305	
DMR31	chr2:232347922-232348794	873	5	*NCL*	−0.1456	Similar to Schenkel et al. (2018)		yes	Peak_20907, Peak_21942 and Peak_44482, Peak_21548, Peak_20907	
DMR32	chr2:25383404-25384809	1406	7	*POMC*	−0.1475	Similar to Schenkel et al. (2018)		yes	Peak_16459, Peak_16126, Peak_17604, Peak_16459	
DMR33	chr20:62679255–62679713	459	3	*SOX18; ZNF512B; TCEA2*	−0.1259			yes	NA, Peak_40855, Peak_31688, NA	
DMR34	chr5:115697214-115697696	483	4	-	−0.1179			yes	Peak_13256, Peak_13442, Peak_13218, Peak_13256	
DMR35	chr5:415575-415885	311	3	-	−0.1097			–		
DMR36	chr6:164092588-164093099	512	4	-	−0.1253	Similar to Schenkel et al. (2018)		yes		
DMR37	chr6:31148332-31148748	417	15	*POU5F1*	−0.1165			yes		
DMR38	chr7:23245900-23246146	247	3	*ENSG00000226816*	−0.1361			yes		

EIF2B3, *C17orf44*, *CACYBP* and *KLHL21*) or at a CpG island (*HENMT1*). We found a more significant overlap (49% of the differentially methylated CpGs) with recently published data that explored the epigenetic signature of *KDM5C* mutations, using the same HM450K platform used in the present study (Schenkel et al., 2018).

Three DMRs have been reported in other patients carrying *KDM5C* mutations (Schenkel et al., 2018), two of them located next to *POMC* and *NCL*. The role of both genes in the clinical phenotype of the individuals carrying *KDM5C* mutations needs further exploration. Finally, three DMRs, two of them located next to *GABRG1* and *MYOF1*, were also detected in the study of ID syndrome-specific epigenetic signatures in peripheral blood (Aref-Eshghi et al., 2018). *GABRG1* has an increasing expression in development, from the 14th week to birth. when reaches a plateau; *MYOF1* appears to have lower expression during neocortex development compared to high expression after birth (Colantuoni et al., 2011). These DMRs were also found in other syndromes suggesting that they might be important for common phenotypes.

Moreover, DNA methylation epi-signatures can help with the diagnosis of patients with intellectual disability, particularly those carrying variants of unknown clinical significance. This was demonstrated in a 9-year-old patient carrying a *KDM5C* variant of unknown clinical significance but that presented a methylation profile similar to other patients with Sotos syndrome (Aref-Eshghi et al., 2019), suggesting that this specific variant may be related to the DNA methylation alterations found in the patients and thus, might not be fully responsible for or not be related to the patient's phenotype. A proper interpretation of DNA methylation together with other types of data (mutation, phenotype, clinical symptoms, etc) will improve diagnosis of patients with intellectual disability.

5. Conclusion

Our results showed that the novel *KDM5C* (pathogenic variant resulted in methylation changes in a subset of CpG sites of the affected male siblings, but not in their clinically normal mother. This mutation led to multilocus hypomethylation (72% of differentially methylated CpG sites), including *FBXL5* and *CACYBP*, which are known to be regulated by KDM5C. Together with the enrichment data of other *KDM5C*-mutated datasets, these data validate our analyses and allow us to suggest that the genes identified here are associated with methylation changes that contribute to the clinical features of the patients. Moreover, we report a remarkable similarity in the methylation patterns in the monozygotic twins and their sibling, that shares 50% of the genetic background, strongly suggesting that these epigenetic changes were mostly driven by the *KDM5C* pathogenic variant.

Declarations

Consent for publication

Participants provided assent in addition to undersigned consent by their legal guardian.

Competing interests

The authors declare that they have no competing interests.

Funding sources

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP [Project CEPID - Human Genome and Stem Cell Research Center - 2013/08028-1]. JVSG and MM were supported by FAPESP 2015/22758-8 and 2015/06281-7, respectively.

Author's contributions

Contributions to the conception and design of the study: AMVM, ACVK, MM. Identification of patients and collection of clinical data: DFO, GFL, JRMO. Exome sequencing analysis and segregation study: JOS, AMVM. X-inactivation analysis: SSC, AMVM. Acquisition, analysis, and interpretation of methylation data: JVSG, ACVK, MM. Authors that drafted the article: JVSG, MM. Revision of the manuscript: ACVK, AMVM, MM. All authors read and approved the final version of this manuscript.

Accession numbers

*KDM5C variant at C*linVar - Submission ID SUB5849796.

KDM5C methylation experiments are available at NCBI/GEO (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE108423).

Acknowledgments

The authors thank the participants and their families. We thank Ana Carolina Tahira for advising in the analysis of ENCODE data. We also thank to ENCODE Consortium and Zhiping Weng and Joseph Costello laboratories for generating the male and female brain datasets.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmg.2019.103737.

