



Università di Foggia

FACULTY OF MEDICINE AND SURGERY

PhD Course in Experimental and Regenerative Medicine
XXX Cycle

*Mutational analysis of Kabuki Syndrome patients and
functional dissection of KMT2D mutations*

Tutor
Prof. V.M Fazio

PhD Student
Dario Cocciadiferro

Supervisor
Dr. Giuseppe Merla

A.A. 2016/2017

INDEX

ABSTRACT	1
INTRODUCTION	
KABUKI SYNDROME: CLINICAL CHARACTERISTICS.....	3
PREVALENCE AND PROGNOSIS.....	4
GENETICS OF KABUKI SYNDROME: <i>KMT2D</i>	5
GENETICS OF KABUKI SYNDROME: <i>KDM6A</i>	8
DIFFERENTIAL DIAGNOSIS.....	10
GENE-SPECIFIC DNA METHYLATION SIGNATURES IN DISEASES.....	12
HISTONE METHYLATION	13
KABUKI SYNDROME AS DISORDER OF THE EPIGENETIC MACHINERY.....	15
MOSAICISM AND KABUKI SYNDROME.....	17
READTHROUGH STRATEGIES FOR SUPPRESSION OF NONSENSE VARIANTS.....	19
NONSENSE MEDIATED mRNA DECAY IN MAMMALS.....	20
AIM OF THE THESIS	22
MATERIAL AND METHODS	
PATIENTS AND SAMPLES PREPARATION.....	23
<i>KMT2D</i> AND <i>KDM6A</i> SEQUENCE MUTATION, MLPA AND PYROSEQUENCING.....	23
IN <i>SILICO</i> ANALYSIS OF <i>KMT2D</i> AND <i>KDM6A</i> VARIANTS.....	24
IN <i>SILICO</i> ANALYSIS OF <i>KMT2D</i> PROTEIN DOMAINS.....	24
CELL LINES, PLASMIDS AND TRANSFECTION ASSAYS.....	25
ESTABLISHMENT OF LYMPHOCYTE CELL LINES AND NMD ASSAY.....	25
QUANTIFICATION OF READTHROUGH LEVELS IN HEK293.....	26
SITE-DIRECTED MUTAGENESIS.....	26
TRANSFORMATION.....	27
SCREENING FOR RECOMBINANTS.....	27
PLASMID DNA EXTRACTION.....	28
TOTAL RNA ISOLATION AND REVERSE TRASCRIPTION (RT-PCR).....	29
REAL-TIME POLYMERASE CHAIN REACTION (qPCR).....	30
<i>IN VITRO</i> HISTONE METHYLTRANSFERASE (HMT) ASSAY AND EPIGENETIC REPORTER ALLELE.....	31
CO-IMMUNOPRECIPITATION ASSAY AND WESTERN BLOTTING ANALYSIS.....	31
RNAi AND NT2D1 CELL DIFFERENTIATION.....	32

RESULTS	
MUTATION SCREENING OF <i>KMT2D</i> AND <i>KDM6A</i>	33
KMT2D NONSENSE AND FRAMESHIFT VARIANTS.....	37
IDENTIFICATION OF <i>KMT2D</i> NONSENSE VARIANTS RESPONSIVE TO GENTAMICIN TREATMENT.....	38
<i>KMT2D</i> SPLICE SITE VARIANTS.....	40
<i>KMT2D</i> MOSAIC VARIANTS.....	43
MISSENSE VARIANTS PATHOGENIC ASSESSMENT BY BIOINFORMATIC APPROACHES.....	44
MISSENSE VARIANTS IMPAIR KMT2D METHYLTRANSFERASE ACTIVITY.....	47
KMT2D-COMPLEX PROTEIN INTERACTION.....	50
KMT2D MEDIATES NEURONAL DIFFERENTIATION OF NT2D1 CELLS BY ACTIVATING DIFFERENTIATION SPECIFIC GENES.....	51
 DISCUSSION	 53
 CONCLUSIONS	 56
 REFERENCES	 57

ABSTRACT

The discovery of histone methyltransferase *KMT2D* and demethylase *KDM6A* genetic alterations in Kabuki Syndrome (KS) expanded and highlighted the role of histone modifiers in causing congenital anomalies and intellectual disability syndromes. KS is a rare autosomal dominant condition characterized by facial features, various organ malformations, postnatal growth deficiency, and intellectual disability. Since 2011 we performed a mutational screening of our KS cohort, that includes now 505 KS patients, by Sanger sequencing and MLPA of *KMT2D*, followed by *KDM6A* analysis in those patients resulted as *KMT2D*-negative. Of these 505 patients, we identified 196/505 (39%) patients with *KMT2D* variants and 208 different *KMT2D* variations; of them 37/208 (18%) never described before. The majority of KS patients carry nonsense and splice site variants, suggesting the loss of function, and therefore haploinsufficiency, as the likely mechanism for the KS phenotype. RT-PCR and direct sequencing on cDNA from Kabuki patients carrying *KMT2D* splice site variants demonstrated that these cause aberrant splicing of the corresponding transcript, resulting in a truncating and not functional translated protein.

Molecular assays also showed that *KMT2D* mRNAs bearing premature stop codon are degraded by the nonsense mediated mRNA decay, contributing to *KMT2D* protein haploinsufficiency. We hypothesized that KS patients may benefit from a readthrough therapy that mediates translational suppression of nonsense variants, restoring the physiologically levels of endogenous *KMT2D* protein. Fourteen *KMT2D* nonsense variants were tested for their response to readthrough treatment through an in vitro dual reporter luciferase vector system, identifying 11/14 variants that displayed high levels of readthrough in response to gentamicin treatment.

Among our cohort we identified three new cases with a mosaic variants in *KMT2D* gene, consisting in single nucleotide change resulting in two already reported nonsense variants, the c.13450C=>T (p.R4484X) and the c.15061C=>T (p.R5021X) and in a new frameshift variant, the c.3596_3597=/del (p.L1199HfsX7) *KMT2D*, respectively.

Moreover, relevant for diagnostic and counselling purposes, we implemented a number of bioinformatics tools to assess the pathogenicity of 69 *KMT2D* missense variants, found overall in our cohort of 505 KS patients, and for 14 of them we adopted a combination of biochemical and cellular approaches to investigate their role and characterize their functional impact in the pathogenesis of the disease. We found 9/14 missense variants showing altered H3K4 methylation activity. We additionally assessed the impact on complex formation with

WRAD protein complex, and we found that the reduced methyltransferase activity could be a consequence of lack of interaction.

Overall this study:

- Expands the number of *KMT2D* and *KDM6A* variants that cause KS, as a result of a 505 KS patients cohort.
- Enlarges the number of KS patients with mosaic mutations by reporting three additional cases with a single nucleotide change in *KMT2D*.
- Adds some insight to the functional mechanisms that cause the disease.
- Provides a preliminary proof-of-concept that naturally occurring nonsense mutations in *KMT2D* can be effectively suppressed by treatment with readthrough inducers, moving us closer to personalized medicine for KS.
- Affords a strategy to estimate the real deleterious effect of *KMT2D* missense variants in KS.

INTRODUCTION

Kabuki Syndrome: clinical characteristics

Kabuki Syndrome (KS, MIM #147920, MIM #300867) also known as Kabuki makeup syndrome, KMS or Niikawa–Kuroki Syndrome is a rare autosomal dominant condition identified in 1981 by Niikawa and Kuroki, whose reported for the first time a syndrome characterized by intellectual disability, short stature and a peculiar facial gestalt that reminded the authors the make-up of actors in the traditional Japanese Kabuki theatre [1, 2] (Figure 1). It is characterized by striking facial features such as elongated palpebral fissures with eversion of the lateral third of the lower eyelid, short columella with depressed nasal tip, skeletal anomalies, dermatoglyphic abnormalities, mild to moderate mental retardation, and postnatal growth deficiency [1, 2]. Additional findings include congenital heart defects, genitourinary anomalies, cleft lip and/or palate, susceptibility to infections, gastrointestinal abnormalities, ophthalmologic defects, ptosis and strabismus, dental anomalies including widely spaced teeth and hypodontia, ear pits, visceral abnormalities, and premature thelarche [3].

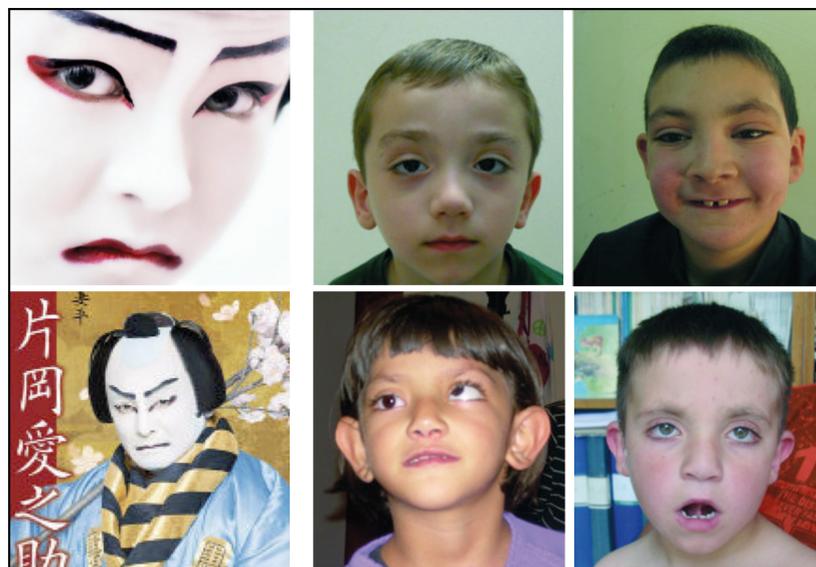


Figure 1. Kabuki patients' phenotype: On the left, Kabuki actors' make-up in the traditional Japanese Kabuki theatre. On the right, craniofacial features of KS patients [4].

Involvement of the connective tissue has been proposed because of the combination of facial laxity, joint hyperlaxity and joint dislocations [5, 6] as well as the association with autoimmune diseases, in particular idiopathic thrombocytopenic purpura, common variable immune deficiency (CVID), and/or hemolytic anemia in rare cases [7, 8]. A susceptibility to

infection leading to recurrent infections of the upper airway tract and the middle ear is also a known and a common feature of KS [9]. KS patients have particular delays in speech and language acquisition and autism or autistic-like behavior, with difficulties in both communication and peer interactions [10, 11]. Consensus clinical diagnostic criteria for Kabuki Syndrome have not been established yet, although individuals with this condition usually have five cardinal manifestations according to Niikawa et al. 1988 [12], including: typical facial features, skeletal anomalies, dermatoglyphic abnormalities (persistence of fetal fingertip pads), mild to moderate intellectual disability, and postnatal growth deficiency (Table1).

Table 1: Clinical features of KS patients

<i>Features</i>	<i>Prevalence %</i>
Facial anomalies	87 [#]
Skeletal anomalies	80 [#]
Dermatoglyphic abnormalities	92 [#]
Developmental delay/Learning difficulties	99 [#]
Postnatal growth deficiency	58 [#]
Elongated palpebral fissures	95 [*]
Sparse eyebrows	82 [*]
Palpebral ptosis	52 [*]
Broad nasal tip	69 [*]
Thin upper and full lower lip	71 [*]
Large dysmorphic ears	90 [*]
Short stature	70 [#]
Feeding difficulties	82 [#]
Cleft palate	37 [#]
Cardiac defects	46 [#]
Urogenital anomalies	42 [#]

Legend: KS, Kabuki Syndrome; * Micale et al. 2011 [4], # Bogershausen, et al. 2016 [18].

Prevalence and Prognosis

KS was initially thought to be specific to Japanese individuals, with an estimated prevalence in Japan of 1/32,0005 [12]. Even if the global incidence of KS is still not known, there are now reports of a variety of ethnic groups, including European, Brazilian, Vietnamese, Filipino, East Indian, Arabic, Chinese, Mexican, and African [3, 13] that presumably approximate the global prevalence to the Japanese one (1/32,0005). Because KS is not classically associated with severe medical complications, the adulthood prognosis is good, particularly if congenital anomalies and infections are properly managed.

Genetics of Kabuki Syndrome: *KMT2D*

In 2010, whole-exome sequencing successfully identified heterozygous loss of function variants in the *KMT2D* (NM_003482.3; MIM #602113, also known as *MLL2* and *MLL4*) gene as the major cause of KS [14]. *KMT2D* encodes a conserved member of the SET1 family of histone lysine methyltransferases (KMTs), which catalyzes the methylation of lysine 4 on histone H3 (H3K4), a modification associated with active transcription [15-17]. The enzymatic function of *KMT2D* depends on a cluster of conserved C-terminal domains, including plant homeodomains (PHD), two phenylalanine and tyrosine (FY)-rich motifs (FY-rich, N-terminal (FYRN) and FY-rich, C-terminal (FYRC) and a catalytic Su(var)3-9, Enhancer-of-zeste, Trithorax (SET) domain. Six different KMTs have been identified in higher eukaryotes, which fall into three subgroups on the basis of homologies in protein sequence and subunit composition: SET1A-SET1B, MLL1-MLL4 (KMT2A-KMT2B) and MLL3-MLL2 (KMT2C-KMT2D) [15-18]. KMT2C-KMT2D subgroup localizes and functions as major histone H3K4 mono, di and trimethyltransferases in a tissue specific manner at enhancers and promoters in mammalian cells [19-24] playing an important role in the epigenetic control of chromatin states [25] acting as transcriptional coactivator involved in the expression control of genes essential for embryogenesis and development, such as the *HOX* genes [26, 27]. The SET1 protein family, including *KMT2D*, is active in the context of a multisubunit complex, sharing four common highly conserved components, namely, WDR5, RbBP5, Ash2L and Dpy30 (a mini complex referred to as WRAD) that are related to yeast Set1 COMPASS complex [15], with the minimal four-component complex, including WDR5, RbBP5 and Ash2L along with the *KMT2D* SET-domain subunit, that can reconstitute most of the H3K4-specific histone methyltransferase activity (Figure 2) [28-31].

In addition, *KMT2D* also interacts with the Histone H3 lysine-27 demethylase KDM6A, who plays an important role for general chromatin remodelling and regulation of homeobox (*HOX*) genes, contributing to the correct reprogramming and tissue-specific differentiation during development [32-34]. During adipogenesis and myogenesis, *KMT2D* exhibits cell type and differentiation-stage-specific binding, co-localizing with lineage determining transcription factors on active enhancers [22]. The localization of *KMT2D* on active enhancers is also required for the enhancer-binding and enhancer activation of H3K27 acetyltransferase p300, one of the major enhancer epigenomic writers with *KMT2D*, both drivers of the enhancer epigenome and transcriptome processes during brown adipogenesis and ESC differentiation [35, 36]. The inactivation of the zebrafish *kmt2d* orthologue by morpholino is associated with significant craniofacial defects with severe hypoplasia of the viscerocranium, including loss of branchial arches, Meckel's cartilage and the ceratohyal

from the cartilaginous structures [37]. Consistent with the involvement of *KMT2D* in critical cellular functions, germline homozygous deletion in mice results in embryonic lethality [22, 38].

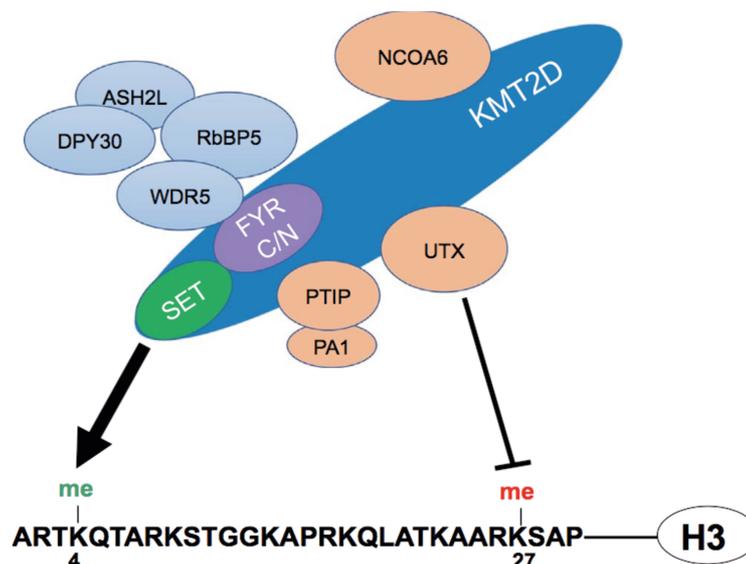


Figure 2. The KMT2D Protein Complex. KMT2D associates with WRAD (WDR5, RbBP5, ASH2L, DPY30), UTX, NCOA6, PA1, and PTIP in one complex. The C-terminal SET domain is the source of KMT2D's H3K4 methyltransferase activity. UTX is the complex's H3K27 demethylase. H3K4 methylation is associated with gene activation, whereas H3K27 methylation is associated with gene repression. Froimchuk et al. 2017 [39].

KMT2D functional domains

KMT2D (located at 12q13 chromosome region) encodes for a large protein of 5537 amino acid residues (54 exons) that consists of seven plant homeodomains (PHD), one high-mobility group domain (HMG), Leu-X-X-Leu-Leu (LXXLL) and Coiled-coil motifs, one each FYRC and FYRN domains, and a single SET domain (Figure 3).

1. PHD and HMG domains

The Plant HomeoDomain (PHD) finger, implicated in protein–protein interactions, is a C4HC3 zinc-finger-like motif of approximately 50-80 aminoacids in length, found in more than 100 human nuclear proteins involved in epigenetics and chromatin mediated transcriptional regulation. The HMG domain is a DNA-binding motif that is observed in non-histone components of chromatin and in specific regulators of transcription and cell differentiation [40].

2. Coiled-coils and LXXLL motifs

Coiled coil domains mediate protein homo-oligomerisation; they may interact with each other to form homotypic oligomers or with other coiled-coil domains to form heterotypic

oligomers. KMT2D contains five LXXLL motifs that participate in protein–protein interactions associated with different aspects of transcriptional regulation [41]. LXXLL motifs were originally observed in cofactor proteins that interact with hormone-activated nuclear receptors. Mo et al. (2006) [42] demonstrated that KMT2D physically interacts with ER α , through two LXXLL motifs after stimulation with the steroid estrogen hormone. As matter of fact, KMT2D suppression by its specific siRNA, decreased estrogen-induced expression of ER α target genes, such as cathepsin D and pS2 [42]. KMT2D may also interact with ER α via different KMT2D interacting proteins (NR-box containing), such as Menin, ASC2 and INI1 that contain multiple LXXLL domains. The tumour suppressor Menin interacts with ER α , recruits KMT2D complex into the promoter of estrogen-responsive genes and regulates their expression in a ligand-dependent manner [43], working as a critical link between activated ER α and the KMT2D-coactivator complex. A similar interaction involving Menin and KMT2D was observed at the Hox loci; in absence of the Menin protein, H3K4 trimethylation of the entire Hox cluster is abolished [44, 45].

3. FYRN and FYRC domains

These motifs are two poorly characterized phenylalanine/tyrosine-rich regions of approximately 50–100 aminoacids, respectively, that are found in a variety of chromatin-associated proteins [46]. The FYRN and FYRC motifs are closely juxtaposed and are involved in heterodimerisation between the N-ter and C-ter fragments of the protein.

4. SET domain

KMT2D is characterized by an evolutionarily conserved SET domain that is found in a number of chromatin-associated proteins with diverse transcriptional activities [47]. The SET domain, is a histone methyltransferase motif named for its presence in three Drosophila chromatin regulators: a modifier of position-effect variegation, Suppressor of variegation 3–9 (Su(var)3–9) [48], the Polycomb-group chromatin regulator Enhancer of zeste (E(z)) [49] and the Trithorax-group chromatin regulator trithorax (Trx) [50]. The role of SET-domain is to transfer a methyl group from the donor S-adenosyl-L-methionine (AdoMet) to the amino-group of a lysine residue on the histone or other proteins, leaving a methylated lysine residue and the cofactor by product S-adenosyl-L-homocysteine (AdoHcy). SET domain proteins can be classified into several families that differ because of the substrate specificity and the presence of associated domains [51]. KMT2D belongs to the SET1 family, which is found in conserved multisubunit complexes that regulate cellular H3K4 methylation levels [52]. The KMT2D SET domain is flanked on the C-terminus by a 22-aminoacid post-SET domain,

which provides several conserved residues that coordinate a zinc atom necessary for KMT2D enzymatic activity.

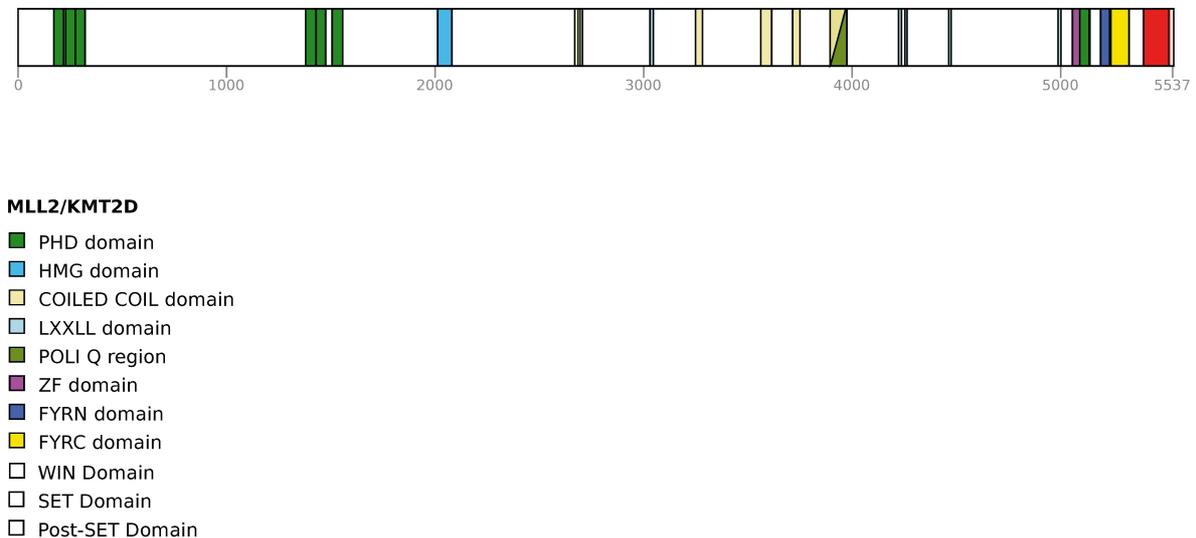


Figure 3. Schematic representation of KMT2D protein and functional domains

Genetics of Kabuki Syndrome: *KDM6A*

The genetic heterogeneity of KS has been confirmed by the discovery of mutations involving the Histone H3 lysine-27 demethylase *KDM6A* (MIM #300128, NM_021140.3), also known as *UTX*, in a subset of *KMT2D*-negative patients [53]. Since then a number of studies revealed that variants in *KMT2D* and *KMD6A* account for approximately 56%-75% and 5%-8% of KS cases, respectively [18]. *KDM6A* escapes X inactivation, but it has been suggested that, although in mice *Kmd6a* expression from inactive X chromosome is significantly lower than that from the active X chromosome, two copies of the gene are necessary for the normal expression in female mouse embryos and adult mice [54] suggesting that *KDM6A* haploinsufficiency might be a pathogenic mechanisms. *KDM6A* belongs to the large family of JumonjiC (JmjC) domain-containing proteins that also includes *UTY* (ubiquitously transcribed tetra tricopeptide repeat, Y chromosome) and *JMJD3* [55]. The *UTY* protein is the Y-linked homologue of *UTX* sharing 88% homology with *UTX* and might compensate for the increased expression of *KDM6A* in females. The third member of this family, *JMJD3*, is predicted to mediate protein-protein interactions and shares extensive homology sequence with *UTX* and *UTY* within and outside of the JmjC domain. *KDM6A* is composed by 29 exons encoding a 1401 residues protein that contains six TPR (tetratricopeptide repeat)

domains in the N-terminal half and the JmjC domain in the C-terminal half of the protein (Figure 4).

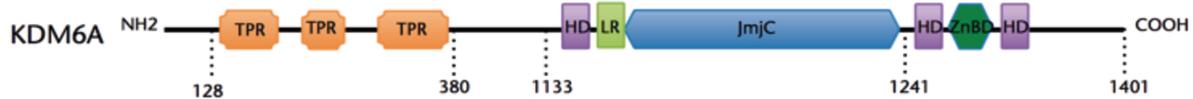


Figure 4. KDM6A protein domains. TPR, tetratricopeptide repeat; LR, linker region; HD, helical domain; JmjC, JumonjiC catalytic domain; ZnBD, Zinc binding domain. The helical and Zn-binding domains, together with the Jmj domain, comprise the KDM6A catalytic domain [40].