References

Abidi, F.E., Holloway, L., Moore, C.A., Weaver, D.D., Simensen, R.J., Stevenson, R.E., Rogers, R.C., Schwartz, C.E., 2008. Mutations in JARID1C are associated with X-linked mental retardation, short stature and hyperreflexia. J. Med. Genet. 45, 787–793. https://doi.org/10.1136/jmg.2008.058990.

Allen, R.C., Zoghbi, H.Y., Moseley, a B., Rosenblatt, H.M., Belmont, J.W., 1992. Methylation of HpaII and HhaI sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome inactivation. Am. J. Hum. Genet. 51, 1229–1239.

Aref-Eshghi, E., Bend, E.G., Colaiacovo, S., Caudle, M., Chakrabarti, R., Napier, M., Brick, L., Brady, L., Carere, D.A., Levy, M.A., Kerkhof, J., Stuart, A., Saleh, M., Beaudet, A.L., Li, C., Kozenko, M., Karp, N., Prasad, C., Siu, V.M., Tarnopolsky, M.A., Ainsworth, P.J., Lin, H., Rodenhiser, D.I., Krantz, I.D., Deardorff, M.A., Schwartz, C.E., Sadikovic, B., 2019. Diagnostic utility of genome-wide DNA methylation testing in genetically unsolved individuals with suspected hereditary conditions. Am. J. Hum. Genet. 104 (4), 685–700. https://doi.org/10.1016/j.ajhg.2019.03.008. https://doi.org/10.1016/j.ajhg.2019.03.008.

Aref-Eshghi, E., Rodenhiser, D.I., Schenkel, L.C., Lin, H., Skinner, C., Ainsworth, P., Paré, G., Hood, R.L., Bulman, D.E., Kernohan, K.D., Boycott, K.M., Campeau, P.M., Schwartz, C., Sadikovic, B., 2018. Genomic DNA methylation signatures enable concurrent diagnosis and clinical genetic variant classification in neurodevelopmental syndromes. Am. J. Hum. Genet. 102, 156–174. https://doi.org/10.1016/j.ajhg.2017.12.008.

Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., Irizarry, R.A., 2014. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 30, 1363–1369. https://doi.org/10.1093/bioinformatics/btu049.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B. 57 (1), 289–300. https://doi.org/10.2307/2346101.

Brookes, E., Laurent, B., Õunap, K., Carroll, R., Moeschler, J.B., Field, M., Schwartz, C.E., Gecz, J., Shi, Y., 2015. Mutations in the intellectual disability gene KDM5C reduce protein stability and demethylase activity. Hum. Mol. Genet. 24, 2861–2872. https://doi.org/10.1093/hmg/ddv046.

Colantuoni, C., Lipska, B.K., Ye, T., Hyde, T.M., Tao, R., Leek, J.T., Colantuoni, E.A., Elkahloun, A.G., Herman, M.M., Weinberger, D.R., Kleinman, J.E., 2011. Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature 478, 519–523. https://doi.org/10.1038/nature10524.

Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., Onate, K.C., Graham, K., Miyasato, S.R., Dreszer, T.R., Strattan, J.S., Jolanki, O., Tanaka, F.Y., Cherry, J.M., 2018. The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res. 46 (D1), D794–D801. https://doi.org/10.1093/nar/gkx1081.

Edgar, R.D., Jones, M.J., Meaney, M.J., Turecki, G., Kobor, M.S., 2017. BECon: a tool for interpreting DNA methylation findings from blood in the context of brain. Transl. Psychiatry 7. https://doi.org/10.1038/tp.2017.171.

Eisinger, B.E., Saul, M.C., Driessen, T.M., Gammie, S.C., 2013. Development of a versatile enrichment analysis tool reveals associations between the maternal brain and mental health disorders, including autism. BMC Neurosci. 14. https://doi.org/10.1186/1471-2202-14-147.

Gonçalves, T.F., Gonçalves, A.P., Fintelman Rodrigues, N., dos Santos, J.M., Pimentel, M.M.G., Santos-Rebouças, C.B., 2014. KDM5C mutational screening among males with intellectual disability suggestive of X-Linked inheritance and review of the literature. Eur. J. Med. Genet. 57, 138–144. https://doi.org/10.1016/j.ejmg.2014.02.011.

Grafodatskaya, D., Chung, B.H., Butcher, D.T., Turinsky, A.L., Goodman, S.J., Choufani, S., Chen, Y.A., Lou, Y., Zhao, C., Rajendram, R., Abidi, F.E., Skinner, C., Stavropoulos, J., Bondy, C.A., Hamilton, J., Wodak, S., Scherer, S.W., Schwartz, C.E., Weksberg, R., 2013. Multilocus loss of DNA methylation in individuals with mutations in the histone H3 Lysine 4 Demethylase KDM5C. BMC Med. Genomics 6. https://doi.org/10.1186/1755-8794-6-1.

Houseman, E.A., Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K., Kelsey, K.T., 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinf. 13, 1–16. https://doi.org/10.1186/1471-2105-13-86.