KDM6A specifically demethylates di- and trimethyl-lysine 27 (H3K27me_{2/3}) on histone H3 [56, 57]. Trimethylation of H3K27 (H3K27me₃) is considered an epigenetic mark important for maintaining embryonic stem cell pluripotency and plasticity during embryonic development, for Polycomb-mediated gene silencing and X chromosome inactivation [58]. Genomic analysis suggests that the presence of H3K27me₃ at transcriptional start sites is generally correlated with the repression of gene expression, even at target genes containing a high level of H3K4me₃. In pluripotent cells, *HOX* genes remain silent and are marked by a continuous pattern of H3K27me₃. During vertebrate embryogenesis, *Hox* genes exhibit temporal and spatial colinearity of expression, first with activated *Hox* genes in the most anterior body structure, and later with activated *Hox* genes in the more posterior body structures. Histone methylation patterns delimit the expression boundaries: active genes are marked by a continuous stretch of H3K4me_{2/3}, whereas H3K27me₃ marks silent genes. *In vivo*, histone demethylases create large complexes with other histone-modifying enzymes [25, 34, 59]. KDM6A is reported to form a complex with the KMT2D histone H3K4 methyltransferase complex, suggesting that KDM6A removes the histone H3K27 methyl mark when the KMT2D complex methylates histone H3K4, which are antagonistic methyl marks for gene expression [34].

Interestingly, KDM6A and KMT2D act together in the epigenetic control of transcriptionally active chromatin by counter acting Polycomb-group (PcG) proteins [60] that are involved in the tri- and dimethylation of H3K27. Similar to KMT2D, KDM6A plays a role in embryogenesis and development. Homozygous *dUTX* *Drosophila* mutants have increased H3K27me₃ levels and display rough eyes, dysmorphic wings and modified sex combs. This phenotype resembles the *trithorax* phenotype, supporting the notion that KDM6A counteracts PcG repression [61]. In addition, KDM6A, together with KMT2D, plays a major role in the regulation of muscle-specific genes during embryogenesis [60, 62]. Members of the lineage-

defining T-box transcription factor family recruit KDM6A to activate specific target gene profile in their natural developmental context like mesoderm, heart and vertebrae, emphasising the role of KDM6A in the developmental processes [63]. Interestingly, most of the pathogenic variants reported for T-box genes in human diseases are located in the KDM6A-interacting T-box domain.

Differential Diagnosis

Some disorders share overlapping features with KS. Evaluation of the patient via clinical assessment as well as molecular analysis is often performed to investigate the causes of clinical findings that are characteristic of KS but may also appear in other genetic syndromes. Disorders with the most overlapping features with Kabuki Syndrome (KS) include the following:

1) Turner syndrome (TS): is a neurogenetic disorder characterized by partial or complete monosomy-X associated with estrogen deficiency, short stature and increased risk for several diseases with cardiac conditions. Genetic analyses identified the Short stature HOmeoboX (*SHOX*) gene as the candidate gene for short stature and other skeletal abnormalities associated with TS but currently the gene or genes associated with cognitive impairments remain unknown [64]. Overlapping features with KS include heart defects, short stature, scoliosis, and similar facial features [65]. Constitutional *KDM6A* mutations cause Kabuki Syndrome with variable degrees in heterozygous females because of low expression levels from the normal wild type allele. Most of 45, X Turner females do not have a complete Kabuki Syndrome probably because of the dosage compensation of a single X chromosome in 45, X patients that might be greater than the one in individuals with two sex chromosomes or because of a differential X chromosome inactivation. As patients with Turner syndrome only share some overlapping features with Kabuki Syndrome, monoallelic *KDM6A* expression is apparently sufficient to prevent this phenotype [66, 67].

2) CHARGE syndrome: (Coloboma of the eye, Heart defects, Atresia of the choanae, Retardation of growth and/or development, Genital and/or urinary abnormalities, and Ear abnormalities) is a genetic disorder characterized by a specific and a recognizable pattern of anomalies. De novo variants in the gene encoding chromodomain helicase DNA binding protein 7 (CHD7) are the major cause of CHARGE syndrome [68].

Overlapping clinical features with KS include: postnatal growth retardation, cleft lip/palate, hearing loss, congenital heart defects, urogenital malformations, developmental delay, and intellectual disability. As well, ocular coloboma, which is a major diagnostic criterion for CHARGE syndrome, has occasionally been reported in individuals with Kabuki Syndrome [69]. Typical facial features and prominent fingertip pads in KS are distinct from those in CHARGE syndrome. Pathogenic variants in CHD7 are causative; inheritance is autosomal dominant [70].

3) 22q11.2 deletion syndrome (22q11.2DS): is the most common chromosomal microdeletion disorder, estimated to result mainly from de novo non-homologous meiotic recombination events occurring in approximately 1 in every 1,000 fetuses [71]. The syndrome is now known to have a heterogeneous presentation that includes multiple additional congenital anomalies and later-onset conditions, such as palatal, gastrointestinal and renal abnormalities, autoimmune disease, variable cognitive delays, behavioural phenotypes and psychiatric illness — all far extending the original description of DiGeorge syndrome [71]. The overlapping clinical features with KS comprise the cleft palate, congenital heart defects, and urinary tract anomalies. However, the different characteristic facial features seen in the two conditions should distinguish them [72].

4) IRF6-related disorders: is a group of orofacial clefting disorders including Van der Woude syndrome (VWS) and popliteal pterygium syndrome (PPS) caused by mutations in the interferon regulatory factor 6 (IRF6) gene [73]. The overlapping clinical features with KS are mainly the cleft lip/palate and lip pits although individuals with IRF6-related disorders do not have atypical growth and development, cardiac malformations, or the typical Kabuki Syndrome facies [70].

5) Branchio-oto-renal (BOR) syndrome: is a rare autosomal dominant disorder characterized by pits or ear tags in front of the outer ear, abnormal passages from the throat to the outside surface of the neck (branchial fistulas), branchial cysts, hearing loss and/or kidney abnormalities [74]. Ear pits, cupped ears, hearing loss, and renal anomalies are overlapping features with KS. However, individuals with BOR syndrome have otherwise normal craniofacies, normal growth, and normal development. Branchial cleft cysts may be present in BOR but have not been reported in KS [70].

6) Hardikar syndrome: is a disorder of multiple anomalies predominantly characterized by cleft lip/palate, liver and biliary tract disease, intestinal malrotation, obstructive uropathy, and

retinopathy [75]. Overlapping features with KS include prolonged hyperbilirubinemia with cleft lip and palate. However, individuals with KS do not typically develop pigmentary retinopathy or sclerosing cholangitis, as seen in Hardikar syndrome [70].

7) Wiedemann Steiner syndrome (WSS): is an autosomal dominant congenital anomaly syndrome characterized by hairy elbows, dysmorphic facial appearances (hypertelorism, thick eyebrows, downslanted and vertically narrow palpebral fissures), pre- and post-natal growth deficiency, and psychomotor delay. WSS is caused by heterozygous mutations in *KMT2A* (also known as MLL), a gene encoding a histone methyltransferase. Overlapping clinical features with KS are postnatal growth restriction, hypertrichosis, and intellectual disability [76, 77].

Gene-specific DNA methylation signatures in diseases

Epigenetics was originally defined as the genes-environment interaction process that leads to manifestations of various phenotypes during development [78]. In the recent years, this conception has evolved into the study of heritable changes in gene expression that occur without a change in DNA sequence. To date, the best understood epigenetic mechanisms are CpG DNA methylation and histone modifications. DNA methylation in particular has been the subject of intense interest because of its recently recognized role in disease and cancer as well as in the development and normal function of organisms. Many diseases are a consequence of alterations in gene expression patterns leading to deranged biological functions in cells [79]. Genes engaged in epigenetic regulation (epigenes), including those involved in chromatin remodeling and histone modifications, are increasingly being identified in the etiology of a variety of neurodevelopmental disorders, causing specific patterns of DNA methylation (DNAm) alterations that constitute unique signatures [80-82]. As a matter of fact, unique DNAm signatures are observed in individuals harboring variants in lysine-specific demethylase 5C (KDM5C [MIM: 314690]), which encodes an H3K4 demethylase and causes non-syndromic intellectual disability [MIM: 300534], DNA methyltransferase 1 (DNMT1 [MIM: 126375]), which cause autosomal dominant cerebellar ataxia with deafness and narcolepsy (ADCA- DN [MIM: 604121]), and nuclear receptor binding SET domain protein 1 (NSD1 [MIM: 606681]) which encodes for a histone methyltransferase and cause Sotos syndrome [MIM:117550] [80-82]. Moreover, for NSD1 has been shown that genes encoding proteins in growth and neurodevelopmental pathways are highly represented in the DNAm signature reflecting the pathophysiology of the disorder [80].

Unique DNAm signatures have also been identified in Kabuki Syndrome patients carrying *KMT2D* variants, determining a characteristic global DNA methylation profile that distinguishes patients with *KMT2D* loss of function, missense or splice site variants from normal controls [69, 83]. Interestingly, the DNA methylation changes found in Kabuki Syndrome, showed similar magnitude and some overlapping target genes with Wiedermann-Steiner and Charge syndrome, all sharing overlapping features with Kabuki Syndrome. For instance, the DNA methylation profile of Kabuki Syndrome patients with *KMT2D* missense variants, clustered with the same of *KMT2A* missense and splice site variants, the main gene associated to Wiedermann-Steiner syndrome [83], while genome-wide DNAm profiles in patients with CHARGE and Kabuki Syndrome with *CHD7*-LOF or *KMT2D*-LOF variants, share 14CpG sites, 11 corresponding to *HOXA5* and 3 to *SLITRK5*, which could account for some of the clinical overlap in CHARGE and Kabuki Syndrome [69].

Diagnosis of CHARGE, Wiedermann-Steiner, Kabuki Syndrome, and a variety of neurodevelopmental disorders in the clinical and molecular setting can be challenging. These data provide new insights into the genotype-epigenotype and phenotype relationship, indicating a cross-talk between histone and DNA methylation machineries exposed by inborn errors of the epigenetic apparatus. DNAm classification signature could be used as a functional molecular test to aid in the interpretation of the pathogenicity of the sequence variants, providing a valuable tool to facilitate in the diagnosis of the syndrome [69].

Histone methylation

The epigenetic control of developmental processes is a mechanism by which spatial and temporal expression of distinct genes and pathways are regulated. Alterations of this epigenetic mechanism, including histone modifications (acetylation, methylation and phosphorylation), have been mainly associated with the pathogenesis of cancer and with a variety of congenital malformation syndromes [84, 85]. In the extended form, chromatin appears as an array of nucleosomes, but in the nucleus the chromatin fibres that form chromosomes undergo several levels of folding, resulting in increasing degrees of condensation [86]. It is known that the histone tails have an important role in this folding process, and the histone tails methylation is important to regulate chromatin structure directly by affecting the higher-order folding of the chromatin fibre [87]. This has important implications for chromatin-templated processes such as transcription and DNA repair, altering the accessibility of DNA to the proteins that mediate these processes [88].

Histone lysine methylation marks, transcriptionally activate and inactivate chromatin domains in eukaryotic genomes, playing a key role in the regulation of transcription, gene silencing, and heterochromatin formation. This process is dynamically regulated by different site-specific histone methyltransferases (HMTs), including KMT2D (Figure 5). HMTs are key enzymes that introduce methyl groups into the lysine side chain of histone proteins and regulate both gene activation and silencing depending on the specific lysine residue that becomes methylated and the level of methylation (mono-, di-, or trimethylation). KMT2D mono- and trimethylation of histone H3 at K4 (H3K4) are highly associated with gene activation; H3K4 (me1) is enriched on enhancer regions and is required for enhancer activation and super-enhancer formation during cell differentiation of brown adipose tissue and skeletal muscle development [22, 89], whereas H3K4 (me3) is increased on active promoters/transcription start sites (TSS) and is important for actively transcribed gene regulation [90, 91]. The main sites of lysine methylation that have been associated with gene activity include K4, K36 and K79 of histone H3. Interestingly, the methylation of all three sites seems to be directly coupled to the transcription process. In the case of H3-K4 and H3-K36 methylation, the enzymes responsible for both modifications have been shown to physically associate with RNA polymerase II (RNAPII) during elongation, resulting in histone methylation in the coding regions [88, 92-95]. At least six different HMTs have been identified in higher eukaryotes, which fall into three subgroups on the basis of homologies in protein sequence and subunit composition: SET1A-SET1B, MLL1-MLL4 (KMT2A-KMT2B) and MLL3-MLL2 (KMT2C-KMT2D) [15-18].

On the contrary, H3-K9 and H3-K27 histone methylation has been linked to several silencing phenomena including homeotic-gene silencing, X inactivation and genomic imprinting [96]. The core of this silencing system is the POLYCOMB GROUP (PcG) of proteins [97], a complex that varies in different species, although the SET-domain-containing protein enhancer of zeste (E(Z); or its human homologue EZH2), extra sex combs (ESC; or its human homologue embryonic ectoderm development (EED) and suppressor of zeste-12 (SUZ12) proteins are present in all complexes that have been isolated so far [88]. Whereas all forms of H3-K27 methylation require EED79, SUZ12 seems to be dispensable for the monomethylation of H3-K27 in vivo, although it is indispensable for the enzymatic activity in vitro [98, 99]. Dynamic changes in the methylation pattern during gene expression indicated the existence of histone demethylase enzymes. The first histone demethylase identified was LSD1 [100]. To date, many histone demethylases have been identified and are classified into two families, the LSD1 demethylase family and the Jumonji (Jmj) domain-containing demethylase (JHDMs) family, depending on their domain structures and reaction mechanisms

[100, 101]. Considering the different biological consequences resulting from the various locations and levels of methylation, the histone demethylase activity needs to be highly specific and should discriminate between the different locations and status of methylation on histone lysine residues. Unlike LSD1, which can only remove mono- and dimethyl lysine modifications, the JmjC-domain-containing histone demethylases can remove all three-histone lysine-methylation states. So far, JHDMs have been shown to reverse H3K36 (JHDM1), H3K9 (JHDM2A), and both H3K9 and H3K36 (JHDM3 and JMJD2A–D) methylation [102], LSD1 directly acts on H3K4 or H3K9 modifications.

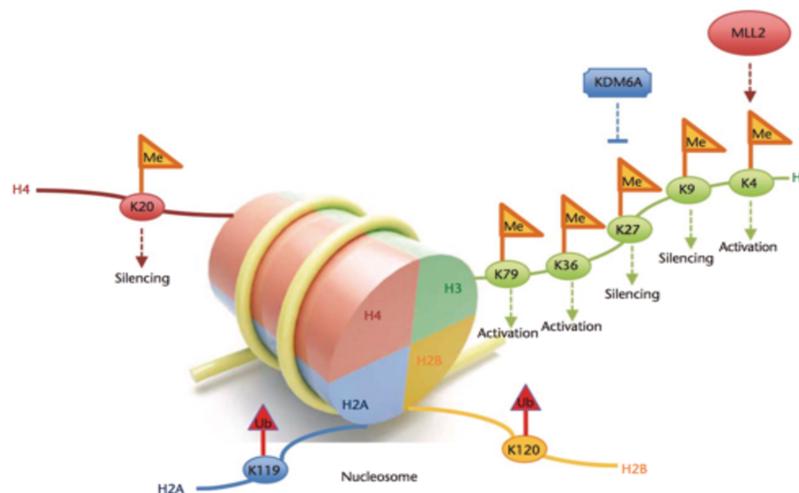


Figure 5. Nucleosomes methylase and demethylase activity. KMT2D H3K4 methylase activity on H3K4 and KDM6A demethylase activity on H3K27 [40].

Kabuki Syndrome as disorder of the epigenetic machinery

Aberrant chromatin regulation is observed in many diseases where it leads to defects in epigenetic gene regulation, involving disruption of the various components of the epigenetic machinery (writers, erasers, readers, and remodelers) causing widespread downstream epigenetic consequences [103]. Neurological dysfunction and, in particular, intellectual disability appears to be a common phenotype consequence, although these disorders can be manifested as multifactorial complexes such as congenital heart disease, autism spectrum disorders, or cancer in which mutations in chromatin regulators are contributing factors. [103, 104]. There is an emerging importance recognition of these variety disorders, that may be termed “chromatinopathies”, study of which may offer a unique opportunity to learn about the role of epigenetics in health and disease [105, 106].

The Epigenetic genes underlying cognitive disorders can be divided into four major categories:

1. Writers, that are involved in enzymatic addition of side groups (DNA methylation, RNA methylation and histone modification);
2. Erasers, are enzymes that remove the side groups;
3. Chromatin Remodelers of the DEAD/H ATPase helicase family, which are involved in the regulation of nucleosome positioning;
4. Other readers and chromatin remodelers, containing proteins that recognize and bind to their cognate chromatin marks, as well as proteins that are found in transcription repressor and activator complexes and other complexes that regulate DNA accessibility by the basic transcription machinery [103, 107] (Figure 6).

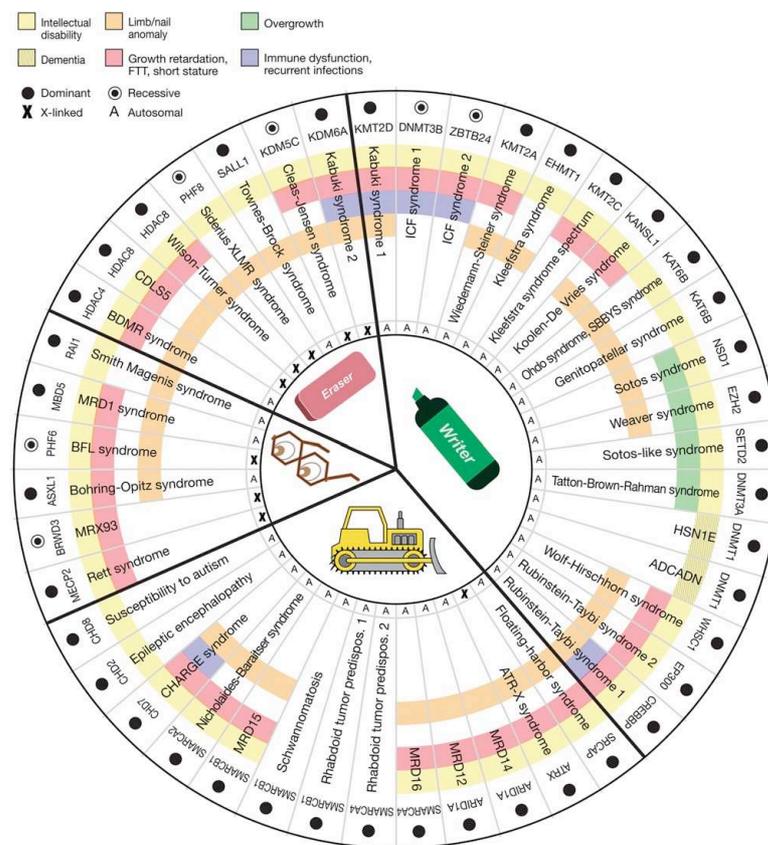


Figure 6. Features of the Mendelian disorders of the epigenetic machinery. The most common phenotype is intellectual disability (yellow). Other features include growth retardation (red), overgrowth (green), immune dysfunction (purple), and various limb abnormalities (orange). The components of the epigenetic machinery (horizontal labels) and genetic syndromes (vertical labels) are divided into four categories (writer, eraser, reader, or remodeler). The majority of these genes demonstrate dosage sensitivity (filled circle). Bjornsson et al. 2014;[103].

As part of the writer epigenetic machinery, the histone acetyltransferases (HATs) acetylate a variety of lysine residues of histone and non-histone proteins [107, 108]. De novo mutations in HATs cause human genetic disorders as the Rubinstein-Taybi syndrome, characterized by intellectual disability, postnatal growth deficiency, microcephaly, broad thumbs and halluces,

and characteristic facial appearance and caused by mutations in two HAT genes, CREBBP and EP300 [109, 110]. In contrast to acetyltransferases, Histone Methyltransferases (HMTs) have a higher degree of specificity for catalyzing the modification of specific lysine residues. This reflects the large number of histone methyltransferases (at least 27 according to the HUGO Gene Nomenclature Committee; [http:// www.genenames.org/genefamilies/KDM-KAT-KMT](http://www.genenames.org/genefamilies/KDM-KAT-KMT)) that are encoded in the mammalian genome. One of the most well characterized syndrome associated with HMT defects and intellectual disability is Kabuki Syndrome (KS), which is caused by mutations in two interacting chromatin modifiers: the writer KMT2D and the eraser KDM6A [1, 12, 103].

Mosaicism and Kabuki Syndrome

Genetic disorders are caused by changes in DNA sequence, copy number, or genomic locations (structural changes), which leads to alterations in gene expression and/or protein function. The DNA changes can be inherited; can occur newly during meiosis as the germ cells are being formed, after fertilization, either during development or in a differentiated cell. [111]. When genetic changes occur somatically, the individual is composed of cells with at least two different genotypes, and this state is known as mosaicism. From the human geneticist's point of view, there are three main types of mosaicism; somatic (occurring only in the cells of the body, but not including the germline), germline (occurring only in the germ cells or their precursors but not found elsewhere in the body) and mixed gonadal and somatic, occurring in the both the cells of the body and the germline (Figure 7) [111]. Can be very difficult to determine the type of mosaicism, since differentiation of germline from somatic mosaicism could require examination of many types of cells, which is not usually possible.

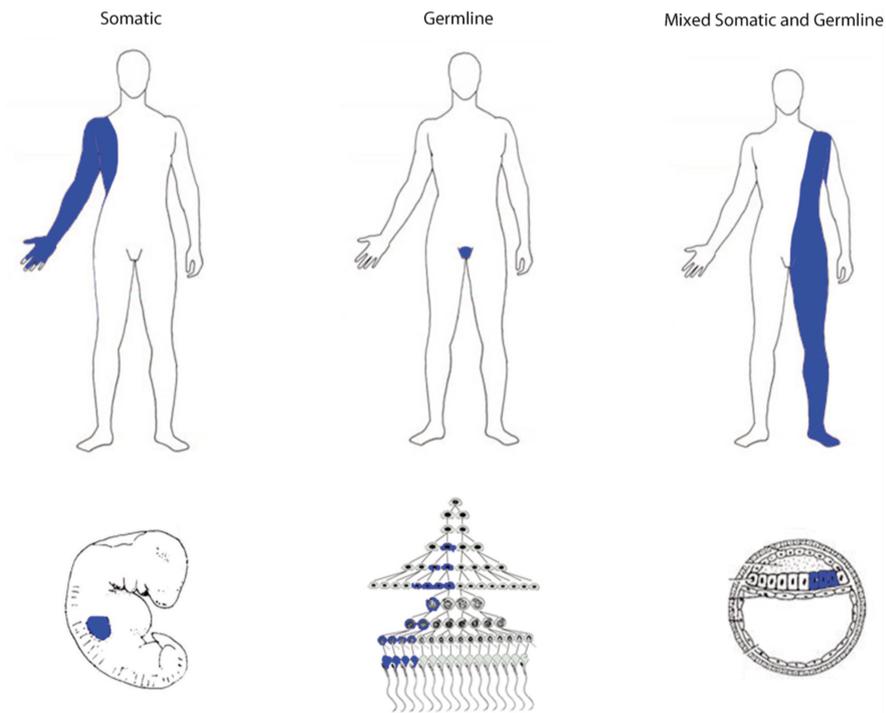


Figure 7. Different types of mosaicism. Along the top row, the variants distribution in adults with purely somatic, purely germline or mixed somatic and germline patterns. The bottom figures demonstrate possible mutation patterns in the late embryo, during spermatogenesis and in the early embryo. Spinner and Conlin 2014 AJMG [111].

To date, *KMT2D* mosaic variants have been described only in six KS patients [112, 113]; four of them with a mosaic *KMT2D* point mutation and two with a mosaic partial/whole gene deletion, while no case has been described with mosaic *KDM6A* mutations. Five out of six patients described disclosed no significant difference in clinical features from one patient that showed the lowest percentage of mutated cells in leukocytes (Figure 13A). Notably, five of the six known patients displayed a stature in the high-normal range, arguing that mosaic mutations could have a mild effect their stature. Interestingly, patients displaying a more severe delay resulted from deletions of the entire *KMT2D* gene or mutations causing a premature truncation of *KMT2D* in the first-half of the protein, predicting a shorter protein lacking the majority of *KMT2D* functional domains, including WIN and FYRC/N protein-protein interaction domains and the SET catalytic domain, therefore hampering its enzymatic activity. Clinical examination and neurobehavioral assessment in the few mosaic KS patients does not allow to draw at present any firm genotype-phenotype correlation. Nevertheless, some distinct facial dysmorphism overlapping those found in non-mosaic KS patients and mild developmental delay can be the handles for targeted genetic testing in these subjects.

Readthrough strategies for suppression of nonsense variants

An important subset of inherited genetic disorders is caused by frameshifts or nonsense variants, which generate premature termination codons (PTCs) and result in the production of truncated proteins [114, 115]. In the past years, there was an attempt to develop mutation specific pharmacological approaches aimed at achieving sufficient levels of functional proteins. One approach that has been gaining prominence is the one using pharmacological agents to promote nonsense suppression or readthrough of PTCs, thus enabling re-expression of full-length functional proteins (Figure 8) [116]. One such class of these pharmacological agents are aminoglycosides such as gentamicin, G418 and amikacin, which promote in vivo readthrough of nonsense variants leading to the recovery of the biological activity of full-length restored proteins. More recently, novel synthetic readthrough nonaminoglycosides agents with improved biocompatibility, such PTC124 and RTC13, have been developed [117-119]. Examples of studied PTC-mediated diseases are cystic fibrosis, Duchenne muscular dystrophy, ataxia telangiectasia, Rett syndrome, Usher syndrome type I, Hurler syndrome, hemophilia, methylmalonic aciduria, obesity, poor drug metabolism, spinal muscular atrophy, peroxisome biogenesis disorder, and cancer linked to the presence of PTC in p-53 tumor suppressor gene [120].