Iwase, S., Brookes, E., Agarwal, S., Badeaux, A.I., Ito, H., Vallianatos, C.N., Tomassy, G.S., Kasza, T., Lin, G., Thompson, A., Gu, L., Kwan, K.Y., Chen, C., Sartor, M.A., Egan, B., Xu, J., Shi, Y., 2016. A mouse model of X-linked intellectual disability associated with impaired removal of histone methylation. Cell Rep. 14, 1000–1009. https://doi.org/10.1016/j.celrep.2015.12.091.

Jensen, L.R., Amende, M., Gurok, U., Moser, B., Gimmel, V., Tzschach, A., Janecke, A.R., Tariverdian, G., Chelly, J., Fryns, J.-P., Van Esch, H., Kleefstra, T., Hamel, B., Moraine, C., Gecz, J., Turner, G., Reinhardt, R., Kalscheuer, V.M., Ropers, H.-H., Lenzner, S., 2005. Mutations in the JARID1C gene, which is involved in transcriptional regulation and chromatin remodeling, cause X-linked mental retardation. Am. J. Hum. Genet. 76, 227–236. https://doi.org/10.1086/427563.

Johnson, W.E., Li, C., Rabinovic, A., 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics 8, 118–127. https://doi.org/10.1093/biostatistics/kxj037.

Maschietto, M., Silva, A.R., Puga, R.D., Lima, L., Pereira, C.B., Nakano, E.Y., Mello, B., Gama, C.S., Belmonte-de-Abreu, P., Carraro, D.M., Palha, J.A., Brentani, H., 2012. Gene expression of peripheral blood lymphocytes may discriminate patients with schizophrenia from controls. Psychiatry Res. 200, 1018–1021. https://doi.org/10.1016/j.psychres.2012.04.030.

Õunap, K., Puusepp-Benazzouz, H., Peters, M., Vaher, U., Rein, R., Proos, A., Field, M., Reimand, T., 2012. A novel c.2T > C mutation of the KDM5C/JARID1C gene in one large family with X-linked intellectual disability. Eur. J. Med. Genet. 55, 178–184. https://doi.org/10.1016/j.ejmg.2012.01.004.

Peters, T.J., Buckley, M.J., Statham, A.L., Pidsley, R., Samaras, K., V Lord, R., Clark, S.J., Molloy, P.L., 2015. De novo identification of differentially methylated regions in the human genome. Epigenet. Chromatin 8, 6. https://doi.org/10.1186/1756-8935-8-6.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., Smyth, G.K., 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43, e47. https://doi.org/10.1093/nar/gkv007.

Rujirabanjerd, S., Nelson, J., Tarpey, P.S., Hackett, A., Edkins, S., Raymond, F.L., Schwartz, C.E., Turner, G., Iwase, S., Shi, Y., Futreal, P.A., Stratton, M.R., Gecz, J., 2010. Identification and characterization of two novel JARID1C mutations: suggestion of an emerging genotype-phenotype correlation. Eur. J. Hum. Genet. 18, 330–335. https://doi.org/10.1038/ejhg.2009.175.

Santos-Rebouças, C.B., Fintelman-Rodrigues, N., Jensen, L.R., Kuss, A.W., Ribeiro, M.G., Campos, M., Santos, J.M., Pimentel, M.M.G., 2011. A novel nonsense mutation in KDM5C/JARID1C gene causing intellectual disability, short stature and speech delay. Neurosci. Lett. 498, 67–71. https://doi.org/10.1016/j.neulet.2011.04.065.

Schenkel, L.C., Aref-Eshghi, E., Skinner, C., Ainsworth, P., Lin, H., Paré, G., Rodenhiser, D.I., Schwartz, C., Sadikovic, B., 2018. Peripheral blood epi-signature of Claes-Jensen syndrome enables sensitive and specific identification of patients and healthy carriers with pathogenic mutations in KDM5C. Clin. Epigenet. 10, 21. https://doi.org/10.1186/s13148-018-0453-8.

Smyth, G.K., 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. 3, 1–25. https://doi.org/10.2202/1544-6115.1027.

Spiers, H., Hannon, E., Schalkwyk, L.C., Smith, R., Wong, C.C.Y., O'Donovan, M.C., Bray, N.J., Mill, J., 2015. Methylomic trajectories across human fetal brain development. Genome Res. 25, 338–352. https://doi.org/10.1101/gr.180273.114.

Tahiliani, M., Mei, P., Fang, R., Leonor, T., Rutenberg, M., Shimizu, F., Li, J., Rao, A., Shi, Y., 2007. The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. Nature 447, 601–605. https://doi.org/10.1038/nature05823.

Walton, E., Hass, J., Liu, J., Roffman, J.L., Bernardoni, F., Roessner, V., Kirsch, M., Schackert, G., Calhoun, V., Ehrlich, S., 2016. Correspondence of DNA methylation between blood and brain tissue and its application to schizophrenia research. Schizophr. Bull. 42, 406–414. https://doi.org/10.1093/schbul/sbv074.

Wang, K., Li, M., Hakonarson, H., 2010. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38. https://doi.org/10.1093/nar/gkq603.