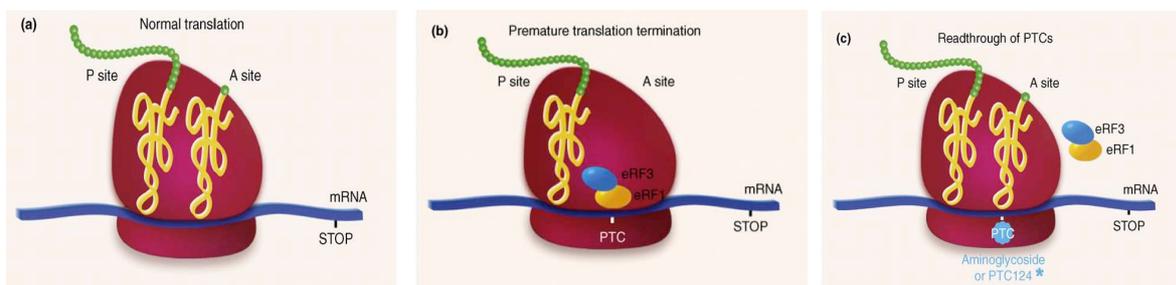


Figure 8. The effect of readthrough drugs on protein translation (a) Normal protein synthesis (b) Premature termination of protein synthesis owing to a PTC (c) Readthrough of PTC by aminoglycosides or PTC124. Modified from Lind and Kerem, 2008 [121].

Aminoglycosides are widely used in clinical practice as bactericidal antibiotics with established effects on translational accuracy or efficiency. Their utility as antibacterial agents arise from binding to the decoding site of the 16S or 18S ribosomal RNA (rRNA) interfering with protein synthesis promoting an accurate codon–anticodon pairing. In the presence of aminoglycoside, the conformation of rRNA becomes altered, inducing codon misreading that causes either incorporation of an erroneous amino acid at a sense codon or failure of recognition of the stop codon, leading to translational readthrough rather than chain

termination [122]. Although aminoglycosides target a conserved region of the rRNA sequence, these agents are highly active against bacterial and mitochondrial ribosomes but have limited interaction with human ribosomes. The specificity is related to the high affinity binding of aminoglycosides to prokaryotic rRNA, which has an adenine at position 1408 of the 16S rRNA (numbered according to the Escherichia coli sequence). In contrast, eukaryotic ribosomes have a guanine at the corresponding position, causing a low affinity towards aminoglycosides [122]. Consistently, the effects of aminoglycosides on protein translation in eucaryotic cells have been demonstrated at concentrations 10 to 15 times higher than the typical therapeutic antibacterial concentrations. The potential for aminoglycosides to induce misreading was observed in various eukaryotic cell-free systems such as yeast, plants, and human cells.

Readthrough of in-frame PTCs enables the protein synthesis to continue to the end of the transcript, thus, generating a full-length protein with either the correct or an abnormal aminoacid (bound to a tRNA that is a near-cognate of the stop codon) at the PTC. Because the binding of aminoglycosides is inefficient, both full length and truncated proteins will be synthesized. Aminoglycosides have minimal effects on normal translation termination because the normal stop codons of eukaryote genes are surrounded by upstream and downstream sequences, which enhance the efficiency of translation termination, whereas nonsense variants are usually not surrounded by these sequences [123].

Nonsense mediated mRNA decay in mammals

The nonsense-mediated decay (NMD) is an evolutionarily conserved mRNA surveillance mechanism that controls the quality of mRNAs by eliminating aberrant transcripts that prematurely terminate translation, regulating many physiological transcripts that harbor an NMD-eliciting element (uORF) or an intron within the 3' untranslated region (UTR) [124, 125]. In mammalian cells, NMD generally degrades mRNAs that terminate translation more than 50-55 nucleotides upstream of a splicing-generated exon-exon junction [126]. Although NMD is important to eliminate PTC-bearing mRNAs to minimize the production of mutant proteins, this process is not 100% efficient, often resulting in generation low levels of truncated proteins produced from residual aberrant transcripts [127]. A key issue is how the NMD machinery discriminates between an authentic termination codon and a PTC. Two different mechanistic NMD models have been proposed, (i) the exon junction complex (EJC)-dependent model and (ii) the *faux* 3' UTR (untranslated region) model [128]. Both mechanisms imply that a signal downstream of the stop codon determines whether a stop

codon is premature or not. In the EJC-dependent model, the multiprotein EJCs are deposited as a consequence of splicing, ~20-24 nucleotides upstream of each exon-exon junction [129]. EJC components include the conserved NMD factors, UPF1 (up-frameshift) UPF2 and UPF3, as well as other factors that are essential for NMD. If the ribosome runs into a stop codon that is located at least 50–55 nucleotides up-stream of an exon-exon junction, it is recognized as a PTC and the SURF complex, (consisting of the PI3K, SMG-1, UPF1, eRF1 and eRF3), is recruited to the premature termination site [130]. The binding of UPF2 to both UPF1 and SMG-1 bridges the connection between EJC and the SURF complex, which triggers the phosphorylation of UPF1 and the dissociation of eRF1 and eRF3 [130]. Phospho-UPF1 then recruits SMG-5/SMG-7 and SMG-6 proteins, resulting in mRNA decay [131, 132]. The other model, the '*faux* 3'-UTR' model, proposes that the mechanism of premature termination is intrinsically different from translation termination at a natural stop codon. Accordingly, during natural termination the ribosome is able to interact with 3'-UTR-bound proteins, while premature termination would impair or delay this interaction [127]. Hence, the position of the stop codon relative to the poly(A)-binding protein (PABP) is critical in discriminating premature from natural stop codons. Even if the exact mechanism used to identify a PTC it's not known, it is clear that the cell can distinguish between a PTC and a natural termination codon. This discrimination suggests that should be possible to identify small molecules that can promote readthrough at a PTC without affecting normal translational termination. NMD mechanism can affect drug-induced readthrough interfering with the mutated mRNAs, therefore NMD inhibition may increase PTC readthrough [121]. Because NMD occurs upstream bulk of translation, its inhibition may represent an attractive way to treat nonsense-mutation mediated genetic diseases [133].

AIM OF THE THESIS

Kabuki Syndrome is a rare autosomal, multiple malformation disorder with an estimated prevalence of 1 in 32000. Mutations in two interacting chromatin modifiers belonging to the ASCOM protein complex, *KMT2D/MLL2* and less frequently *KDM6A/UTX* have been identified as the major cause of KS, although mutations in the *RAP1A* and *RAP1B* genes have been recently proposed as additional causatives events. The large prevalence of *KMT2D* variants is predicted to lead a truncated protein, suggesting a loss of function event, and therefore haploinsufficiency as the likely mechanism for the KS phenotype. However, ~30% of *KMT2D* variants in our data set and the 16.1% from recent published mutational analysis review [18] harbour missense variants whose classification and interpretation is a challenge in molecular diagnostics and genetic counselling.

Overall this thesis aimed to: i) expand the mutation spectrum of *KMT2D* and *KDM6A* in Kabuki Syndrome ii) add some insight to the functional mechanisms that cause the disease iii) investigate the potential drug-mediated readthrough to restore *KMT2D* function as therapeutical approach and iv) afford a strategy to estimate the real deleterious effect of *KMT2D* missense variants in KS.

MATERIAL AND METHODS

Patients and samples preparation

Our cohort comprised 505 index patients clinically diagnosed as affected by Kabuki Syndrome (Table 3) recruited from 2011 to 2017. Patients were enrolled after obtaining appropriate informed consent by the physicians in charge and approval by the local ethics committees. Genomic DNAs were extracted from fresh and/or frozen peripheral blood leukocytes from patients and their available family members using an automated DNA extractor and commercial DNA extraction Kits (EZ1, Qiagen, Hilden, Germany). Total RNA was extracted from peripheral blood leukocytes using TRIZOL reagent (Life Technology) and reverse transcribed using the Quantitect Transcription kit (Qiagen), according to the manufacturer's instructions.

KMT2D and *KDM6A* sequence mutation, MLPA and pyrosequencing analysis

Mutation screening was performed for all 54 coding exons of the *KMT2D* (MIM #602113, NM_003482.3, also known as *MLL2* and *MLL4*) gene and 29 coding exons of *KDM6A* (MIM #300128, NM_021140.3). PCR amplifications were carried out in a final volume of 50 μ l consisting of:

- 50 ng genomic DNA
- 2.5 mM deoxyribonucleotides
- 15 pmol/ μ l of sense and antisense primers
- 1X Reaction Buffer (200mM Tris-HCl/ pH 8.8 at 25°), 100 mM KCl, 100mM (NH₄)₂SO₄, 1.0% Triton X-100 and 21 mg/ml nuclease-free BSA) and
- 1 U of Taq DNA polymerase.

Denaturation was carried out at 95° for 30 s, annealing at 60°C for 30 s, extension at 72° for 2 minutes/kb, for 35 cycles.

Primers were designed using the Primer 3 Output program (<http://frodo.wi.mit.edu/primer3/>) to amplify the 54 coding exons of *KMT2D* and 29 coding exons of *KDM6A* gene including the intronic flanking sequences. Amplicons and primers were checked both by BLAST and BLAT against the human genome to ensure specificity. The amplified products were subsequently purified (Exo Sap) and sequenced with a ready reaction kit (BigDyeTerminator v1.1 Cycle, Applied Biosystems). The fragments obtained were purified using DyeEx plates (Qiagen) and resolved on an automated sequencer (3130xl Genetec analyzer DNA Analyzer, ABI Prism). Sequences were analyzed using the Sequencer software (Gene Codes, Ann

Arbor, Michigan). Whenever possible the mutations identified were confirmed on a second independent blood sample. All existing and new mutations were described following the recommendations of the Human Genome Variation Society (<http://www.hgvs.org/mutnomen>). MLPA analysis was performed as reported in Priolo et al. (2012) [134] using probe mixture (SALSA MLPA KIT P389-A1 KMT2D; SALSA MLPA KIT P445-A1 KDM6A MRC-Holland, Amsterdam, The Netherlands) that contains 27 probes targeting exons across the *KMT2D* gene and 32 probes targeting exons across the *KDM6A* gene [134]. For Pyrosequencing analysis, PCR reaction was performed three times independently and after checking for amplification on a 2% agarose gel, and 10 µl of each PCR product was pyrosequenced twice on a PyroMark Q24 (Qiagen, Germany) using Pyromark PCR Kit according to the manufacturer's instructions (Qiagen, Germany). Primers were designed using PyroMark Assay Design 2.0 (Qiagen, Germany) and are available upon request.

***In silico* analysis of *KMT2D* and *KDM6A* variants**

The putative causal and functional effect of each identified nucleotide variant was estimated by using the following *in silico* prediction tools: Polyphen-2 version 2.2.2 (<http://genetics.bwh.harvard.edu/pph>) [135], Align GVGD (http://agvgd.iarc.fr/agvgd_input.php) [136], PROVEAN v1.1 (<http://provean.jcvi.org/index.php>) [137], SIFT v1.03 (<http://sift.jcvi.org/>) [138], UMD-predictor (<http://www.umd.be/>) [139], and Mutation Taster (<http://www.mutationtaster.org/>) [140] using default parameters. Splice-site variants were evaluated for putative alteration of regulatory process at the transcriptional or splicing level with Human Splice Finder (<http://www.umd.be/HSF3/>) [141], NNSPLICE (http://www.fruitfly.org/seq_tools/splice.html) [142], and NetGene2 (<http://www.cbs.dtu.dk/services/NetGene2>) [143].

***In silico* analysis of *KMT2D* protein domains**

We implemented a number of algorithms to analyze the different *KMT2D* protein domains. Initial models were generated by threading approach using five web-servers: I-Tasser (zhanglab.ccmb.med.umich.edu/I-TASSER/) {Yang, 2015 #10853, Phyre2 (www.sbg.bio.ic.ac.uk/phyre2/) {Kelley, 2015 #10854}, Genesilico (genesilico.pl/meta/) [144], RaptorX (raptorx.uchicago.edu/) [145], Multicom (http://sysbio.rnet.missouri.edu/multicom_toolbox/) [146]. Final model was obtained using

Modeller (version 9.9; salilab.org/modeller/) [147] with the five server-generated models as templates. Energy minimization of models was carried out using KoBaMIN web server (<http://csb.stanford.edu/kobamin/>) [148]. Quality of the resulting models was assessed by Qmean server (swissmodel.expasy.org/qmean/) [149]. Effects of mutations on domain structure and stability were predicted by Rosetta Backrub server (kortemmelab.ucsf.edu/backrub/) [150], FoldX software (foldx.embl.de/) [151].

Cell lines, plasmids, and transfection assays

The pFlag-CMV2 FUSION-KMT2D vector (cDNA spanning the PHD4-5-6 domains (amino acids 1358–1572) and ZF-PHD7- FYRN-FYRC- WIN-SET-post SET domains-(amino acids 4507–5537) was a gift of Professor Min Gyu Lee, Department of Molecular and Cellular Oncology, The University of Texas [152]. We sub-cloned this partial KMT2D clone into the p3XFlag-CMV14 vector (Sigma) using standard procedures. The expression plasmids harbouring patients *KMT2D* missense mutations were generated by site-directed mutagenesis according to the manufacturer's instructions (Agilent). The *KMT2D* C-terminal ORF was assembled into the pcDNA3-Myc-EGFP vector (Reymond et al. 2001) by PCR site directed amplification using human cDNA and pFlag-CMV2 FUSION-KMT2D vector as templates respectively. The *WDR5* full length ORF was assembled into the p3XFlag-CMV14 vector (Sigma) by PCR amplification reactions using as template cDNA from HEK 293T cells. HEK 293, 293T and NT2D1 cells were cultured in Dulbecco's Modified Eagle Medium (Life Technologies) with 10% FBS, penicillin (100 U/ml) and streptomycin (100 µg/ml). Patients' Fibroblasts were cultured in Dulbecco's Modified Eagle Medium F12 with 10% FBS, penicillin (100 U/ml) and streptomycin (100 µg/ml). HEK 293 and HEK293T cells were transiently transfected using the polyethylenimine method, following published protocols [153]. Cells were harvested 48 h after transfection and used for protein extraction and all the assays.

Establishment of Lymphocyte cell lines and NMD assay

Approximately 10 ml of blood in sodium heparin was processed within 24 hours from taking. Lymphocytes were isolated from whole blood of Kabuki patient by centrifugation using Lymphoprep or similar reagents, transformed by infection with EBV and maintained in RPMI medium (Gibco) containing 20% fetal bovine serum (FBS) and 1% penicillin/streptomycin, in a humidified atmosphere containing 5% carbon dioxide (CO₂) at 37°C. NMD was assayed by

treating patient lymphoblast cell lines with puromycin at a concentration of 200 ug/ml. After 8h of incubation, total RNA was obtained from lymphoblast cell lines, using the RNasy mini Kit (Qiagen) according to manufacturer instructions and a Quantitect Reverse Transcription kit (Qiagen) was used for cDNA synthesis.

Quantification of readthrough levels in HEK293 cells

A dual gene reporter pCRFL (gently provided by Prof. J-P Rousset) was used to quantify the effect of gentamicin on stop mutation readthrough in culture cells. Sequence to be analyzed spanning 27 nucleotides centered on the different stop mutations were inserted in-frame between the Renilla (LucR) and Firefly Luciferase (LucF) coding sequences. Readthrough efficiency was estimated by calculating the ratio of Renilla and Firefly activity. Both reporter proteins are expressed from the same message, eliminating in particular potential variation in mRNA stability between the different targets analyzed. A 100% activity control was provided by a construct (TQ) with no stop codon between the coding sequences of the two reporters [154]. A pCRFL reporter vector harboring the 319d Duchenne muscular dystrophy mutation (pCRFL319) exhibiting the highest gentamicin-induced readthrough efficiency and the highest induction factor in NIH3T3 cultured cells assays [155] was used as positive control.

Site-directed mutagenesis

In vitro approaches to site-directed mutagenesis can be grouped generally into three categories:

- i) methods that restructure fragments of DNA, such as cassette mutagenesis;
- ii) localized random mutagenesis;
- iii) oligonucleotide-directed mutagenesis;

All oligonucleotide-directed mutagenesis is based on the same concept: an oligonucleotide pair encoding the desired mutation(s), is annealed to the DNA of interest and serves as a primer for initiation of DNA synthesis. The p3XFlag-KMT2D-FUSION-CMV14 was mutated to insert in-frame missense variant/s corresponding to the variant/s found in our patients with Kabuki Syndrome. All site-directed mutagenesis reactions were performed by using Pfu polymerase from Promega.

50 ng of dsDNA template was added to 10X reaction buffer, 2.5mM dNTP mix, 125 ng of each oligonucleotide primer and 2.5 U/μl Pfu DNA polymerase.

After initial denaturation at 95° for 1 min, the cycling parameters were:

- 30 seconds at 95°C followed by
- 1 min at 55°C and
- 2 minutes/kb of plasmid length at 72°C (12 cycles).

Then the parental, supercoiled double-stranded DNA was digested with 10U of Dpn I restriction enzyme at 37°C for 1 h before being transformed.

Transformation

Bacterial transformation is the process by which electrocompetent or chemically competent bacterial cells take up naked molecules. If the foreign DNA has an origin of replication recognize by the host cell DNA polymerases, the bacteria will replicate the foreign DNA along with their own DNA. When transformation is coupled with antibiotic selection techniques, bacteria can be induced to uptake certain DNA molecules, and those bacteria can be selected for that incorporation. All constructs were transformed in DH5 alpha Competent Cells (Invitrogen). 10 µl of the mutagenesis reaction mixture was added to a 120 µl of cells, and the tubes were incubated for 30' on ice. Cells were heat-shocked in a 42° water bath for 45 seconds, then incubated on ice for 2 minutes, before being additionated with 250 µl LB (10g Bacto-Tryptone, 5g Bacto-yeast extract, 10g NaCl, dd H2O to 1 litre) medium and incubated at 37°C for 1 hour with shaking at 225-250 rpm. 150 µl of the transformation mixture were plated on LB agar plates containing the appropriate antibiotic (Ampicillin, 100 µg/ml).

Screening for recombinants

Screening for recombinants was performed by colony PCR or plasmid miniprep followed by restriction digestion.

i) Colony PCR: this method is designed to quickly screen for plasmid inserts directly from E.coli colonies. Typical colony PCR reaction:

Mix together the following on ice; always adding enzyme last. For multiple samples make a large master mix and aliquot 50 ul in each PCR tube:

- 38 µl sterile distilled water
- 5 µl 10X PCR buffer
- 3 µl 25mM MgCl₂
- 1 ul 10mM dNTPs
- 1 µl 20µM forward primer

-1 µl 20µM reverse primer

-0.2 µl Taq polymerase

To each cold PCR tube containing the PCR reaction, a small amount of colony was added.

ii) plasmid miniprep followed by restriction digestion: well-isolated colonies were picked from a plate and cultured for 8 hours at 37°C in LB culture medium containing the appropriate antibiotic for selection. Miniprep were performed by using GenElute Plasmid Miniprep Kit (Sigma-Aldrich).

Plasmid DNA extraction

Plasmid purification protocols are based on a modified alkaline lysis procedure followed by binding of plasmid DNA to anion-exchange resin under appropriate low-salt and pH conditions. RNA, proteins, dyes, and low-molecular-weight impurities are removed by a medium-salt wash. Plasmid DNA is eluted in a high-salt buffer and then concentrated and desalted by isopropanol precipitation. Starter culture was prepared by inoculating positive colony into 2-5ml of selective LB medium and incubated for 8 hours at 37°C with vigorous shaking (approximately 300 rpm). 1ml of starter culture was diluted in 100 ml of selective LB medium, and grown for 12-16 hours at 37°C, then bacterial cells were harvested by centrifugation at 3500g for 30 minutes at 4°C. Plasmid purification was performed by using QIAGEN Plasmid Midi Kit. Bacterial pellet was resuspended in 4ml of P1 buffer (Resuspension Solution) which contains Rnase A (100 µg/ml). Lysis was performed by incubating for 5 minutes the resuspended pellet with 4 ml of P2 buffer (Lysis buffer) after inverting the tube for mixing, and then reaction was stopped by the addition of 4ml of P3 buffer (Neutralization Buffer). Tubes were incubated on ice for 15 minutes, then cell debris, proteins and genomic DNA were precipitated by centrifugation at 15000g for 30 minutes at 4°C. Column resin was equilibrated applying 4ml of QBT buffer (Equilibration buffer), allowing the column to empty by gravity flow. After centrifugation, supernatant containing plasmid DNA was promptly removed and loaded into the column, allow the lysate to enter the resin by gravity flow. Column resin was washed twice with 10 ml of QC buffer (Wash buffer) to remove all contaminants in the DNA plasmid preparation. Plasmid DNA was eluted by adding 5 ml of QF Elution buffer in a new tube, then precipitated with 3,5 ml of room-temperature isopropanol, mixed and centrifuged at 15000g for 30 minutes at 4°C. DNA pellet was washed with 2 ml of room-temperature 70% ethanol, centrifuged at 15000 g for 10 minutes. After removal of supernatant, pellet was air-dried, then resuspended in 100-200 µl of H₂O.

Total RNA Isolation and Reverse Transcription (RT-PCR)

Total RNA was isolated using the TRIzol® protocol (Gibco) from lymphoblast cell lines.

RNA extraction with TRIzol is a common method for total RNA extraction from cells. The correct name of the method is Guanidinium thiocyanate-phenol-chloroform extraction and uses guanidinium isothiocyanate, a powerful chaotrope used as a protein denaturant, which causes the inactivation of RNases, and acidic phenol/chloroform, which lead to the partitioning of RNA into aqueous supernatant for separation. Cells were homogenized in 1 ml Trizol and incubated for 5 min at RT. 200 µl of chloroform was added, then shaken vigorously by hand for 15", incubated for 2-3 minutes at RT and centrifuge at 12,000 rpm for 15 minutes at 4°C. Aqueous (top) phase was transferred to a fresh sterile 1,5 ml microcentrifuge tube. 500 µl isopropyl alcohol was added, then samples were incubated for 10' RT and centrifuged at 12,000 rpm for 5 minutes at 4°C. Pellet was washed in 1 ml 75% EtOH, RNA pellet was briefly air dry, and dissolved in 50 µl RNase free water. 1 µg RNA was reverse transcribed using the Quantitect Transcription kit (Qiagen). RT-PCR produces DNA copies (complementary DNA, or cDNA) of a RNA template, by using the enzyme reverse transcriptase, and the resulting single-stranded cDNA can be amplified using traditional or real-time PCR. Reverse transcriptase enzyme, in general, has 3 distinct enzymatic activities: an RNA-dependent DNA polymerase, a hybrid-dependent exoribonuclease (Rnase H), and a DNA-dependent DNA polymerase. For reverse transcription in vitro, the first 2 activities are utilized to produce single-stranded cDNA: RNA-dependent DNA-polymerase activity (reverse transcription) transcribes cDNA from an RNA template, and RNase H activity specifically degrades only the RNA in RNA: DNA hybrids. The purified RNA samples were incubated in 1X gDNA Wipeout Buffer at 42°C for 2 minutes to remove contaminating genomic DNA. After genomic DNA elimination, the RNA samples were reverse transcribed using a master mix prepared from Reverse Transcriptase (1U), 1X RT Buffer, and RT Primer Mix, a optimized blend of oligo-dT and random primers. The entire reaction was performed at 42°C for 30 minutes and then inactivated at 95°. Each cDNA sample was measured by using a Nanodrop spectrophotometer (NanoDrop Technologies, Willmington, Delaware, USA) and used in qPCR for NMD-assay and in PCR for studying splicing mutations.

Real-time polymerase chain reaction (qPCR)

qPCR is a quantitative PCR method which enables both detection and quantification (as absolute number of copies or relative amount when normalized to DNA input or additional normalizing genes) of one or more specific sequences in a DNA sample.

Based on the molecule used for the detection, the real time PCR techniques can be categorically placed under two heads:

- non-specific fluorescent dyes that intercalate with double-stranded DNA, such as SYBR Green, which binds to the minor groove of the DNA double helix, and is the most widely used double-strand DNA-specific dye reported for real time PCR
- sequence-specific DNA probes consisting of oligonucleotides that are labeled with a fluorescent reporter, which permits detection only after hybridization of the probe with its complementary DNA target, such as Molecular Beacons, TaqMan Probes, FRET Hybridization Probes, Scorpion Primers. Oligos for qPCR were designed using the Primer3 program⁶ with default parameters. Amplicons and primer pairs were checked both by Blast and Blat against the human genome to ensure specificity. Target genes expression was examined by amplification with the primer sets described in the table below (Table 2). *EEF1A1* and *GAPDH* were used as housekeeping genes. The reactions were run in triplicate in 10 µl of final volume with 10 ng of sample cDNA, 0.3 mM of each primer, and 1X Power SYBR Green PCR Master Mix (Applied Biosystems). Reactions were set up in a 384-well plate format with a Biomeck 2000 (Beckmann Coulter, Milan, Italy) and run in an ABI Prism7900HT (Applied Biosystems) with default amplification conditions. Raw Ct values were obtained using SDS 2.3 (Applied Biosystems). Calculations were carried out by the comparative Ct method.

Table 2: Oligos used for splicing mutations analysis

Primer	Sequence 5'-3'
KMT2D_EX2_F	TCTCTGTCCTTAGTTCTGGGAGT
KMT2D_EX5_R	GGAAGTGGTAAAGCCGTGGA
KMT2D_EX40_F	TACAGAAGGCAAGCGACAGG
KMT2D_EX44_R	AGGAATGAGGGGGTGAC
KMT2D_EX16_F	TATGCAGTGTGGGGCTGCTTC
KMT2D_EX19_R	AGCTCATCGGTGTCCAGGTGG
KMT2D_EX42_F	CAAACCTGGTAGGTGGGAGGA

KMT2D_EX42_R	GGCCCCATAAGGTTTGGTAT
KMT2D_EX43+44_F	TTCTCCTGACAGCATTGTGC
KMT2D_EX45_R	TCTAGCCCAGGCTTTCACAT
KMT2D_EX45_F	TTCCCAGATACCAAACCTTATG
KMT2D_EX48_R	GCACTCCTTCCATTTCTTGAG

***In vitro* histone methyltransferase (HMT) assay and epigenetic reporter allele**

Partially purified FLAG-KMT2D wild-type and mutant derivative proteins were obtained from transfected HEK 293T cells by lysis in co-IP buffer (50 mM Tris, pH 7.5, 250 mM NaCl, 1% TritonX-100, 1 mM EDTA), followed by overnight incubation with EZview™ Red Anti-FLAG Affinity Gel (Sigma-Aldrich) at 4 °C and final elution in BC100 buffer (20 mM Tris pH 7.5, 10% Glycerol, 0.2 mM EDTA, 1% TritonX-100, 100 mM NaCl) containing FLAG peptide (Sigma-Aldrich). KMT2D protein amounts were quantified by Coomassie staining and immunoblot analysis using mouse monoclonal FLAG antibodies. Enzymatic activity against native nucleosomes was measured following a published method [152]. Briefly, equal amounts of wild-type or mutant FLAG-KMT2D proteins were incubated at 37 °C for 4 h with HeLa nucleosomes (Reaction Biology) in KMT buffer (50 mM Tris pH 8.5, 100 mM KCl, 5 mM MgCl₂, 10% glycerol, 4 mM DTT) supplemented with *S*-adenosyl methionine (New England BioLabs). Reactions were stopped by adding equal volumes of 2× Laemmli buffer and heated at 100 °C for 5 min before loading onto Tris-glycine 4–20% gradient gels. All assays were performed at least twice independently. Epigenetic reporter allele, a kind gift from Dr Hans T. Bjornsson (McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore) has been used as described in [103].

Co-immunoprecipitation assay and Western Blotting analysis

Co-immunoprecipitations were performed using Dynabeads magnetic beads (Thermo Fisher Scientific) following manufacturer's instructions. Complexes were analyzed by western blot using the indicated antibodies. Protein extracts were resolved on NuPAGE Tris-acetate 3–8% gels (for KMT2D) or Tris-glycine 4–20% gels (for histone H3) (Life Technologies) and transferred to nitrocellulose membranes (GE Healthcare) according to the manufacturer's instructions. Antibodies used were: mouse monoclonal antibody to α -tubulin (clone DM1A, Sigma-Aldrich), rabbit polyclonal anti-H3K4me1 (Abcam), anti-H3K4me2 (Active Motif), anti-H3K4me3 (Abcam), rabbit monoclonal anti-Histone H3 (clone D1H2, Cell Signaling

Technology), mouse monoclonal anti-Flag (Sigma cat# F3165), rabbit monoclonal anti GFP (Santa Cruz), rat monoclonal anti-HA (Roche), and mouse monoclonal anti Myc (Roche), rabbit polyclonal anti Ash2 (Bethyl) and rabbit polyclonal anti RbBP5 (Bethyl). Horseradish peroxidase conjugated anti-mouse (Santa Cruz), anti-rabbit (Santa Cruz) antibodies, and the ECL chemiluminescence system (GE Healthcare) was used for detection. Quantitation of band signal intensity was obtained by the ImageJ software (<http://imagej.nih.gov/ij/>). Values are expressed as fold differences relative to the wild-type protein sample, set at 1, after normalization for the loading control.

RNAi and NT2D1 cell differentiation

HEK293 cells were transfected with shKMT2D #1 (SHCLNV-NM_003482; TRCN0000235745), shKMT2D #2 (SHCLNV-NM_003482; TRCN0000013140), shKMT2D #1 + shKMT2D #2, or control shRNA (shLuc) that were all purchased from Sigma (Mission shRNA Lentiviral transduction particles). These shRNA plasmids, which contain a puromycin-resistant marker, were cotransfected along with a packing plasmid (deltaR8) and an envelope plasmid (VSV-G) into HEK293 cells using a calcium phosphate method. Thirteen hours later, the medium was replaced with DMEM supplemented with 10% FBS. Virus particles containing shKMT2D or shLuc were generated for 2 days and used to infect NT2D1 mammalian cells. NT2D1 cells were infected by virus-containing medium, and the infected cells were selected in medium containing 2.5 mg/ml puromycin for 2 days. Knockdown efficiency was examined by Western blot analysis and qRT-PCR. Knockdown cells were treated with 10 mM RA for 0–6 days, monitored for cell differentiation patterns, and harvested for further analysis.

RESULTS

Mutation Screening of *KMT2D* and *KDM6A*

Since 2011 we performed the mutational screening of our KS cohort, that include now 505 KS patients, by Sanger sequencing and MLPA of coding sequence of *KMT2D*, followed by *KDM6A* analysis in those patients resulted as *KMT2D*-negative. Of these 505 patients, 303 were analysed in previous studies where 140 pathogenic *KMT2D* variants were identified [4, 134, 156-161]. Here, we extended the *KMT2D* mutational analysis to a new cohort of 202 individuals identifying 64 patients carrying causative *KMT2D* mutations. Overall, we identified 196/505 (39%) patients with *KMT2D* mutations and 208 different *KMT2D* mutations (Table 3); of them 37/208 (18%) have never been described before. The following types of *KMT2D* mutations have been identified: 54 nonsense (54/208, 26%), 59 frameshift (59/208, 28%), 69 missense (69/208, 33%), 13 splice site (13/208, 6%), 12 indel (12/208, 6%) and 1 gross deletion (Table 3, Figure 9). Among the *KMT2D* identified variants, 66 occurred *de novo* and 32 mutations were inherited from an apparently asymptomatic parent. In our case series, by Sanger sequencing and MLPA we identified 14 *KDM6A* variants, of which 12 are predicted being pathogenic and 7 were never described before. Seven of the mutations are *de novo*, for the remaining we did not have access to the parental DNA. Still unsolved remains 294 cases clinically diagnosed as Kabuki Syndrome affected.

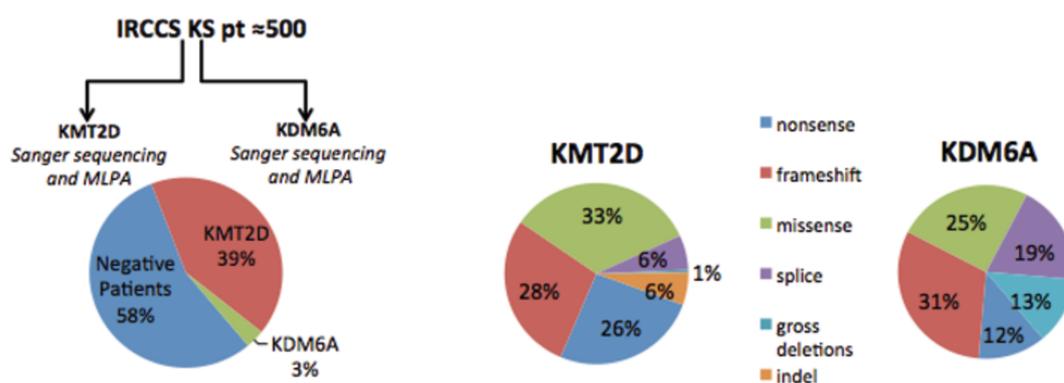


Figure 9. Percentage of *KMT2D* and *KDM6A* variants found in our cohort. On the left, the percentage of patients mutated in *KMT2D* or *KDM6A*. On the right, the different types of variants found in our data set.

Table 3: *KMT2D* and *KDM6A* variants identified in our cohort.

ID	Inheritance	Exon/ Intron	Variant	AA change	Reference	ACMG Classification
KMT2D						
Nonsense						
KB49	NA	ex 5	c.669T>G	p.(Tyr223*)	4	P
KB343	NA	ex 8	c.1016G>A	p.(Trp339*)	This study	P
KB35	NA	ex 10	c.1921G>T	p.(Glu641*)	4	P
KB33	NA	ex 16	c.4419G>A	p.(Trp1473*)	4	P
KB63	NA	ex 19	c.4895delC	p.(Ser1632*)	4	P
KB317	NA	ex 22	c.5212G>T	p.(Glu1738*)	13	P
KB336	de novo	ex 22	c.5269C>T	p.(Arg1757*)	11	P
KB262	NA	ex 26	c.5674C>T	p.(Gln1892*)	13	P
KB429	NA	ex 26	c.5707C>T	p.(Arg1903*)	6, 16, 11, 23	P
KB26	NA	ex 31	c.6295C>T	p.(Arg2099*)	1, 4, 23	P
KB502	de novo	ex 31	c.7228C>T	p.(Arg2410*)	6, 5, 7, 9	P
KB66	NA	ex 31	c.7246C>T	p.(Gln2416*)	4	P
KB59	NA	ex 31	c.7903C>T	p.(Arg2635*)	4, 23	P
KB153	de novo	ex 31	c.7903C>T	p.(Arg2635*)	4, 23	P
KB226	de novo	ex 31	c.7903C>T	p.(Arg2635*)	4, 23	P
KB338	de novo	ex 31	c.7933C>T	p.(Arg2645*)	2, 22	P
KB198	de novo	ex 31	c.7936G>T	p.(Glu2646*)	13	P
KB352	NA	ex 32	c.8227C>T	p.(Gln2743*)	This study	P
KB323	NA	ex 33	c.8311C>T	p.(Arg2771*)	2, 7, 22	P
KB289	NA	ex 34	c.8743C>T	p.(Arg2915*)	3, 15, 20, 23	P
KB422	de novo	ex 34	c.9396C>A	p.(Cys3132*)	This study	P
KB186	de novo	ex 34	c.9961C>T	p.(Arg3321*)	1, 5, 7, 20	P
KB56	de novo	ex 34	c.10135C>T	p.(Gln3379*)	4	P
KB168	de novo	ex 39	c.10750C>T	p.(Gln3584*)	13	P
KB46	de novo	ex 39	c.10841C>G	p.(Ser3614*)	4	P
KB41	NA	ex 39	c.11119C>T	p.(Arg3707*)	4	P
KB44	NA	ex 39	c.11119C>T	p.(Arg3707*)	4	P
KB42	de novo	ex 39	c.11269C>T	p.(Gln3757*)	4, 23	P
KB25	NA	ex 39	c.11434 C>T	p.(Gln3812*)	4	P
KB244	de novo	ex 39	c.11674 C>T	p.(Gln3892*)	7	P
KB178	NA	ex 39	c.11704C>T	p.(Gln3902*)	13	P
KB425	de novo	ex 39	c.11731C>T	p.(Gln3911*)	This study	P
KB461#	NA	ex 39	c.11749C>T	p.(Gln3917*)	This study	P
KB463	NA	ex 39	c.11845C>T	p.(Gln3949*)	This study	P
KB181	NA	ex 39	c.11869C>T	p.(Gln3957*)	13	P
KB358	NA	ex 39	c.11944C>T	p.(Arg3982*)	2, 11, 23	P
KB40	NA	ex 39	c.12274C>T	p.(Gln4092*)	4, 11	P
KB114	de novo	ex 39	c.12274C>T	p.(Gln4092*)	4, 11	P
KB65	NA	ex 39	c.12076C>T	p.(Gln4026*)	4	P
KB333	NA	ex 39	c.12703C>T	p.(Gln4235*)	1	P
KB 410	NA	39	c.12760C>T	p.(Gln4254*)	23	P
KB82	de novo	ex 39	c.12844C>T	p.(Arg4282*)	13, 22	P
KB350	de novo	ex 39	c.12844C>T	p.(Arg4282*)	13, 22	P
KB189	de novo	ex 39	c.12955A>T	p.(Arg4319*)	13, 23	P
KB183	de novo	ex 39	c.13450C>T	p.(Arg4484*)	2, 9, 16, 23	P
KB450	NA	ex 39	c.13450C>T	p.(Arg4484*)	2, 9, 16, 23	P
KB175	de novo	ex 39	c.13507C>T	p.(Gln4503*)	13	P
KB73	de novo	ex 40	c.13666A>T	p.(Lys4556*)	4	P
KB83	NA	ex 48	c.15022G>T	p.(Glu5008*)	13	P
KB377	NA	ex 48	c.15061C>T	p.(Arg5021*)	7, 11	P
KB45	NA	ex 48	c.15079C>T	p.(Arg5027*)	2, 4, 23	P
KB72	NA	ex 48	c.15079C>T	p.(Arg5027*)	2, 4, 23	P
KB362	NA	ex 50	c.16018C>T	p.(Arg5340*)	2	P
KB130	NA	ex 52	c.16360C>T	p.(Arg5454*)	1, 2, 5, 23, 22	P
Frameshift						
KB454	NA	ex 3	c.234_235delGC	p.(Gln79Alafs*7)	This study	P
KB469	NA	ex 3	c.345dupA	p.(Ser116Ilefs*7)	This study	P
KB337	NA	ex 4	c.446_449delTATG	p.(Val149Glyfs*58)	This study	P
KB75	de novo	ex 4	c.472delT	p.(Cys158Valfs*50)	4	P
KB8	de novo	ex 5	c.588delC	p.(Cys197Alafs*11)	9	P
KB58	NA	ex 6	c.705delA	p.(Glu237Serfs*24)	4	P
KB57	NA	ex 8	c.1035_1036delCT	p.(Cys346Serfs*17)	4	P
KB89	NA	ex 10	c.1345_1346delCT	p.(Leu449Valfs*5)	4, 16	P
KB156	de novo	ex 10	c.1503dupT	p.(Pro502Serfs*7)	13	P
KB116	NA	ex 10	c.1634delT	p.(Leu545Argfs*385)	7	P
KB349	NA	ex 10	c.1634delT	p.(Leu545Argfs*385)	7	P
KB 545	NA	10	c.2091dupC	p.(Thr698Hisfs*6)	This study	P
KB369	NA	ex 11	c.3596_3597del	p.(Leu1199Hisfs*7)	This study	P
KB48	de novo	ex 11	c.2993dupC	p.(Met999Tyrfs*69)	4	P
KB203	NA	ex 11	c.3161_3171del	p.(Pro1054Hisfs*10)	17, 13, 18	P
KB309	NA	ex 11	c.3730delG	p.(Val1244Serfs*86)	13	P
KB142	de novo	ex 13	c.4021delG	p.(Val1341Leufs*35)	13	P

KB311	NA	ex 14	c.4135_4136delAT	p.(Met1379Valfs*52)	13, 14, 23	P
KB 524	NA	14	c.4135_4136delAT	p.(Met1379Valfs*52)	13,14, 23	P
KB188	de novo	ex 16	c.4454delC	p.(Pro1485Leufs*21)	13	P
KB159	NA	ex 19	c.4896_4905del	p.(Asp1633Alafs*86)	13	P
KB443#	de novo	ex 25	c.5575delG	p.(Asp1859Thrfs*17)	This study	P
KB3	NA	ex 26	c.5652dup	p.(Lys1885Glnfs*18)	9	P
ID	Inheritance	Exon/ Intron	Variant	AA change	Reference	ACMG Classification
Frameshift						
KB84	NA	ex 26	c.5779delC	p.(Gln1927Lysfs120*)	4	P
KB146	de novo	ex 27	c.5857delC	p.(Leu1953Trpfs*94)	13	P
KB208	NA	ex 28	c.5954delC	p.(Thr1985Lysfs*62)	13	P
KB221	de novo	ex 29	c.6149_6150del	p.(Arg2050Lysfs*6)	13	P
KB 525	NA	30	c.6212_6213del	p.(His2071Profs*10)	This study	P
KB152	de novo	ex 31	c.6583delA	p.(Thr2195Profs*69)	13	P
KB267	NA	ex 31	c.6594delC	p.(Tyr2199Ilefs*65)	5	P
KB79	de novo	ex 31	c.6595delT	p.(Tyr2199Ilefs*65)	1, 3, 4, 5, 7, 9, 19, 20, 23	P
KB102	de novo	ex 31	c.6595delT	p.(Tyr2199Ilefs*65)	1, 3, 4, 5, 7, 9, 19, 20, 23	P
KB342	NA	ex 31	c.6595delT	p.(Tyr2199Ilefs*65)	1, 3, 4, 5, 7, 9, 19, 20, 23	P
KB67	de novo	ex 31	c.6638_6641del	p.(Gly2213Alafs*50)	4	P
KB176	NA	ex 31	c.6738delA	p.(Lys2246Asnfs*18)	13	P
KB253	NA	ex 31	c.6794delG	p.(Gly2265Glufs*21)	13, 23	P
KB278	NA	ex 31	c.7481dupT	p.(Ala2496Serfs*10)	13, 20	P
KB313	de novo	ex 32	c.8196delG	p.(Ser2733Valfs*24)	13	P
KB80	NA	ex 33	c.8273delG	p.(Gly2758Alafs*29)	4	P
KB243	de novo	ex 34	c.8430_8431insAA	p.(Gln2811Asnfs*41)	13	P
KB182	NA	ex 34	c.9203delA	p.(Gln3068Glyfs*3)	13	P
KB30	NA	ex 38	c.10606delC	p.(Arg3536Alafs*122)	4	P
KB101	de novo	ex 39	c.11066_11078del	p.(Ala3689Valfs*56)	4	P
KB504	de novo	ex 39	c.11093dupG	p.(Phe3699Leufs*14)	This study	P
KB495	de novo	ex 39	c.11715delG	p.(Gln3905Hisfs*74)	This study	P
KB172	NA	ex 39	c.12647delC	p.(Pro4216Leufs*62)	13	P
KB192	de novo	ex 39	c.12966delA	p.(Gln4322Hisfs*62)	13	P
KB54	NA	ex 39	c.13129dupT	p.(Trp4377Leufs*33)	4	P
KB121	de novo	ex 39	c.13277dupT	p.(Ala4428Serfs*59)	13	P
KB 540	NA	Ex 41	c.13780delG	p.(Ala4594Profs*23)	23	P
KB123	de novo	ex 42	c.13884dupC	p.(Thr4629Hisf*18)	13	P
KB481	NA	ex 42	c.13895dupC	p.(Ser4633Ilefs*14)	This study	P
KB197	de novo	ex 47	c.14592dupG	p.(Pro4865Alafs*48)	13	P
KB125	NA	ex 48	c.15031delG	p.(Glu5011Serfs*40)	13	P
KB16	de novo	ex 48	c.15374dupT	p.(Phe5126Leufs*12)	9	P
KB 535	NA	Ex 50	c.16043_16044del	p.(His5348Leufs*14)	This study	P
KB355	NA	ex 53	c.16438_16441del	p.(Asn5480Val*6)	5	P
KB64	NA	ex 53	c.16469_16470del	p.(Lys5490Argfs*21)	4	P
KB 533	NA	Ex 53	c.16469_16470del	p.(Lys5490Argfs*21)	4	P
Missense						
KB21#	inherited M	ex 3	c.346T>C	p.(Ser116Pro)	13	LB
KB21#^	inherited M	ex 4	c.510G>C	p.Gln170His	9, 13	P
KB256	NA	ex 5	c.626C>T	p.(Thr209Ile)	13	VOUS
KB458	NA	ex 8	c.1076G>C	p.(Arg359Pro)	This study	VOUS
KB269	inherited M	ex 10	c.1940C>A	p.(Pro647Gln)	3,9,13	LB
KB126	NA	ex 10	c.2074C>A	p.(Pro692Thr)	2	LB
KB374#	inherited M	ex 10	c.2074C>A	p.(Pro692Thr)	2	LB
KB487	NA	ex 10	c.2654C>T	p.(Pro885Leu)	This study	VOUS
KB370#	inherited P -M	ex 11	c.2837C>G	p.(Ala946Gly)	This study	LB
KB215	inherited M	ex 11	c.3392C>T	p.(Pro1131Leu)	13	LB
KB341	inherited M	ex 11	c.3392C>T	p.(Pro1131Leu)	13	LB
KB222	inherited P	ex 11	c.3572C>T	p.(Pro1191Leu)	13	LB
KB32	inherited P	ex 11	c.3773G>A	p.(Arg1258Gln)	4	LB
KB307	de novo	ex 14	c.4171G>A	p.(Glu1391Lys)	13, 23	LP
KB28#	inherited M	ex 15	c.4249A>G	p.(Met1417Val)	4	LB
KB28#	inherited M	ex 15	c.4252C>A	p.(Leu1418Met)	4	LB
KB174	inherited M	ex 15	c.4283T>C	p.(Ile1428Thr)	13	LB
KB138	NA	ex 16	c.4427C>G	p.(Ser1476Cys)	13	VOUS
KB34	inherited P	ex 16	c.4565A>G	p.(Gln1522Arg)	4	LB
KB 535	NA	Ex 25	c.5549G>A	p.(Gly1850Asp)	This study	LB
KB119	inherited M	ex 31	c.6638G>A	p.(Gly2213Asp)	13	LB
KB204#	inherited M	ex 31	c.6638G>A	p.(Gly2213Asp)	13	LB
KB330#	NA	ex 31	c.6733C>G	p.(Leu2245Val)	This study	VOUS
KB326#	inherited M	ex 31	c.6811C>T	p.(Pro2271Ser)	13	LB
KB107#	NA	ex 31	c.6970C>A	p.(Pro2324Thr)	13	VOUS
KB430	NA	ex 31	c.7378C>T	p.(Arg2460Cys)	2	LB
KB122	inherited M	ex 31	c.7829T>C	p.(Leu2610Pro)	6*, 13	LB
KB287	inherited M	ex 31	c.7829T>C	p.(Leu2610Pro)	6*, 13	LB
KB27	NA	ex 34	c.8521C>A	p.(Pro2841Thr)	4	VOUS
KB330#	NA	ex 34	c.8774C>T	p.(Ala2925Val)	This study	VOUS
KB443#	inherited M	ex 34	c.9971G>T	p.(Gly3324Val)	This study	LB
KB326#	inherited P	ex 34	c.10192A>G	p.(Met3398Val)	13	LB

KB357	inherited P	ex 34	c.10192A>G	p.(Met3398Val)	13	LB
KB297	NA	ex 35	c.10256A>G	p.(Asp3419Gly)	2	LB
KB378	NA	ex 35	c.10256A>G	p.(Asp3419Gly)	2	LB
KB292	de novo	ex 37	c.10499G>T	p.(Gly3500Val)	13	LP
KB86#	NA	ex 39	c.10966C>T	p.(Arg3656Cys)	13	VOUS
KB374#	inherited P	ex 39	c.11380C>T	p.(Pro3794Ser)	This study	LB
KB293	inherited P	ex 39	c.11794C>G	p.(Gln3932Glu)	13	LB
KB204#	inherited P	ex 39	c.12070A>G	p.(Lys4024Glu)	13	LB
KB385	NA	ex 39	c.12302A>C	p.(Gln4101Pro)	This study	VOUS
KB247	NA	ex 39	c.12485G>A	p.(Arg4162Gln)	13	VOUS
KB107#	NA	ex 39	c.12488C>T	p.(Pro4163Leu)	13	VOUS
ID	Inheritance	Exon/ Intron	Variant	AA change	Reference	ACMG Classification
Missense						
KB170	inherited M	ex 39	c.13256C>T	p.(Pro4419Leu)	13	LB
KB 512	NA	ex 45	c.14381A>G	p.(Lys4794Arg)	This study	VOUS
KB86#	NA	ex 48	c.14893G>A	p.(Ala4965Thr)	13	VOUS
KB38#	de novo	ex 48	c.15084C>G	p.(Asp5028Glu)	4	LP
KB154	NA	ex 48	c.15088C>T	p.(Arg5030Cys)	9, 11	VOUS
KB185	de novo	ex 48	c.15088C>T	p.(Arg5030Cys)	9, 11	LP
KB423	inherited P	ex 48	c.15089G>A	p.(Arg5030His)	This study	LB
KB38#	de novo	ex 48	c.15100T>G	p.(Phe5034Val)	4	LP
KB76	de novo	ex 48	c.15176A>C	p.(His5059Pro)	4	LP
KB129	NA	ex 48	c.15292A>C	p.(Thr5098Pro)	13	VOUS
KB462	de novo	ex 48	c.15310T>C	p.(Cys5104Arg)	This study	LP
KB171	inherited M	ex 48	c.15565G>A	p.(Gly5189Arg)	4, 11, 23	VOUS
KB264	NA	ex 48	c.15640C>T	p.(Arg5214Cys)	5, 7, 9, 23	LP
KB376	NA	ex 48	c.15640C>T	p.(Arg5214Cys)	5, 7, 9, 23	LP
KB24	de novo	ex 48	c.15641G>A	p.(Arg5214His)	1	LP
KB219	NA	ex 48	c.15641G>A	p.(Arg5214His)	1,5	LP
KB408	de novo	ex 48	c.15641G>A	p.(Arg5214His)	1,5	LP
KB109	de novo	ex 48	c.15649T>C	p.(Trp5217Arg)	13	LP
KB17	de novo	ex 50	c.16019G>A	p.(Arg5340Gln)	4, 23	LP
KB169	de novo	ex 51	c.16273G>A	p.(Glu5425Lys)	13,15, 23	P
KB90	NA	ex 51	c.16295G>A	p.(Arg5432Gln)	8, 21, 23	LP
KB467	NA	ex 51	c.16295G>A	p.(Arg5432Gln)	8, 21, 23	LP
KB480	NA	ex 52	c.16385A>G	p.(Asp5462Gly)	This study	VOUS
KB177	de novo	ex 52	c.16412G>T	p.(Arg5471Met)	13	LP
KB489	NA	ex 53	c.16498C>T	p.(Arg5500Trp)	15	P
KB120	de novo	ex 54	c.16528T>G	p.(Tyr5510Asp)	13	LP
Indel						
KB404	inherited M	ex 10	c.2283_2309del	p.(Ala765_Gln773del)	This study	LB
KB274	NA	ex 10	c.2532_2591del	p.(Arg845_Pro864del)	13	VOUS
KB370*	de novo	ex 14	c.4202_4210del	p.(Ser1401_Cys1403del)	This study	LP
KB384	NA	ex 39	c.11220_11222dup	p.(Gln3745dup)	This study	LB
KB461*#	NA	ex 39	c.11220_11222dup	p.(Gln3745dup)	This study	LB
KB281	inherited M	ex 39	c.11714_11716dup	p.(Gln3905dup)	13	LB
KB71	inherited M	ex 39	c.11819_11836dup	p.(Leu3940_Gln3945dup)	4	LB
KB227	inherited P	ex 39	c.11843_11860del	p.(L3948_Q3953del)	13	LB
KB228	inherited P	ex 39	c.11854_11874dup	p.(Q3952_Q3958dup)	13	LB
KB77	NA	ex 48	c.15163_15168dup	p.(Asp5055_Leu5056dup)	4, 11	VOUS
KB403	de novo	ex 53	c.16489_16491del	p.(Ile5497del)	4, 5, 7, 23	LP
KB53	NA	ex 53	c.16489_16491del	p.(Ile5497del)	4, 5, 7, 23	LP
Splice site						
KB286	de novo	int 2-3	c.177-2A>C	p.S59Rfs*86	13	P
KB31	NA	int 3-4	c.400+1 G>A	p.Ser59Argfs*86	4	P
KB20	de novo	int 3-5	c.401-3 A>G	p.Gly134Glufs*75	4	P
KB442	NA	int 6-7	c.840-6delC	p.?	This study	VOUS
KB 519	NA	int 13-14	c.4132-2A>G	p.?	This study	P
KB 529	NA	int 16-17	c.4584-6C>G	p.?	This study	VOUS
KB210	de novo	int 17-18	c.4693+1G>A	p.Val1561Argfssplice*11	10	P
KB29	NA	int 42-43	c.13999+5 G>A	p.Asn4614Ilefs*5	2, 4, 22	P
KB290	de novo	int 47-48	c.14643+1G>A	p.Glu4882Profs*36	13	P
KB195	de novo	int 47-48	c.14644-3C>G	p.Gln4882Profs*36	13	P
KB7	de novo	int 44-45	c.14252-6_14252-5ins	p.Val4751Glufs*22	9	P
KB360	NA	int 47-48	c.14644-2A>T	p.?	2, 22	P
KB496	NA	int 53-54	c.16520_16521+1del	p.?	This study	P
Gross deletion						
KB43 [Ⓢ]	NA	ex 48-51	c.15785-238_16168delins	p.?	13	P
KDM6A						
Nonsense						
KB215 [§]	de novo	ex 6	c.514C>T	p.(Arg172*)	12, 13, 23	P
KB341	de novo	ex 6	c.514C>T	p.(Arg172*)	12, 23	P
Frameshift						
KB39	NA	ex 16	c.1846_1849del	p.(Thr616tyrfs*8)	13	P
KB141	NA	ex 17	c.2118_2119ins	p.(G707Hfs*13)	This study	P
KB434	NA	ex 17	c.2515_2518del	p.(Asn839Valfs*27)	25	P
KB381	NA	ex 20	c.3044delC	p.(Thr1015Metfs*33)	This study	P

Missense						
KB415	NA	ex 16	c.1843C>T	p.(Leu615Phe)	This study	VOUS
KB272	NA	ex 17	c.2326G>T	p.(Asp776Tyr)	This study	VOUS
KB131	de novo	ex 20	c.2939A>T	p.(Asp980Val)	13	P
KB380	de novo	ex 26	c.3743A>G	p.(Gln1248Arg)	This study	P
Gross deletion						
KB11	NA	ex 1-2		p.?	This study	P
KB50	de novo	ex 5-9		p.?	24	P
Splice site						
KB314	de novo	int 11-12	c.975-1G>A	p.Cys293IlefsX26	This study	P
KB127	de novo	int 22-23	c.3384+3 3384+6del	p.Asn1070 Lys1094del	13	P

Legend:

* detected together with other variant
^ falls in the last base of exon, predicted to disrupt the donor splice site
& identified by MLPA
detected together with pathogenic mutation
f compound heterozygosity
§ detected together with missense in *KMT2D*
P pathogenic
LP likely Pathogenic
LB likely benign
VOUS variant of unknown significance

References:

- Ng et al., 2010 [14]
- Paulussen et al., 2011 [162]
- Li et al., 2011 [163]
- Micale et al., 2011 [4]
- Hannibal et al., 2011 [164]
- Pasqualucci et al., 2012 [165]
- Banka et al., 2012 [166]
- Kokitsu-Nakata et al., 2012 [167]
- Makrythanasis et al., 2013 [159]
- Ratbi et al., 2013 [161]
- Miyake et al., 2013b [168]
- Banka et al., 2014 [169]
- Micale et al., 2014 [160]
- Cheon et al., 2014 [170]
- Ju-Li Lina et al., 2014 [171]
- Dentici ML et al., 2014 [172]
- Cappuccio et al., 2014 [157]
- Dentici ML et al., 2014 [172]
- Morgan et al., 2015 [173]
- Van Laarhoven et al., 2015 [37]
- Shuan Liu et al., 2015 [174]
- Schott et al., 2016 [175]
- Bögershausen et al., 2016 [18]
- Lederer et al., 2012 [53]
- Lederer et al., 2014 [176]

***KMT2D* nonsense and frameshift variants**

The large prevalence of *KMT2D* variants found in our cohort is predicted to lead to truncated protein suggesting a loss of function event, and therefore haploinsufficiency as the likely mechanism for the KS phenotype. *KMT2D* mRNA carrying truncating variants may result in the partial transcripts degradation through nonsense-mediated mRNA decay pathway, contributing to protein haploinsufficiency. In order to investigate this possibility, we measured by qPCR *KMT2D* mRNA levels in 4 KS- lymphoblastoid cell lines (KS-LCLs) and in 3-KS fibroblasts cell lines carrying nonsense variants before and after treatment with puromycin, a known indirect NMD inhibitor. After puromycin treatment, we observed an increased *KMT2D* transcript levels of 2.18, 2.35 e 4.01 folds in fibroblast cell lines of KB186, KB 153 e KB 3, respectively (Figure 10A), 1, 93 and 1.69 in KB 48 and KB 83 lymphoblastoid cell lines compared to untreated cells (Figure 10C). The reduced amount of *KMT2D* mRNA levels in mutated cell lines and its recovery following puromycin treatment indicate that endogenous transcripts might be subject to NMD although with variable efficiency. The physiological NMD substrate SC35 1.7 Kb was included as positive control to

ensure the functional activity of the NMD machinery (Figure 10 B, D) [177]. These data have also been confirmed by direct sequencing of rtPCR-mRNA derived from KB 186, KB 153 and KB 3 patients' fibroblasts (Figure 10E).

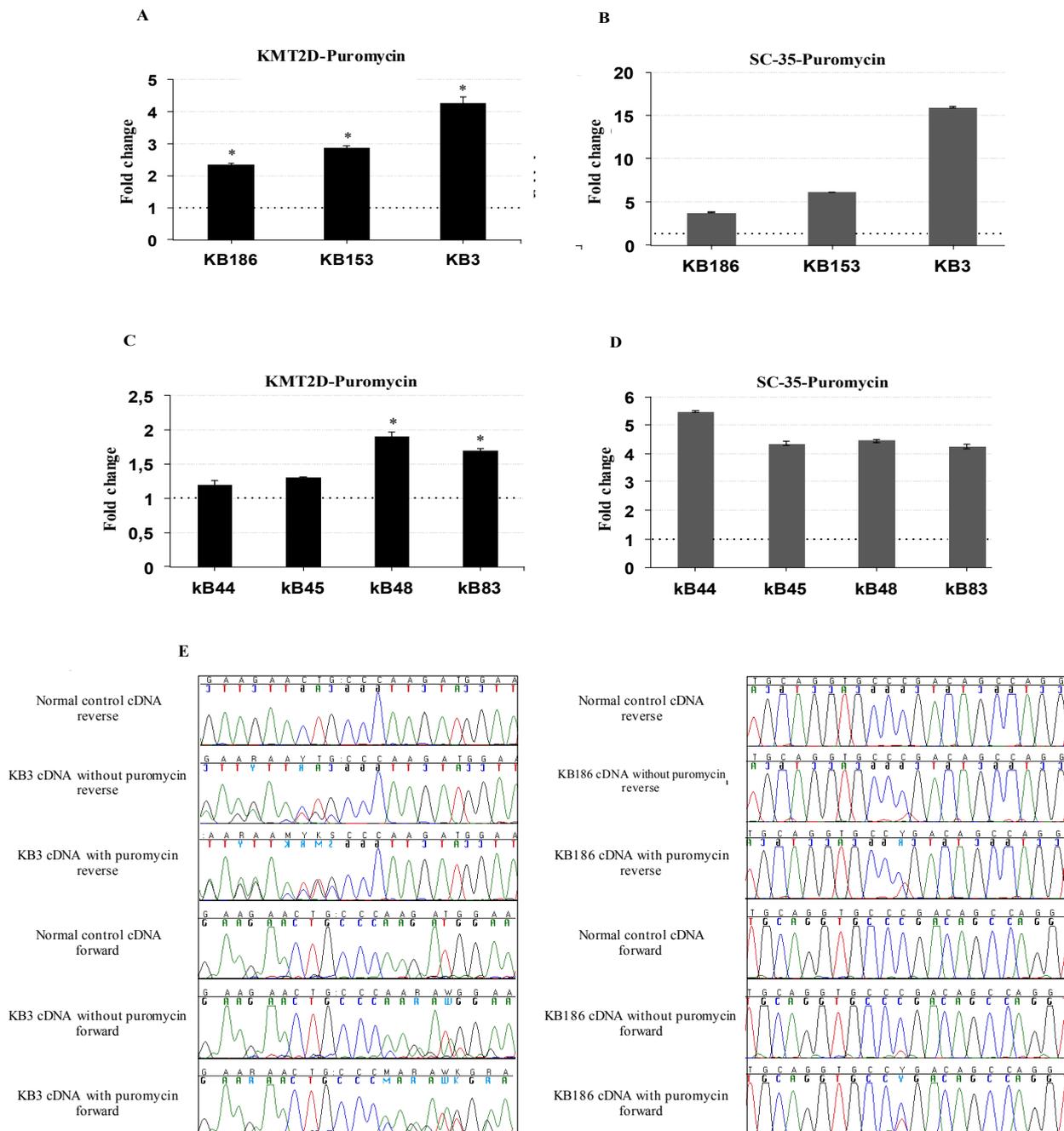


Figure 10. Nonsense-mediated mRNA decay results A-D) *KMT2D* mRNA levels of fibroblasts and lymphoblastoid cell lines (KS-LCLs) after treatment with puromycin (200ug/ml for 8hs) measured by qPCR. (E) The inhibition of the NMD process restores the mutant alleles in presence of puromycin (c.9961C>T, KB186, c.7903C>T, KB153; and c.5652dup, KB3) Micale et al.; 2014 [160].

Identification of *KMT2D* nonsense variants responsive to gentamicin treatment

Fourteen *KMT2D* nonsense variants were tested for gentamicin readthrough treatment through an in vitro dual reporter luciferase vector system. Readthrough efficiency has been shown to depend on the nature of the sequences surrounding the stop codon. We therefore tested those nonsense mutations where the nucleotide context is in line with reported previous studies [155, 178]. Sequences spanning 27 nucleotides of the different stop mutations were inserted into the dual reporter vector pCRFL in-frame between Renilla and Firefly Luciferase coding sequences. Readthrough levels were quantified in HEK293 cells transiently transfected with the dual reporter vector, in the presence or absence of incremental doses of gentamicin. Preliminary experiments with not-transfected cells demonstrated that no signs of cytotoxicity were detectable after 48 h of treatment with up to 1200 µg/ml gentamicin. The difference between basal and gentamicin-induced readthrough level was statistically significant ($P < 0.01$) for eleven *KMT2D* nonsense mutations (Figure 11 and Table 4), although with a wide variability in strength. Nonsense mutations with a readthrough rates $> 0,09\%$ were considered as responsive to gentamicin treatment. In agreement with the previous report [179], we found that all the eleven mutations responsive to gentamicin have a uracil in -1 residue immediately upstream the stop codon and/or a cytosine in $+4$ position confirming that the presence of these residues promotes higher basal and gentamicin-induced readthrough than other nucleotides. We observed that the readthrough levels were increased when higher concentrations of gentamicin used, suggesting that their treatment response is dose dependent.

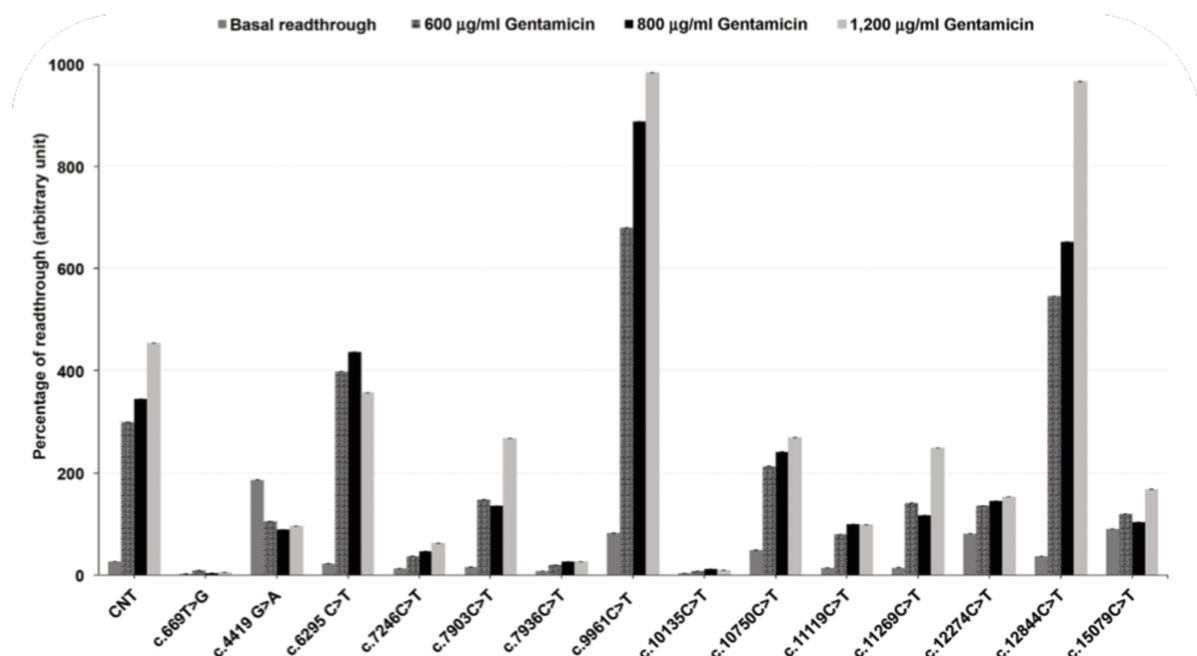


Figure 11. Comparison between basal and gentamicin-induced readthrough level.

The distribution of readthrough level driven by different stop codons in the absence of gentamicin or in presence respectively of 600, 800, 1200 µl/mg. Each value corresponds to the mean of four to six independent experiments. * P-value<0,01.

Table 4. Basal and gentamicin-induced readthrough for KB mutations

Gene	WT codon	Mutation	Nt change	Nucleotidic context of stop mutation	Readthrough at gentamicin (µg/ml)=%				Increase factor between doses 0 and 600	Increase factor between doses 0 and 800	Increase factor between doses 0 and 1200
					0	600	800	1200			
KMT2D	TAT	p.Tyr223X	c.669T>G	TGGAGGGGGCTGCATAGCTGGAGGAGGCTCCA	0,002	0,009*	0,004	0,004	4	2	2
KMT2D	TGG	p.Trip1473X	c.4419G>A	TGGTGGAAAGTGCAGTGA1GTGTGTCTGT CCA	0,187	0,104	0,089	0,096	1	0	1
KMT2D	CGA	p.Arg2099X	c.6295 C>T	TGGCGTAAGACTGACTGACCGGCCATCATCCA	0,022	0,398*	0,436*	0,356*	18	19	16
KMT2D	CAG	p.Gln2416X	c.7246C>T	TGGCCTGCCAGTCTAGTCCAGCTCCAGCCCA	0,013	0,037*	0,047*	0,062*	3	4	5
KMT2D	CGA	p.Arg2635X	c.7903C>T	TGGGGAGCCCTCACAGTGA1CAGGCATCACC CCA	0,015	0,147*	0,135*	0,267*	10	9	18
KMT2D	GAA	p.Glu2646X	c.7936C>T	TGGGTGCGAAAAGCGATAAGACCCAGGACTCCA	0,008	0,019*	0,026*	0,026*	3	3	3
KMT2D	CGA	p.Arg3321X	c.9961 C>T	TGGCTTGCAGGTGCC1TGAACAGCCAGGTTTGCCA	0,082	0,681*	0,867*	0,965*	8	11	12
KMT2D	CAG	p.Gln3379X	c.10136C>T	TGGCCAGTGCTATCATAGAAGCCATGGGCCCA	0,003	0,008*	0,011*	0,013*	2	3	3
KMT2D	CAG	p.Gln3584X	c.10750C>T	TGGCAGTCCGGAAATAGCAGAAAGGAGCACCCA	0,049	0,213*	0,241*	0,269*	4	5	6
KMT2D	CGA	p.Arg3707X	c.11119C>T	TGGAACCTTGCTCTTTGAAGCCTCGGACCTCCA	0,014	0,079*	0,099*	0,098*	6	7	7
KMT2D	CAG	p.Gln3757X	c.11269C>T	TGGATCCAGCAGCAATAGCAGCAGGCTCCTCCA	0,015	0,141*	0,117*	0,249*	10	8	17
KMT2D	CAG	p.Gln4092X	c.12274C>T	TGGTCTCTGCAGCTGTAGCCACCCTCTGAGGCCA	0,080	0,135	0,145	0,152	2	2	2
KMT2D	CGA	p.Arg4282X	c.12844C>T	TGGGGGCGAGGGCCCTTGACCTCAGGGCCACCCA	0,036	0,545*	0,652*	0,967*	15	18	27
KMT2D	CGA	p.Arg5027X	c.15079C>T	TGGGACAAGGTACCGTGAACATGCGTGCGCCA	0,090	0,119	0,103	0,168	1	1	2
	TQ: in frame Ctrl			GCAGGAACACAACAGCAATTACAG	18,122	14,865	15,635	12,372	1	1	1
	CGA*	p.Ser319X	c.966C>G	AGCCCATTTCTTGACAGCATTGGAA	0,027	0,299*	0,344*	0,454*	11	13	17

The readthrough response was defined as the factor of increase between basal and gentamicin-induced readthrough levels.

TQ: construct with no stop codon (100% of activity)

*: p-value <0.01.

***KMT2D* splice site variants**

Six splice site variants identified in our study have been analyzed by NetGene2 and Fruitfly softwares to estimate the possible translational consequences of the nucleotide change.

The RNA availability allowed us to perform RT-PCR to experimentally determine the alteration effects of the regulatory process at the transcriptional or splicing level. RT-PCR confirmed the bioinformatics prediction (Figure 12), in particular:

A) c.400+1G>A (p.Ser59ArgfsX86): this variant occurring within the GT donor splice site in intron 3 results in the disruption of the canonical splice site. RT-PCR of the carrier patient with primers that encompass exons 2 and 4 revealed two bands, one of the expected size and a second smaller band. Sequencing of the smaller fragment showed a skip of the entire exon 3 (224bp) (Figure 12A) that resulted in a frameshift with a generation of a premature stop codon.

B) c.400-3A>G (p.Gly134GlufsX74): this variant creates a novel splice acceptor site in the intron 3. Sequencing of the normal-sized fragment amplified from cDNA's patient revealed a transcript with an insertion of 2bp that results in a frameshift and a predicted premature stop codon (Figure 12B).

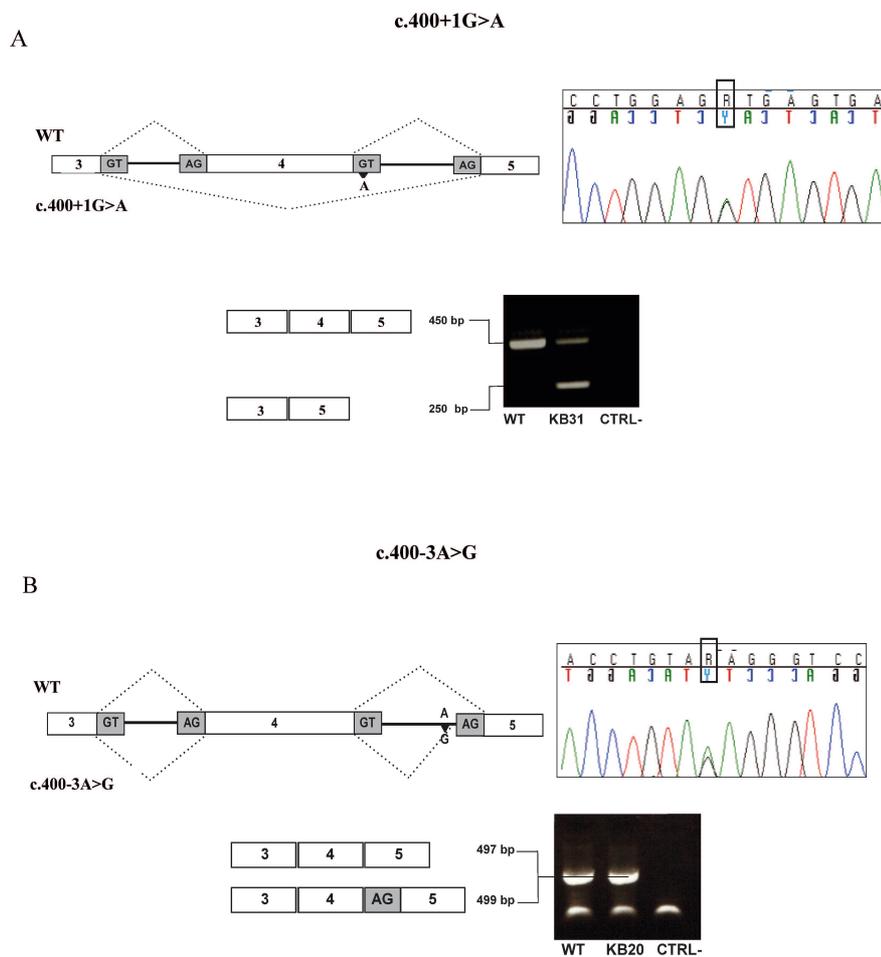
C) c.4693+1G>A (p.Val1561ArgfsX11): this variant affects the essential nucleotide +1 of the donor splice site of *KMT2D* exon 17. Sequencing of the expected size band revealed a transcript with a 13 nucleotide deletion of exon 17 generated by the use of an alternative cryptic donor within exon 17 (Figure 12C). Yet this variant results in a premature stop codon.

D) c.13999+5G>A (p.Asn4614IlefsX5): a shorter PCR product compatible with exon 42 skipping was observed at the cDNA level in KB29 patient using primers that encompass exons 41 and 43 (Figure 12D).

E) c.14252-6_14252-5insGAAA (p.Val4751_Lys4794fsX22): in the patient carrying the mutated allele we identified a small insertion (GAAA) within intron 44 that leads to skipping of exon 45 (Figure 12E).

F) c.14644-3C>G (p.Glu4882ProfsX36): this variant creates a novel acceptor splice site within exon 48. The PCR products from cDNA patient showed a normal and an additional prominent lower band (550 bp and 450 bp, respectively). Sequencing of the 450 bp band revealed a deletion of 232 bp of the exon 48 (Figure 12F).

In summary the six intronic *KMT2D* variants cause aberrant splicing of the corresponding transcript that result in a truncated and functionally altered protein.



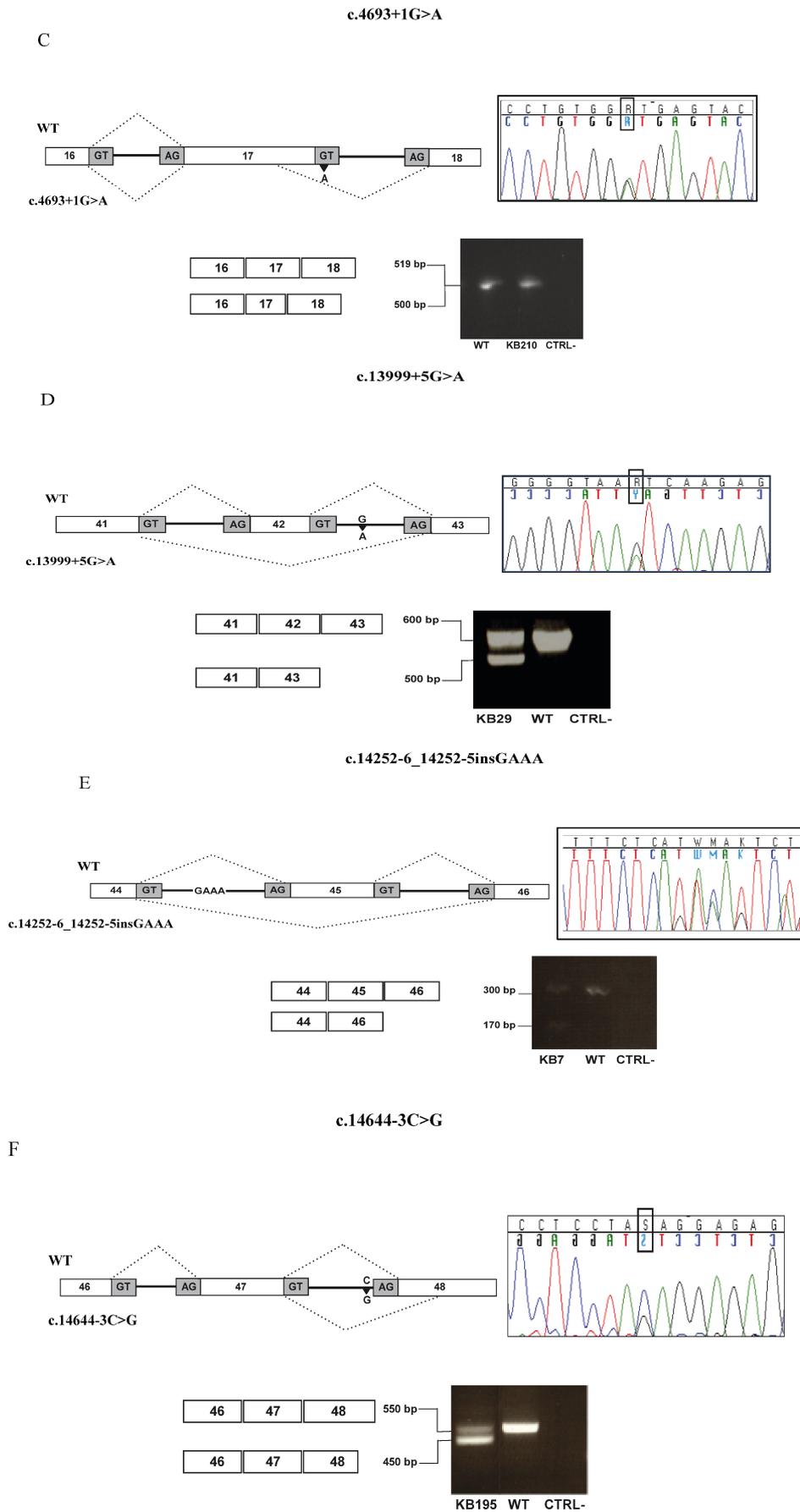


Figure 12. Representation of six *KMT2D* splice site variants.

On the left, schematic representation of six *KMT2D* splice site variants. On the right, electropherograms showing the heterozygous mutations. On the bottom, RT-PCR analysis on RNAs from Kabuki patients and control (wt: wild type).

***KMT2D* mosaic variants**

In patient GM13-3816, using a target resequencing panel of Kabuki genes, we identified a c.15061C=>T, p.R5021X heterozygous variant in *KTM2D* gene, already reported as causative [180]. The variant was detected in a small fraction of alleles from proband's DNA extracted peripheral blood. The variant has been confirmed and quantified by Pyrosequencing analysis, that quantitatively estimated a mosaicism level of 16% in the DNA extracted from proband's peripheral blood representing a mosaicism of ~32% in leukocytes (Figure 13). In patient KB450 and KB369 by Sanger Sequencing we identified the already reported nonsense variant c.13450C=>T (p.R4484X) [162] and the never reported before frameshift variant c.3596_3597=/del (p.L1199HfsX7), both in mosaic, respectively. Pyrosequencing was used to confirm and quantitatively estimate the levels of mosaicism in different tissues, when available. The assay was validated by measuring the levels of mutant allele in the heterozygous proposita DNA, where the *KMT2D* c.13450C=>T (p.R4484X) mutation was estimated to be at a level of ~34% in peripheral blood alleles (mosaicism of ~68% in leucocytes) and ~46% in saliva (~92% in epithelial cells). The c.3596_3597=/del (p.L1199HfsX7) mutation was estimated to be at a level of 20% in the peripheral blood alleles (mosaicism of ~40% in leucocytes) and ~27% in saliva (~54% in epithelial cells) (Figure 13).

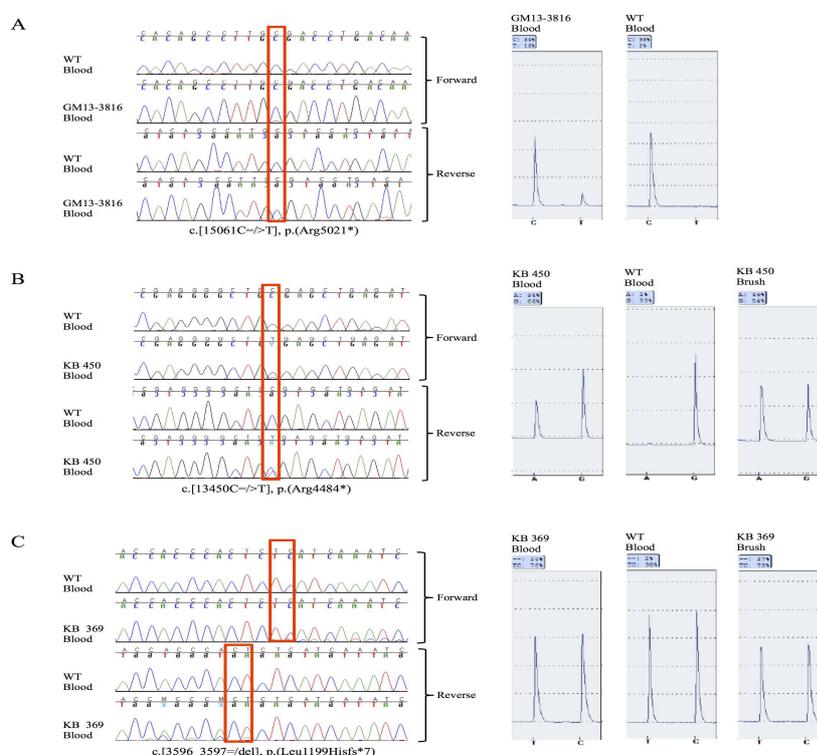


Figure 13. Sanger sequencing data and pyrosequencing results

On the left, blood Sanger electropherogram for patients and controls. On the right, blood pyrosequencing peak profile for patient 1 and control; blood and buccal brush for patient 2 and patient 3.

Missense variants pathogenic assessment by bioinformatic approaches

Most *KMT2D* variants in KS are inactivating mutations that are functionally defective due to the loss of the catalytic SET activity. However, ~30% of *KMT2D* variants in our data set are missense that affect various residues along the *KMT2D* protein (69 of 208 mutational events, 33%; and 100 of 621 (16.1%) events from recent published mutational analysis review [18].

Assigning pathogenicity of missense variants is challenging in genetic diagnosis and counselling. For the 58 new different missense variants, found in 69 KS patients in our cohort we adopted an *in silico* approach that combines already published software with comparative analysis and motif/domain search tools with threading/homology modelling protocols to predict functional and structural effect of missense *KMT2D* variants (Table 5).

Table 5 . Bioinformatics analysis of *KMT2D* missense variants

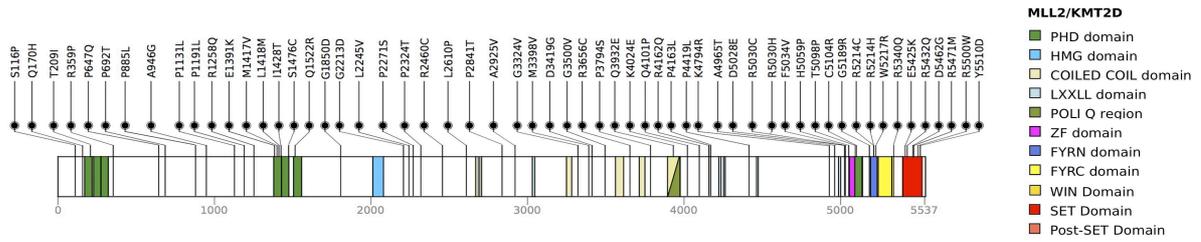
CODE	Variant	AA change	ACMG Variant Class	PREDICTION SOFTWARES											
				Polyphen		Align GVG D		Provean		SIFT		UMD predictor		Mutation Taster	
				Score	Result	Score	Result	Score	Result	Score	Result	Score	Result	Score	Result
KB21#	c.346T>C	p.(Ser116Pro)	LB	PD	0.997	73.35	Class C65	-3.35	D	0.001	D	90	P	DC	0.994
KB21#^	c.510G>C	p.Gln170His	P	PD	1.00	24.08	Class C15	-2.32	N	0.016	D	100	P	DC	0.998
KB256	c.626C>T	p.(Thr209Ile)	VOUS	B	0.096	89.28	Class C65	-2.23	N	0.009	D	75	P	PM	0.980
KB458	c.1076G>C	p.(Arg359Pro)	VOUS	B	0.011	102.7	Class C65	0.04	N	0.2056	T	33	PM	PM	0.999
KB269	c.1940C>A	p.(Pro647Gln)	LB	PD	0.59375	75.14	Class C65	-0.90	N	0.0771	T	78	P	PM	0.999
KB126, KB374#	c.2074C>A	p.(Pro692Thr)	LB	B	0.008	37.56	Class C35	-0.92	N	0.011	D	66	PP	PM	0.999
KB487	c.2654C>T	p.(Pro885Leu)	VOUS	B	0.000	97.78	Class C65	-1.53	N	0.003	D	60	PPM	PM	0.999
KB370#	c.2837C>G	p.(Ala946Gly)	LB	B	0.000	60.00	Class C55	-0.78	N	0.032	D	30	PM	PM	0.999
KB215, KB341	c.3392C>T	p.(Pro1131Leu)	LB	B	0.196	97.78	Class C65	-0.30	N	0.001	D	39	PM	PM	0.989
KB222	c.3572C>T	p.(Pro1191Leu)	LB	PD	0.764	97.78	Class C65	-2.50	D	0.001	D	35	PM	DC	0.999
KB32	c.3773G>A	p.(Arg1258Gln)	LB	PD	0.997	42.81	Class C35	-1.37	N	0.002	D	75	P	PM	0.691
KB307	c.4171G>A	p.(Glu1391Lys)	LP	PD	1.000	56.87	Class C55	-3.29	D	0.001	D	78	P	DC	0.999
KB28#	c.4249A>G	p.(Met1417Val)	LB	PD	0.476	20.52	Class C15	00:19	N	0.4097	T	60	PPM	DC	0.559
KB28#	c.4252C>A	p.(Leu1418Met)	LB	PD	1.000	14.30	Class C0	-1.71	N	0.004	D	69	PP	DC	0.999
KB174	c.4283T>C	p.(Ile1428Thr)	LB	PD	0.889	89.28	Class C65	01:14	N	0.5847	T	93	P	DC	0.996
KB138	c.4427C>G	p.(Ser1476Cys)	VOUS	B	0.005	111.6	Class C65	0.05	N	0.1771	T	99	P	PM	0.885
KB34	c.4565A>G	p.(Gln1522Arg)	LB	PD	0.997	42.81	Class C35	-3.43	D	0.021	D	100	P	DC	0.999
KB 535#	c.5549G>A	p.(Gly1850Asp)	LB	PD	0.535	93.77	Class C65	-0.72	N	0.103	T	87	P	PM	0.997
KB119, KB204#	c.6638G>A	p.(Gly2213Asp)	LB	PD	0.454	93.77	Class C65	-0.70	N	0.004	D	42	PM	DC	0.917
KB330#	c.6733C>G	p.(Leu2245Val)	VOUS	PD	0.69	30.92	Class C25	-0.46	N	0.046	D	69	PP	PM	0.996
KB326#	c.6811C>T	p.(Pro2271Ser)	LB	B	0.024	73.35	Class C65	-2.00	N	0.004	D	75	P	DC	0.850
KB107#	c.6970C>A	p.(Pro2324Thr)	VOUS	PD	0.985	37.56	Class C35	-2.47	N	0.003	D	81	P	PM	0.998
KB430	c.7378C>T	p.(Arg2460Cys)	LB	PD	1.000	179.5	Class C65	-2.22	N	0.023	D	96	P	DC	0.999
KB122, KB287	c.7829T>C	p.(Leu2610Pro)	LB	B	0.076	97.78	Class C65	-1.20	N	0.0819	T	72	PP	DC	0.998
KB27	c.8521C>A	p.(Pro2841Thr)	VOUS	B	0.009	37.56	Class C35	-1.45	N	0.010	D	69	PP	DC	0.997
KB330#	c.8774C>T	p.(Arg2925Val)	VOUS	B	0.004	65.28	Class C65	-1.07	N	0.004	D	84	P	PM	0.958
KB443#	c.9971G>T	p.(Gly3324Val)	LB	PD	0.988	109.5	Class C65	-2.24	N	0.011	D	78	P	DC	0.999
KB326#	c.10192A>G	p.(Met3398Val)	LB	B	0.000	20.52	Class C15	-2.84	D	0.002	D	41	PM	PM	0.958
KB357															
KB297, KB378	c.10256A>G	p.(Asp3419Gly)	LB	PD	1.000	93.77	Class C65	-1.81	N	0.000	D	100	P	DC	0.999
KB292	c.10499G>T	p.(Gly3500Val)	LP	PD	0.6875	109.5	Class C65	-8.18	D	0.000	D	100	P	DC	0.999
KB86#	c.10966C>T	p.(Arg3656Cys)	VOUS	PD	1.000	179.5	Class C65	-1.87	N	0.001	D	84	P	DC	0.999
KB374#	c.11380C>T	p.(Pro3794Ser)	LB	B	0.016	73.35	Class C65	-1.04	N	0.029	D	66	PP	PM	0.993
KB293	c.11794C>G	p.(Gln3932Glu)	LB	B	0.002	29.27	Class C25	-0.68	N	0.047	D	45	PM	PM	0.999

KB204#	c.12070A> G	p.(Lys4024Glu)	LB	B	0,296	56.87	Class C55	00:08	N	0.000	D	63	PPM	PM	0,997
KB385	c.12302A> C	p.(Gln4101Pro)	VOUS	B	0.000	75.14	Class C65	-0.03	N	0,2014	T	66	PP	PM	0,939
KB247	c.12485G> A	p.(Arg4162Gln)	VOUS	B	0.003	42.81	Class C35	00:30	N	0.000	D	72	PP	PM	0,995
KB107#	c.12488C> T	p.(Pro4163Leu)	VOUS	PD	0,991	97.78	Class C65	-2.29	N	0.000	D	72	PP	DC	0,983
KB170	c.13256C> T	p.(Pro4419Leu)	LB	PD	1.000	97.78	Class C65	-4.16	D	0.000	D	87	P	DC	0,999
KB 512	c.14381A> G	p.(Lys4794Arg)	VOUS	PD	0,999	26.00	Class C25	-1.43	N	0.005	D	69	PP	DC	0,999
KB86#	c.14893G> A	p.(Ala4965Thr)	VOUS	PD	0,941	58.02	Class C55	-1.17	N	0,0847	T	75	P	DC	0,997
KB38#	c.15084C> G	p.(Asp5028Glu)	LP	PD	0,999	44.60	Class C35	-3.83	D	0.001	D	66	PP	DC	0,999
KB154, KB185	c.15088C> T	p.(Arg5030Cys)	VOUS	PD	1.000	179.5	Class C65	-7.56	D	0.000	D	99	P	DC	0,999
KB423	c.15089G> A	p.(Arg5030His)	LB	PD	1.000	28.82	Class C25	-4.79	D	0.000	D	81	P	DC	0,999
KB38#	c.15100T> G	p.(Phe5034Val)	LP	PD	0,999	48.95	Class C45	-6.17	D	0.001	D	84	P	DC	0,999
KB76	c.15176A> C	p.(His5059Pro)	LP	PD	1.000	76.28	Class C65	-9.60	D	0.001	D	93	P	DC	0,999
KB129	c.15292A> C	p.(Thr5098Pro)	VOUS	PD	0,992	37.56	Class C35	-4.21	D	0,1035	T	93	P	DC	0,999
KB462	c.15310T> C	p.(Cys5104Arg)	LP	PD	1.000	179.5	Class C65	-	D	0.000	D	96	P	DC	0,999
KB171	c.15565G> A	p.(Gly5189Arg)	VOUS	PD	1.000	125.1	Class C65	11.51 -7.68	D	0.000	D	100	P	DC	0,999
KB264, KB376	c.15640C> T	p.(Arg5214Cys)	LP	PD	1.000	179.5	Class C65	-7.68	D	0.000	D	96	P	DC	0,999
KB24, KB219, KB408	c.15641G> A	p.(Arg5214His)	LP	PD	1.000	28.82	Class C25	-4.80	D	0.000	D	78	P	DC	0,999
KB109	c.15649T> C	p.(Trp5217Arg)	LP	PD	1.000	101.2	Class C65	-13.4	D	0.000	D	96	P	DC	0,999
KB17	c.16019G> A	p.(Arg5340Gln)	LP	PD	1.000	42.81	Class C35	-3.84	D	0.000	D	81	P	DC	0,999
KB169	c.16273G> A	p.(Glu5425Lys)	P	PD	1.000	56.87	Class C55	-3.70	D	0.000	D	75	P	DC	0,999
KB90, KB467	c.16295G> A	p.(Arg5432Gln)	LP	PD	1.000	42.81	Class C35	-3.70	D	0.000	D	84	P	DC	0,999
KB480	c.16385A> G	p.(Asp5462Gly)	VOUS	PD	1.000	93.77	Class C65	-6.69	D	0.000	D	90	P	DC	0,999
KB177	c.16412G> T	p.(Arg5471Met)	LP	PD	1.000	91.64	Class C65	-5.72	D	0.000	D	100	P	DC	0,999
KB489	c.16498C> T	p.(Arg5500Trp)	P	PD	1.000	101.2	Class C65	-7.44	D	0.000	D	99	P	DC	0,999
KB120	c.16528T> G	p.(Tyr5510Asp)	LP	PD	1.000	159.9	Class C65	-9.52	D	0.000	D	96	P	DC	0,999

detected together with other variant
P pathogenic
LP likely Pathogenic
LB likely benign
VOUS variant of unknown significance
PD probably damaging
B benign
D deleterious/damaging
N neutral
T tolerated
PP probably pathogenic
PPM probable polymorphism
PM polymorphism
DC disease causing

The missense variants we found are distributed across the entire length of the *KMT2D* gene (Figure 14A) and according to the AMCG, 20/69 are classified as pathogenic or likely pathogenic, 30/69 likely benign and 19/69 VOUS (Table 5). Frequency was found as pathogenic in dbSNP (<http://www.ncbi.nlm.nih.gov/projects/SNP/>, 1000 Genomes Project (<http://www.internationalgenome.org/>), EVS (<http://evs.gs.washington.edu/EVS/>) and ExAC (<http://exac.broadinstitute.org/>). Combined comparative and conservative analysis also revealed that *KMT2D* missense variants alter evolutionarily conserved aminoacid positions. Free energy variations and conformational changes of the protein structure are also observed for the analyzed variants.

A



B

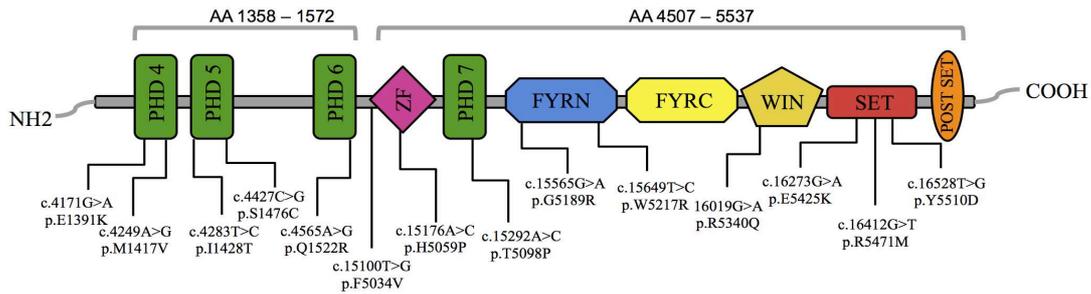


Figure 14. Missense variants distribution across the entire length or fusion *KMT2D* gene.

A) Schematic representation of the KMT2D protein (PHD, plant homeodomain; HMG, high-mobility group; Coiled Coil domain; LXXLL domain, motifs with the consensus sequence L-X-X-L-L motif; Poli Q region, Poli Q reach region; ZF domain, Zinc Finger Domain; FYRN, FY-rich, N-terminal; FYRC, FY-rich, C-terminal; WIN, WDR5 Interaction domain, SET, Su(var)3-9, Enhancer-of-zeste, Trithorax). Underlined, the Poli-Q region identified in this study. In black, the missense variants found in our cohort of KS patients. B) Functionally analysed missense variants distribution across the Fusion-KMT2D construct composed of PHD4–5-6, (amino acids 1358–1572) and ZF-PHD7-FYRN-FYRC-WIN-SET-post SET domains (amino acids 4507–5537).

Missense variants impair KMT2D methyltransferase activity

Up to now, very few *KMT2D* missense variants have been experimental tested in order to assess their pathogenicity (Table 6).

Table 6. This study/published functional tests performed on *KMT2D* missense variants

KMT2D mutation	Domain	H3K4 (me1)	H3K4 (me2)	H3K4 (me3)	Interaction with WRAD	Ref	
R228G	ZF, PHD	NT	NT	NT	NT	2	
S1065T	none	Unch	Unch	Unch	NT	2	
P1104T	none	Unch	Unch	Unch	NT	2	
D1121H	none	Unch	Unch	Unch	NT	2	
C1430R	PHD5	Altered	Altered	Altered	Unch	1	
C1471Y	PHD5	Altered	Altered	Altered	Unch	1	
T1506S	PHD	Unch	Unch	Unch	NT	2	
P1508A	PHD6	NT	NT	NT	NT	1	
Y1514A	PHD6	NT	NT	NT	NT	1	
E1516A; E1517A; D1518A; E1544A	PHD6	Altered	Altered	Altered	Unch	1	
C1523A	PHD6	Altered	Altered	Altered	Unch	1	
W1529A	PHD6	NT	NT	NT	NT	1	
W1529A; E1544A	PHD6	NT	NT	NT	Unch	1	
R2095S	none	Unch	Unch	Unch	NT	2	
K2468T	none	Unch	Unch	Unch	NT	2	
R4659P	none	Unch	Unch	Unch	NT	2	
V4799M	none	Unch	Unch	Unch	NT	2	
R5027L	none	Altered	Altered	Altered	NT	2	
L5056P	none	Altered	Altered	Altered	NT	2	
C5092S	ZF, PHD	Altered	Altered	Altered	NT	2	
C5092Y	ZF, PHD	Altered	Altered	Altered	NT	2	
D5257V	FYRC	Altered	Altered	Altered	NT	2	
R5303H	FYRC	Unch	Unch	Unch	NT	2	
S5404F	SET	NT	Altered	NT	NT	2	
R5432W	SET	Altered	Altered	Altered	NT	2	
N5437S	SET	Altered	Altered	Altered	NT	2	
G5467D	SET	Altered	Altered	Altered	NT	2	
KMT2D mutation	Domain	H3K4 (me1)	H3K4 (me2)	H3K4 (me3)	Interaction with WRAD	Interaction with KDM6A	Ref
E1391K	PHD4	Unch	Unch	Unch	Altered		This study
M1417V	PHD4	Altered	Unch	Altered	Altered		This study
I1428T	PHD5	Unch	Unch	Unch	Altered		This study
S1476C	PHD5	Altered	Unch	Unch	Altered		This study
Q1522R	PHD6	Unch	Unch	Altered	Altered		This study
F5034V	none	Unch	Unch	Unch	Altered		This study
H5059P	ZF	Unch	Unch	Unch	Altered		This study
T5098P	PHD7	Unch	Unch	Unch	Altered		This study
G5189R	FYRN	Altered	Altered	Altered	Altered		This study
W5217R	FYRN	Altered	Altered	Altered	Altered		This study
R5340Q	WIN	Altered	Altered	Unch	Altered		This study
E5425K	SET	Altered	Altered	Altered	Altered		This study
R5471M	SET	Altered	Altered	Altered	Altered		This study
Y5510D	SET	Altered	Altered	Unch	Altered		This study

Legend:

NT= Not tested

Unch= Unchanged

References

1 Shilpa S. Dhar, et al., 2012

2 Zhang J. et al., 2015

To test the functional impact of 14 *KMT2D* missense variants found in our cohort, we generated FLAG-tagged versions of 14 *KMT2D* missense mutant alleles (Figure 14B) using FUSION-KMT2D as template (described in the Methods section). All tested alleles produced similar amounts of both mRNA and protein when transiently transfected in HEK 293T, indicating that missense variants do not interfere with the stability and level of the KMT2D protein (data not shown). As the protein encoded by the FUSION-KMT2D wild type vector retains dose-dependent mono, di and tri methyltransferase activity in vitro at the same extent of full-length protein (Figure 15), we chosen to perform further assays using FUSION-KMT2D that guaranteed higher transfection efficiency. We measured the in vitro ability of the mutated proteins to catalyse mono, di, and trimethylation of H3K4 using purified nucleosomal histones. In this assay, 9/14 missense variants impaired H3K4 monomethylation (H3K4me1), dimethylation (H3K4me2) or trimethylation (H3K4me3) enzymatic activity, (Figure 16A-B), providing experimental evidence that such missense variants are KMT2D loss-of-function variants. To confirm the reduced activity with a different method we used an elsewhere-established epigenetic fluorescent reporter system that monitors H3K4 trimethylation activity [103] (Figure 16C). Fluorescence data reveal a trimethylation H3K4 reduction levels between 40-60% for all the missense variants tested in this assay compared to *KMT2D* wild type counterpart activity confirming the reduced trimethylation activity (Figure 16D).

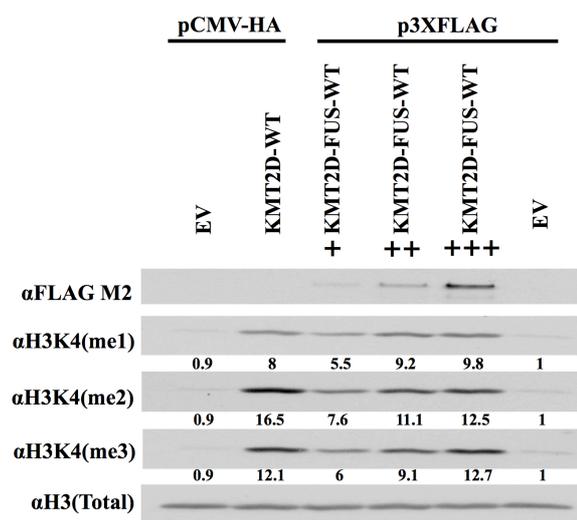


Figure 15. FUSION-KMT2D wild type vector retains dose-dependent mono, di and tri methyltransferase activity Lysine methyltransferase activity of increasing FUSION-KMT2D-WT protein on HeLa nucleosomes at the same extent of full-length protein. Quantification of signal intensity after normalization to total H3 is shown below each lane. For normalization purposes, the ratio of (H3K4 methylation level)/(total H3 levels) in the FLAG-EMPTY-VECTOR sample was set to 1.

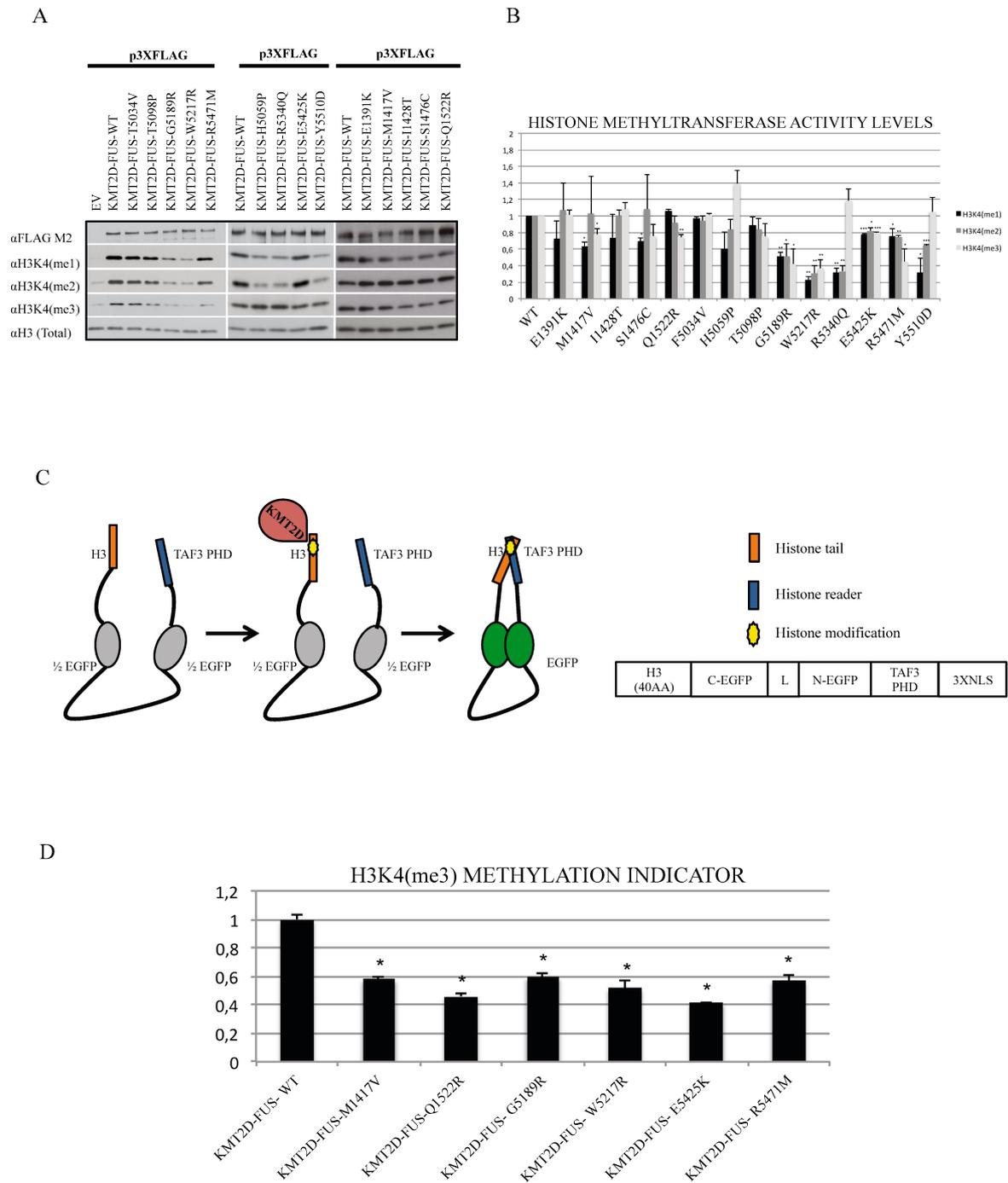


Figure 16. *KMT2D* missense variants are associated with defective methyltransferase activity and diminished H3K4 methylation.

A) Lysine methyltransferase activity of mutated FLAG-*KMT2D* proteins on HeLa nucleosomes. Quantification of signal intensity after normalization to total H3 is shown below each lane. For normalization purposes, the ratio of (H3K4 methylation level)/(total H3 levels) in the wild-type FLAG-*KMT2D* sample was set to 1. B) H3K4 methyltransferase activity average of semipurified FLAG-*KMT2D* proteins on HeLa nucleosomes for two different biological replicates. C) Reporter allele encoding for the two halves of green fluorescent protein (GFP) are separated by a flexible linker region with a histone tail (H3) and a histone reader (TAF3 PHD) at the N and C termini, respectively. When the histone tail is modified by trimethylation, GFP structure and function are reconstituted. D) The H3K4me3 indicator demonstrated a dose-dependent response to the FLAG-*KMT2D* protein, and a decreased H3K4me3 activity for the 6 tested missense variants.

KMT2D-complex protein interaction

ASH2L and RbBP5 engage a direct high affinity interaction with the C-terminal domains of KMT2D (amino acids 4997–5537), co-regulating KMT2D target genes [152, 181]. To define if *KMT2D* missense variants also alter the interaction with ASH2L and RbBP5, we immunoprecipitated protein lysates of HEK 293T cells transfected with Flag-tagged FUSION-KMT2D constructs. Western blot analysis of immunoprecipitation eluates showed that the missense variants located at C-terminal of the protein, p.H5059P, p.T5098P, p.G5189R, p.W5217R, p.R5340Q, p.E5425K, p.R5471M, and Y5510D showed a weak interaction with both ASH2L and RbBP5, while p.E1391K and p.F5034V only distress the interaction between KMT2D and ASH2L (Figure 17A). In addition, the missense variants p.I1428T and p.Q1522R exhibit and increased interaction with both ASH2L and RbBP5, while the p. M1417V and p.S1476C only increase the interaction between KMT2D and ASH2L.

The WDR5 protein interacts with the KMT2D Win motif (AA 5337-5342), a C-terminal interaction motif [182]. Combined conformational 3D modeling and free energy interaction variations revealed that the KS missense variant p.R5340Q alters the protein domain structure, changing an evolutionarily conserved aminoacid position expected to tolerate Arginine and minimally Methionine. To determine if *KMT2D* missense variant p.R5340Q, located within the *KMT2D* Win motif (AA 5337-5342), altered the interaction of KMT2D with WDR5, we co-transfected HEK293 cells with an expression plasmid encoding FLAG-tagged WDR5 with Myc-EGFP-tagged *KMT2D* p.R5340Q variant. Western blot analysis of WDR5 immunoprecipitated showed that the *KMT2D* missense variant p.R5340Q completely abrogates the interaction with WDR5 (Figure 17B).

Overall, according to reduced H3K4 methylation levels of these variants, these data show that the affected protein activity could be a direct consequence of the misplaced interaction with the WRAD complex.

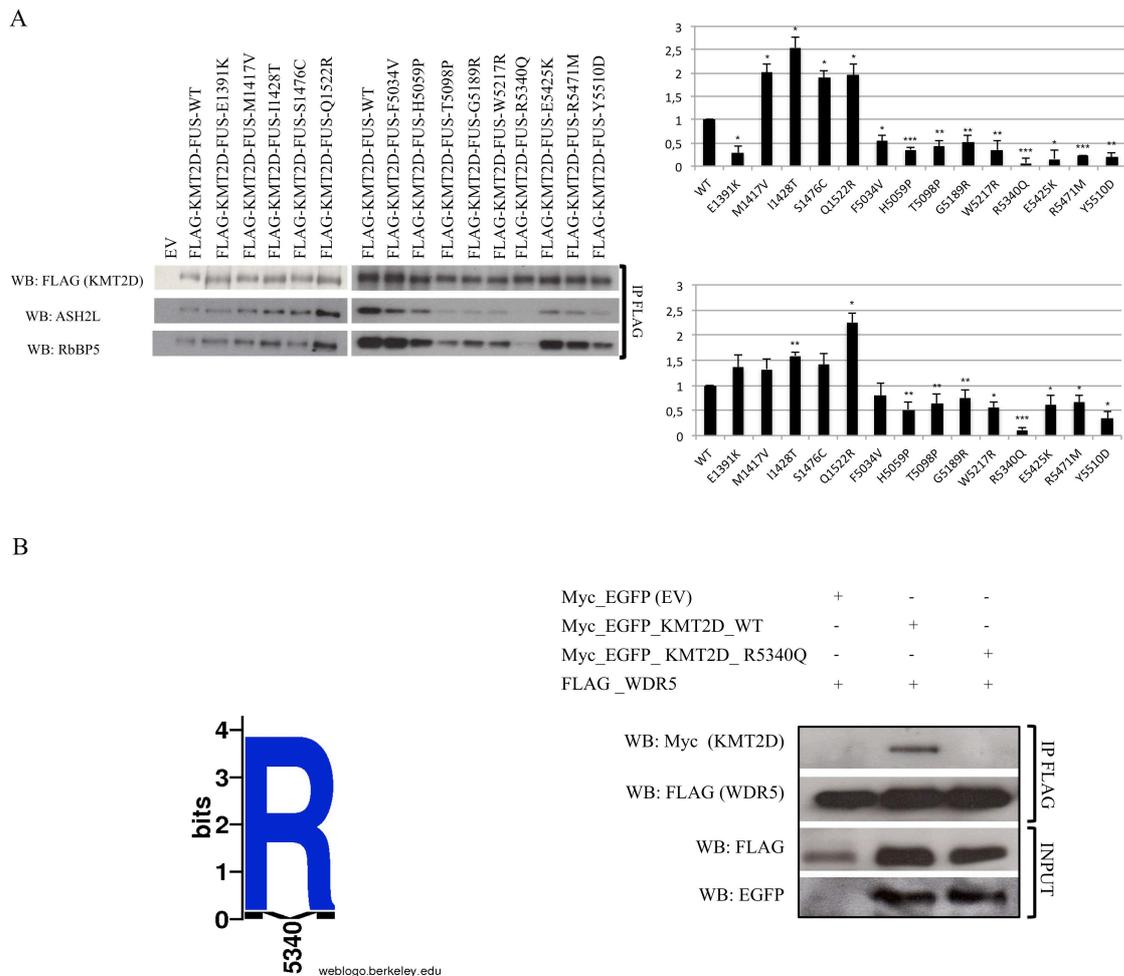


Figure 17. *KMT2D* missense variants interaction with KDM6A, ASH2L, RbBP5 and WDR5

A) On the left, immunoblot analysis of KMT2D, ASH2L and RbBP5 in HEK293T cell line transfected with wild type *KMT2D* or *KMT2D* missense after immunoprecipitation with anti-Flag antibody. On the right, *KMT2D* missense variants interaction average with ASH2L and RbBP5 in two different biological replicates. B) On the left, the predicted aa tolerance for KMT2D p.R5340Q. On the right, immunoblot analysis of KMT2D and WDR5 in HEK293 cell line transfected with wild type *WDR5*, wild-type C-Ter *KMT2D* or the p.R5340Q *KMT2D* missense mutant before (input) and after immunoprecipitation with the anti-Flag antibody

KMT2D mediates neuronal differentiation of NT2D1 cells by activating differentiation-specific genes

A previous study showed that *KMT2D* depletion by shRNAs drastically suppressed RA-induced NT2D1 neuronal differentiation, by regulating the expression of genes involved in the neuronal lineage commitment [152]. Therefore we tested if *KMT2D* missense variants p.R5471M and p.Y5510D, showing a reduced methyltransferase activity, alter the NT2D1 neural commitment. To test this we transfected KMT2D stable-silenced NT2D1 cells with FUSION-KMT2D expression plasmids encoding FLAG-tagged versions of p.R5471M and p.Y5510D missense alleles and evaluated the expression levels of the KMT2D-target key development and differentiation genes HOXA1–3 [152]. qRT-PCR analysis showed altered

levels of HOXA1 and HOXA3 for the missense variant p.Y5510D, while no variations were observed for the p.R5471M (Figure 18).

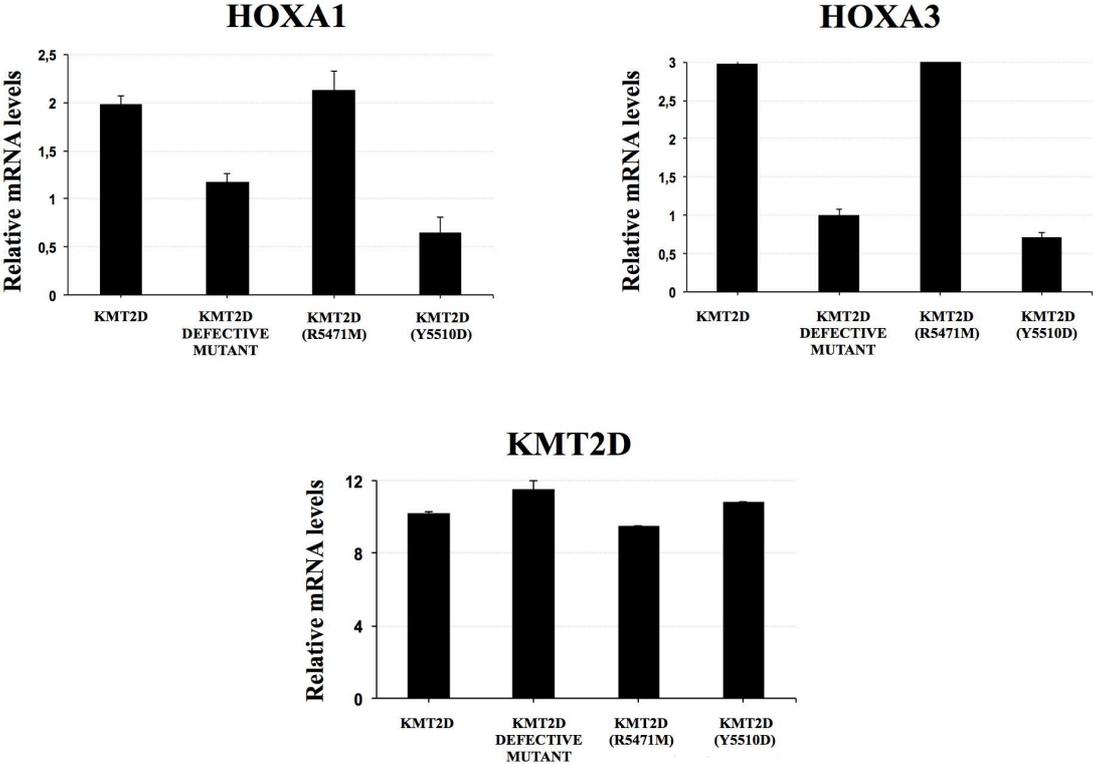


Figure 18. *KMT2D* missense mutations alter transcription levels of its target genes during NT2/D1 cells neuronal differentiation. qPCR analysis of target key development and differentiation genes HOXA1–3 and NANOG in NT2/D1 knockdown cell line transiently transfected with WT and mutant *KMT2D*.

DISCUSSION

In this survey study, we reported 208 *KMT2D* and 14 *KDM6A* variants in our cohort of 505 patients clinically diagnosed as KS-affected. For approximately 30% of individuals with a clinical diagnosis of KS, the genetic cause remains unknown, suggesting that other causative genes or epigenetics mechanisms may be involved, misclassification of variants, or as yet unrecognized promoter or deep intronic mutations affecting normal splicing. Overall, the mutation detection rates for *KMT2D* and *KDM6A* in our KS cohort are lower than that reported in a recent survey study [18]. A possible explanation may be the fact that in our cohort a certain number of patients might have been clinically misdiagnosed and present conditions partially overlapping with KS that should be considered in differential diagnosis with KS. Such a case is represented by Au-Kline syndrome (MIM#616580), a condition partially overlapping or suggestive of Kabuki Syndrome which presents with psychomotor developmental delay, intellectual disability, characteristic face, cardiac, urogenital, skeletal abnormalities, and hypotonia, caused by *HNRNPK* haploinsufficiency [183, 184]. In view of the possible misleading diagnosis, we are now processing our yet molecularly undiagnosed patients' cohort with a NGS panel focused on genes associated to diseases that are in differential diagnosis with KS and that affect the function or the structure of chromatin (unpublished data).

Approximately 30% of human inherited diseases are the consequence of nonsense mutations. Excluding our approach published in 2014, [160] no additional studies addressing stop mutations in *KMT2D* gene as potential target for aminoglycoside-mediated rescue have been published so far. In this study we demonstrate the ability of gentamicin to induce the readthrough of some of naturally occurring stop variants in *KMT2D* gene by using an in vitro assay. Of note, NMD has been reported to be a major mechanism of nonsense transcript elimination and governs the patient response to readthrough. We showed that some of *KMT2D* nonsense transcripts suffer NMD process, thus it is possible that NMD minimize the effect of readthrough treatment for some *KMT2D* nonsense transcripts. To overcome this undesirable contingent issue, the readthrough strategy must be combined with inhibition of NMD by specific inhibitors and/or siRNA directed against NMD key factors as UPF1 or UPF2. Developing a method to specifically disrupt NMD of a disease-causing transcript without influencing normally degraded mRNA might greatly improve the readthrough efficiency of targeted nonsense mutation by readthrough compounds. Our results provide a first preliminary proof-of-concept that naturally occurring nonsense mutations in *KMT2D* can be effectively suppressed by readthrough drugs with different efficiencies. The performed assay only offers evidence for the readthrough response of some *KMT2D* nonsense variants

but does not provide whether the biological function of the protein is restored. Thus, more and specific biochemical and cellular assays implicated in the verification of the readthrough effect are needed. Therefore, the transferability of our data from in vitro to in vivo needs to be further investigate in a mouse model or in the patient's cell lines.

Among our cohort we identified three new cases with mosaic variants in *KMT2D* gene, consisting in single nucleotide change resulting in two already reported nonsense and in a new frameshift variant. To date, *KMT2D* mosaic variants have been described only in three KS female patients [112]. The small number of patients does not allow any comparison between subjects with point mutation mosaic and deletion, but we cannot exclude that mosaic *KTM2D* deletions result in major neurodevelopmental delay, as suggested by the Banka's patients [112].

In our screening we also found 58 different missense variants distributed across the entire length of the *KMT2D* gene including residues in PHD and SET domain and 4 missense variants in KDM6A whose pathogenicity is relevant for diagnostic and counselling purposes. We adopted first a in silico approach that combines comparative analysis and motif/domain search tools with threading/homology modelling protocols to predict functional and structural effect of *KMT2D* missense variants identified in our cohort so far. This analysis predicts that many of them might be pathogenic, although the exact biological role they play in the disease is still speculative. We ascertained the pathogenicity of 14 *KMT2D* missense variants, located within PHD4-5-6 domains and ZF-PHD7- FYRN-FYRC- WIN-SET- domains by dissecting their capability to alter HMT, using two different methods, and to interact with known *KMT2D* partners. HMT assays revealed that 9 *KMT2D* missense variants had altered HMT activity, put forward for consideration that these missense variants are loss of function events and then causative for the observed phenotypes. H3K4 (me1) is enriched on enhancer regions and is required for enhancer activation and super-enhancer formation during cell differentiation of brown adipose tissue and skeletal muscle development [22, 89], whereas H3K4 (me3) is increased on active promoters/transcription start sites (TSS) and is important for actively transcribed gene regulation [90, 91]. Interestingly, the *KMT2D* missense variants functionally tested in this study, affect the methyltransferase activity in different manner. For instance, the missense variants localized within the PHD 4-6 seems to barely reduce the monomethylation H3K4 (me1) levels of H3K4 whereas the missense variants localized within the PHD7, FYRN, WIN and SET domains, also affect the H3K4(me2) and H3K4(me3) levels, suggesting a more relevant role in the aetiology of the disease.

Because *KMT2D* is active in the context of a multisubunit complex, with the minimal four-component complex, including WDR5, RbBP5 and Ash2L along with the *KMT2D* SET-

domain subunit that can reconstitute most of the H3K4-specific HMT activity, we tested the ability of the 14 *KMT2D* missense variants to interact with RbBP5 and Ash2L respectively. In both assays, the missense variants localized in the C-terminal domains impaired/reduced the interaction with KMT2D partners compared to that of the wild-type protein, whereas most of the missense variants localized within the PHD 4-6 increase the interaction levels. Interestingly, the missense variant R5340Q, localized in the WIN domain, also disrupt the interaction with WDR5, suggesting that for these missense variants, the reduced activity can be due to the misleading KMT2D interaction with the WRAD complex.

Because our in vitro methyltransferase and co-immunoprecipitation assays showed reduced activity for *KMT2D* missense variant Y5510D, we determined whether the expression of this transiently transfected *KMT2D* missense variant could rescue the decreased levels of HOXA1–3 mRNA during RA treatment in NT2D1 cell line knock down for KMT2D. Consistent with previous results, the missense variant Y5510D did not restore the expression levels of target genes, while the missense variant R5471M completely rescued defective differentiation mRNA levels, suggesting alternative effects of this variant on the pathogenicity of the disease.

Although additional studies will be required to dissect the precise mechanisms underlying the pathogenic role of the missense variants in KS, together these findings indicate that missense variants in Kabuki Syndrome may affect the disease status by affecting: i) the methyltransferase activity of the protein, ii) the interaction with the WRAD complex or/and iii) gene expression during cellular differentiation. Classification and interpretation of missense variants is a challenge in molecular diagnostics and genetic counselling. A number of prediction tools have been implemented in the last years, mainly as support for NGS data. However, in silico analyses should not to be considered as conclusive evidence in the assessment of variants of unknown clinical significance. The American College of Medical Genetics and Genomics and the Association for Molecular Pathology Standards and Guidelines for the interpretation of sequence variants recommended that these predictions should be used as support to additional findings, including functional evidence that can prove the pathogenicity of variants found. Recently two different DNA methylation profiling studies showed that KS patients present a highly specific and univocal DNA methylation signature that has the potential to be used as a diagnostic method for differentiating the samples with pathogenic mutations from those with benign variants and therefore enabling the functional assessment of genetic variants of unknown clinical significance [69].

The biochemical approaches that we presented here could be of support to these tools with the advantage that we can finally address the functional effect of the variants, although we are

fully aware that not all the Diagnostic Laboratories can be perform routinely such combination of bioinformatics and functional tools.

CONCLUSIONS

This study expands the number of *KMT2D* and *KDM6A* mutations that cause KS, investigates the potential drug-mediated readthrough therapeutical approach and more relevant for diagnostic and counselling purposes, provides a proof-of-concept that a number of naturally occurring missense mutations in *KMT2D*, by affecting *KMT2D* interaction and H3K4 methylation activity, should be considered as loss of function variants, while the remaining tested nucleotide variants may represent benign changes.

ACKNOWLEDGMENTS

We acknowledge the family that agreed to participate and made this study possible. We are grateful to the Genomic Disorder Biobank and Telethon Network of Genetic Biobanks (Telethon Italy grant GTB12001G) for banking some of the biospecimens used into the study. The financial support of Telethon - Italy (Grant no. GGP13231), Italian Ministry of Health (Ricerca Corrente), Jérôme Lejeune Foundation, and Daunia Plast (Private Donor) to GM is gratefully acknowledged. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors declare no conflicts of interest with the exception GM who was a paid consultant for Takeda Pharmaceutical Company for 2016.

REFERENCES

1. Kuroki, Y.; Suzuki, Y.; Chyo, H.; Hata, A.; Matsui, I., A new malformation syndrome of long palpebral fissures, large ears, depressed nasal tip, and skeletal anomalies associated with postnatal dwarfism and mental retardation. *J Pediatr* **1981**, *99*, (4), 570-3.
2. Niikawa, N.; Matsuura, N.; Fukushima, Y.; Ohsawa, T.; Kajii, T., Kabuki make-up syndrome: a syndrome of mental retardation, unusual facies, large and protruding ears, and postnatal growth deficiency. *J Pediatr* **1981**, *99*, (4), 565-9.
3. Cheon, C. K.; Ko, J. M., Kabuki syndrome: clinical and molecular characteristics. *Korean J Pediatr* **2015**, *58*, (9), 317-24.
4. Micale, L.; Augello, B.; Fusco, C.; Selicorni, A.; Loviglio, M. N.; Silengo, M. C.; Reymond, A.; Gumiero, B.; Zucchetti, F.; D'Addetta, E. V.; Belligni, E.; Calcagni, A.; Digilio, M. C.; Dallapiccola, B.; Faravelli, F.; Forzano, F.; Accadia, M.; Bonfante, A.; Clementi, M.; Daolio, C.; Douzgou, S.; Ferrari, P.; Fischetto, R.; Garavelli, L.; Lapi, E.; Mattina, T.; Melis, D.; Patricelli, M. G.; Priolo, M.; Prontera, P.; Renieri, A.; Mencarelli, M. A.; Scarano, G.; della Monica, M.; Toschi, B.; Turolla, L.; Vancini, A.; Zatterale, A.; Gabrielli, O.; Zelante, L.; Merla, G., Mutation spectrum of MLL2 in a cohort of Kabuki syndrome patients. *Orphanet J Rare Dis* **2011**, *6*, 38.
5. Burke, L. W.; Jones, M. C., Kabuki syndrome: underdiagnosed recognizable pattern in cleft palate patients. *Cleft Palate Craniofac J* **1995**, *32*, (1), 77-84.
6. Armstrong, L.; Abd El Moneim, A.; Aleck, K.; Aughton, D. J.; Baumann, C.; Braddock, S. R.; Gillessen-Kaesbach, G.; Graham, J. M., Jr.; Grebe, T. A.; Gripp, K. W.; Hall, B. D.; Hennekam, R.; Hunter, A.; Keppler-Noreuil, K.; Lacombe, D.; Lin, A. E.; Ming, J. E.; Kokitsu-Nakata, N. M.; Nikkel, S. M.; Philip, N.; Raas-Rothschild, A.; Sommer, A.; Verloes, A.; Walter, C.; Wieczorek, D.; Williams, M. S.; Zackai, E.; Allanson, J. E., Further delineation of Kabuki syndrome in 48 well-defined new individuals. *Am J Med Genet A* **2005**, *132A*, (3), 265-72.
7. Kawame, H.; Hannibal, M. C.; Hudgins, L.; Pagon, R. A., Phenotypic spectrum and management issues in Kabuki syndrome. *J Pediatr* **1999**, *134*, (4), 480-5.
8. Torii, Y.; Yagasaki, H.; Tanaka, H.; Mizuno, S.; Nishio, N.; Muramatsu, H.; Hama, A.; Takahashi, Y.; Kojima, S., Successful treatment with rituximab of refractory idiopathic thrombocytopenic purpura in a patient with Kabuki syndrome. *Int J Hematol* **2009**, *90*, (2), 174-6.
9. Hoffman, J. D.; Ciprero, K. L.; Sullivan, K. E.; Kaplan, P. B.; McDonald-McGinn, D. M.; Zackai, E. H.; Ming, J. E., Immune abnormalities are a frequent manifestation of Kabuki syndrome. *Am J Med Genet A* **2005**, *135*, (3), 278-81.
10. Ho, H. H.; Eaves, L. C., Kabuki make-up (Niikawa-Kuroki) syndrome: cognitive abilities and autistic features. *Dev Med Child Neurol* **1997**, *39*, (7), 487-90.
11. Oksanen, V. E.; Arvio, M. A.; Peippo, M. M.; Valanne, L. K.; Sainio, K. O., Temporo-occipital spikes: a typical EEG finding in Kabuki syndrome. *Pediatr Neurol* **2004**, *30*, (1), 67-70.
12. Niikawa, N.; Kuroki, Y.; Kajii, T.; Matsuura, N.; Ishikiriyama, S.; Tonoki, H.; Ishikawa, N.; Yamada, Y.; Fujita, M.; Umemoto, H.; et al., Kabuki make-up (Niikawa-Kuroki) syndrome: a study of 62 patients. *Am J Med Genet* **1988**, *31*, (3), 565-89.
13. Adam, M. P.; Hudgins, L., Kabuki syndrome: a review. *Clin Genet* **2005**, *67*, (3), 209-19.
14. Ng, S. B.; Bigham, A. W.; Buckingham, K. J.; Hannibal, M. C.; McMillin, M. J.; Gildersleeve, H. I.; Beck, A. E.; Tabor, H. K.; Cooper, G. M.; Mefford, H. C.; Lee, C.;

- Turner, E. H.; Smith, J. D.; Rieder, M. J.; Yoshiura, K.; Matsumoto, N.; Ohta, T.; Niikawa, N.; Nickerson, D. A.; Bamshad, M. J.; Shendure, J., Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* **2010**, *42*, (9), 790-3.
15. Shilatifard, A., The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem* **2012**, *81*, 65-95.
 16. Kouzarides, T., Chromatin modifications and their function. *Cell* **2007**, *128*, (4), 693-705.
 17. Li, B.; Carey, M.; Workman, J. L., The role of chromatin during transcription. *Cell* **2007**, *128*, (4), 707-19.
 18. Bogershausen, N.; Gatinois, V.; Riehmer, V.; Kayserili, H.; Becker, J.; Thoenes, M.; Simsek-Kiper, P. O.; Barat-Houari, M.; Elcioglu, N. H.; Wiczorek, D.; Tinschert, S.; Sarrabay, G.; Strom, T. M.; Fabre, A.; Baynam, G.; Sanchez, E.; Nurnberg, G.; Altunoglu, U.; Capri, Y.; Isidor, B.; Lacombe, D.; Corsini, C.; Cormier-Daire, V.; Sanlaville, D.; Giuliano, F.; Le Quan Sang, K. H.; Kayirangwa, H.; Nurnberg, P.; Meitinger, T.; Boduroglu, K.; Zoll, B.; Lyonnet, S.; Tzschach, A.; Verloes, A.; Di Donato, N.; Touitou, I.; Netzer, C.; Li, Y.; Genevieve, D.; Yigit, G.; Wollnik, B., Mutation Update for Kabuki Syndrome Genes KMT2D and KDM6A and Further Delineation of X-Linked Kabuki Syndrome Subtype 2. *Hum Mutat* **2016**.
 19. Ang, S. Y.; Uebersohn, A.; Spencer, C. I.; Huang, Y.; Lee, J. E.; Ge, K.; Bruneau, B. G., KMT2D regulates specific programs in heart development via histone H3 lysine 4 di-methylation. *Development* **2016**, *143*, (5), 810-21.
 20. Herz, H. M.; Mohan, M.; Garruss, A. S.; Liang, K.; Takahashi, Y. H.; Mickey, K.; Voets, O.; Verrijzer, C. P.; Shilatifard, A., Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. *Genes Dev* **2012**, *26*, (23), 2604-20.
 21. Hu, D.; Gao, X.; Morgan, M. A.; Herz, H. M.; Smith, E. R.; Shilatifard, A., The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol Cell Biol* **2013**, *33*, (23), 4745-54.
 22. Lee, J. E.; Wang, C.; Xu, S.; Cho, Y. W.; Wang, L.; Feng, X.; Baldrige, A.; Sartorelli, V.; Zhuang, L.; Peng, W.; Ge, K., H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife (Cambridge)* **2013**, *2*, e01503.
 23. Guo, C.; Chen, L. H.; Huang, Y.; Chang, C. C.; Wang, P.; Pirozzi, C. J.; Qin, X.; Bao, X.; Greer, P. K.; McLendon, R. E.; Yan, H.; Keir, S. T.; Bigner, D. D.; He, Y., KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. *Oncotarget* **2013**, *4*, (11), 2144-53.
 24. Kaikkonen, M. U.; Spann, N. J.; Heinz, S.; Romanoski, C. E.; Allison, K. A.; Stender, J. D.; Chun, H. B.; Tough, D. F.; Prinjha, R. K.; Benner, C.; Glass, C. K., Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol Cell* **2013**, *51*, (3), 310-25.
 25. Issaeva, I.; Zonis, Y.; Rozovskaia, T.; Orlovsky, K.; Croce, C. M.; Nakamura, T.; Mazo, A.; Eisenbach, L.; Canaani, E., Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. *Mol Cell Biol* **2007**, *27*, (5), 1889-903.
 26. Ansari, K. I.; Mandal, S. S., Mixed lineage leukemia: roles in gene expression, hormone signaling and mRNA processing. *FEBS J* **2010**, *277*, (8), 1790-804.
 27. Eissenberg, J. C.; Shilatifard, A., Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Dev Biol* **2010**, *339*, (2), 240-9.

28. Patel, A.; Dharmarajan, V.; Vought, V. E.; Cosgrove, M. S., On the mechanism of multiple lysine methylation by the human mixed lineage leukemia protein-1 (MLL1) core complex. *J Biol Chem* **2009**, 284, (36), 24242-56.
29. Dou, Y.; Milne, T. A.; Ruthenburg, A. J.; Lee, S.; Lee, J. W.; Verdine, G. L.; Allis, C. D.; Roeder, R. G., Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol* **2006**, 13, (8), 713-9.
30. Steward, M. M.; Lee, J. S.; O'Donovan, A.; Wyatt, M.; Bernstein, B. E.; Shilatifard, A., Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. *Nat Struct Mol Biol* **2006**, 13, (9), 852-4.
31. Ali, A.; Veeranki, S. N.; Tyagi, S., A SET-domain-independent role of WRAD complex in cell-cycle regulatory function of mixed lineage leukemia. *Nucleic Acids Res* **2014**, 42, (12), 7611-24.
32. Van der Meulen, J.; Speleman, F.; Van Vlierberghe, P., The H3K27me3 demethylase UTX in normal development and disease. *Epigenetics* **2014**, 9, (5), 658-68.
33. Agger, K.; Cloos, P. A.; Christensen, J.; Pasini, D.; Rose, S.; Rappsilber, J.; Issaeva, I.; Canaani, E.; Salcini, A. E.; Helin, K., UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **2007**, 449, (7163), 731-4.
34. Lee, M. G.; Villa, R.; Trojer, P.; Norman, J.; Yan, K. P.; Reinberg, D.; Di Croce, L.; Shiekhhattar, R., Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* **2007**, 318, (5849), 447-50.
35. Wang, C.; Lee, J. E.; Lai, B.; Macfarlan, T. S.; Xu, S.; Zhuang, L.; Liu, C.; Peng, W.; Ge, K., Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. *Proc Natl Acad Sci U S A* **2016**, 113, (42), 11871-11876.
36. Lai, B.; Lee, J. E.; Jang, Y.; Wang, L.; Peng, W.; Ge, K., MLL3/MLL4 are required for CBP/p300 binding on enhancers and super-enhancer formation in brown adipogenesis. *Nucleic Acids Res* **2017**.
37. Van Laarhoven, P. M.; Neitzel, L. R.; Quintana, A. M.; Geiger, E. A.; Zackai, E. H.; Clouthier, D. E.; Artinger, K. B.; Ming, J. E.; Shaikh, T. H., Kabuki syndrome genes KMT2D and KDM6A: functional analyses demonstrate critical roles in craniofacial, heart and brain development. *Hum Mol Genet* **2015**, 24, (15), 4443-53.
38. Zhang, J.; Dominguez-Sola, D.; Hussein, S.; Lee, J. E.; Holmes, A. B.; Bansal, M.; Vlasevska, S.; Mo, T.; Tang, H.; Basso, K.; Ge, K.; Dalla-Favera, R.; Pasqualucci, L., Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat Med* **2015**, 21, (10), 1190-8.
39. Froimchuk, E.; Jang, Y.; Ge, K., Histone H3 lysine 4 methyltransferase KMT2D. *Gene* **2017**, 627, 337-342.
40. Micale, L.; Merla, G., Molecular Genetics of Kabuki Syndrome *eLs* **2012**.
41. Davis, M. B.; SanGil, I.; Berry, G.; Olayokun, R.; Neves, L. H., Identification of common and cell type specific LXXLL motif EcR cofactors using a bioinformatics refined candidate RNAi screen in *Drosophila melanogaster* cell lines. *BMC Dev Biol* **2011**, 11, 66.
42. Mo, R.; Rao, S. M.; Zhu, Y. J., Identification of the MLL2 complex as a coactivator for estrogen receptor alpha. *J Biol Chem* **2006**, 281, (23), 15714-20.
43. Dreijerink, K. M.; Mulder, K. W.; Winkler, G. S.; Hoppener, J. W.; Lips, C. J.; Timmers, H. T., Menin links estrogen receptor activation to histone H3K4 trimethylation. *Cancer Res* **2006**, 66, (9), 4929-35.

44. Hughes, C. L.; Liu, P. Z.; Kaufman, T. C., Expression patterns of the rogue Hox genes Hox3/zen and fushi tarazu in the apterygote insect *Thermobia domestica*. *Evol Dev* **2004**, 6, (6), 393-401.
45. Wang, P.; Lin, C.; Smith, E. R.; Guo, H.; Sanderson, B. W.; Wu, M.; Gogol, M.; Alexander, T.; Seidel, C.; Wiedemann, L. M.; Ge, K.; Krumlau, R.; Shilatifard, A., Global analysis of H3K4 methylation defines MLL family member targets and points to a role for MLL1-mediated H3K4 methylation in the regulation of transcriptional initiation by RNA polymerase II. *Mol Cell Biol* **2009**, 29, (22), 6074-85.
46. Garcia-Alai, M. M.; Allen, M. D.; Joerger, A. C.; Bycroft, M., The structure of the FYR domain of transforming growth factor beta regulator 1. *Protein Sci* **2010**, 19, (7), 1432-8.
47. Jenuwein, T.; Laible, G.; Dorn, R.; Reuter, G., SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell Mol Life Sci* **1998**, 54, (1), 80-93.
48. Tschiersch, B.; Hofmann, A.; Krauss, V.; Dorn, R.; Korge, G.; Reuter, G., The protein encoded by the *Drosophila* position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J* **1994**, 13, (16), 3822-31.
49. Jones, R. S.; Gelbart, W. M., The *Drosophila* Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. *Mol Cell Biol* **1993**, 13, (10), 6357-66.
50. Stassen, M. J.; Bailey, D.; Nelson, S.; Chinwalla, V.; Harte, P. J., The *Drosophila* trithorax proteins contain a novel variant of the nuclear receptor type DNA binding domain and an ancient conserved motif found in other chromosomal proteins. *Mech Dev* **1995**, 52, (2-3), 209-23.
51. Dillon, S. C.; Zhang, X.; Trievel, R. C.; Cheng, X., The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol* **2005**, 6, (8), 227.
52. Malik, S.; Bhaumik, S. R., Mixed lineage leukemia: histone H3 lysine 4 methyltransferases from yeast to human. *FEBS J* **2010**, 277, (8), 1805-21.
53. Lederer, D.; Grisart, B.; Digilio, M. C.; Benoit, V.; Crespin, M.; Ghariani, S. C.; Maystadt, I.; Dallapiccola, B.; Verellen-Dumoulin, C., Deletion of KDM6A, a Histone Demethylase Interacting with MLL2, in Three Patients with Kabuki Syndrome. *Am J Hum Genet* **2012**, 90, (1), 119-24.
54. Xu, J.; Deng, X.; Watkins, R.; Distech, C. M., Sex-specific differences in expression of histone demethylases Utx and Uty in mouse brain and neurons. *J Neurosci* **2008**, 28, (17), 4521-7.
55. Tsukada, Y.; Fang, J.; Erdjument-Bromage, H.; Warren, M. E.; Borchers, C. H.; Tempst, P.; Zhang, Y., Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **2006**, 439, (7078), 811-6.
56. Lan, F.; Bayliss, P. E.; Rinn, J. L.; Whetstine, J. R.; Wang, J. K.; Chen, S.; Iwase, S.; Alpatov, R.; Issaeva, I.; Canaani, E.; Roberts, T. M.; Chang, H. Y.; Shi, Y., A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* **2007**, 449, (7163), 689-94.
57. Hong, S.; Cho, Y. W.; Yu, L. R.; Yu, H.; Veenstra, T. D.; Ge, K., Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc Natl Acad Sci U S A* **2007**, 104, (47), 18439-44.
58. Schuettengruber, B.; Chourrout, D.; Vervoort, M.; Leblanc, B.; Cavalli, G., Genome regulation by polycomb and trithorax proteins. *Cell* **2007**, 128, (4), 735-45.

59. Pasini, D.; Bracken, A. P.; Agger, K.; Christensen, J.; Hansen, K.; Cloos, P. A.; Helin, K., Regulation of stem cell differentiation by histone methyltransferases and demethylases. *Cold Spring Harb Symp Quant Biol* **2008**, 73, 253-63.
60. Aziz, A.; Liu, Q. C.; Dilworth, F. J., Regulating a master regulator: establishing tissue-specific gene expression in skeletal muscle. *Epigenetics* **2010**, 5, (8), 691-5.
61. Herz, H. M.; Madden, L. D.; Chen, Z.; Bolduc, C.; Buff, E.; Gupta, R.; Davuluri, R.; Shilatifard, A.; Hariharan, I. K.; Bergmann, A., The H3K27me3 demethylase dUTX is a suppressor of Notch- and Rb-dependent tumors in Drosophila. *Mol Cell Biol* **2010**, 30, (10), 2485-97.
62. Seenundun, S.; Rampalli, S.; Liu, Q. C.; Aziz, A.; Palii, C.; Hong, S.; Blais, A.; Brand, M.; Ge, K.; Dilworth, F. J., UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. *EMBO J* **2010**, 29, (8), 1401-11.
63. Miller, S. A.; Mohn, S. E.; Weinmann, A. S., Jmjd3 and UTX play a demethylase-independent role in chromatin remodeling to regulate T-box family member-dependent gene expression. *Mol Cell* **2010**, 40, (4), 594-605.
64. Kesler, S. R., Turner syndrome. *Child Adolesc Psychiatr Clin N Am* **2007**, 16, (3), 709-22.
65. Matsumoto, N.; Niikawa, N., Kabuki make-up syndrome: a review. *Am J Med Genet C Semin Med Genet* **2003**, 117C, (1), 57-65.
66. Deng, X.; Berletch, J. B.; Nguyen, D. K.; Disteche, C. M., X chromosome regulation: diverse patterns in development, tissues and disease. *Nat Rev Genet* **2014**, 15, (6), 367-78.
67. Disteche, C. M.; Berletch, J. B., X-chromosome inactivation and escape. *J Genet* **2015**, 94, (4), 591-9.
68. Zentner, G. E.; Layman, W. S.; Martin, D. M.; Scacheri, P. C., Molecular and phenotypic aspects of CHD7 mutation in CHARGE syndrome. *Am J Med Genet A* **2010**, 152A, (3), 674-86.
69. Butcher, D. T.; Cytrynbaum, C.; Turinsky, A. L.; Siu, M. T.; Inbar-Feigenberg, M.; Mendoza-Londono, R.; Chitayat, D.; Walker, S.; Machado, J.; Caluseriu, O.; Dupuis, L.; Grafodatskaya, D.; Reardon, W.; Gilbert-Dussardier, B.; Verloes, A.; Bilan, F.; Milunsky, J. M.; Basran, R.; Papsin, B.; Stockley, T. L.; Scherer, S. W.; Choufani, S.; Brudno, M.; Weksberg, R., CHARGE and Kabuki Syndromes: Gene-Specific DNA Methylation Signatures Identify Epigenetic Mechanisms Linking These Clinically Overlapping Conditions. *Am J Hum Genet* **2017**, 100, (5), 773-788.
70. Adam, M. P.; Hudgins, L.; Hannibal, M., Kabuki Syndrome. **1993**.
71. McDonald-McGinn, D. M.; Sullivan, K. E.; Marino, B.; Philip, N.; Swillen, A.; Vorstman, J. A.; Zackai, E. H.; Emanuel, B. S.; Vermeesch, J. R.; Morrow, B. E.; Scambler, P. J.; Bassett, A. S., 22q11.2 deletion syndrome. *Nat Rev Dis Primers* **2015**, 1, 15071.
72. McDonald-McGinn, D. M.; Emanuel, B. S.; Zackai, E. H., 22q11.2 Deletion Syndrome. **1993**.
73. Schutte, B. C.; Saal, H. M.; Goudy, S.; Leslie, E., IRF6-Related Disorders. **1993**.
74. Smith, R. J. H., Branchiootorenal Spectrum Disorders. **1993**.
75. Poley, J. R.; Proud, V. K., Hardikar syndrome: new features. *Am J Med Genet A* **2008**, 146A, (19), 2473-9.
76. Miyake, N.; Tsurusaki, Y.; Koshimizu, E.; Okamoto, N.; Kosho, T.; Brown, N. J.; Tan, T. Y.; Yap, P. J.; Suzumura, H.; Tanaka, T.; Nagai, T.; Nakashima, M.; Saitsu, H.; Niikawa, N.; Matsumoto, N., Delineation of clinical features in Wiedemann-Steiner syndrome caused by KMT2A mutations. *Clin Genet* **2016**, 89, (1), 115-9.
77. Jones, W. D.; Dafou, D.; McEntagart, M.; Woollard, W. J.; Elmslie, F. V.; Holder-Espinasse, M.; Irving, M.; Sagar, A. K.; Smithson, S.; Trembath, R. C.; Deshpande,

- C.; Simpson, M. A., De novo mutations in MLL cause Wiedemann-Steiner syndrome. *Am J Hum Genet* **2012**, 91, (2), 358-64.
78. Liu, L.; Li, Y.; Tollefsbol, T. O., Gene-environment interactions and epigenetic basis of human diseases. *Curr Issues Mol Biol* **2008**, 10, (1-2), 25-36.
 79. Conerly, M.; Grady, W. M., Insights into the role of DNA methylation in disease through the use of mouse models. *Dis Model Mech* **2010**, 3, (5-6), 290-7.
 80. Choufani, S.; Cytrynbaum, C.; Chung, B. H.; Turinsky, A. L.; Grafodatskaya, D.; Chen, Y. A.; Cohen, A. S.; Dupuis, L.; Butcher, D. T.; Siu, M. T.; Luk, H. M.; Lo, I. F.; Lam, S. T.; Caluseriu, O.; Stavropoulos, D. J.; Reardon, W.; Mendoza-Londono, R.; Brudno, M.; Gibson, W. T.; Chitayat, D.; Weksberg, R., NSD1 mutations generate a genome-wide DNA methylation signature. *Nat Commun* **2015**, 6, 10207.
 81. 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., Multilocus loss of DNA methylation in individuals with mutations in the histone H3 lysine 4 demethylase KDM5C. *BMC Med Genomics* **2013**, 6, 1.
 82. Kernohan, K. D.; Cigana Schenkel, L.; Huang, L.; Smith, A.; Pare, G.; Ainsworth, P.; Boycott, K. M.; Warman-Chardon, J.; Sadikovic, B., Identification of a methylation profile for DNMT1-associated autosomal dominant cerebellar ataxia, deafness, and narcolepsy. *Clin Epigenetics* **2016**, 8, 91.
 83. Sobreira, N.; Martha Brucato; Li Zhang; Christine Ladd-Acosta; Chrissie Ongaco; Jane Romm; Kimberly F Doheny; Regina C Mingroni-Netto; Debora Bertola; Chong A Kim; Ana BA Perez; Maria I Melaragno; David Valle; Bjornsson, V. A. M. H. T.; , Patients with a Kabuki syndrome phenotype demonstrate DNA methylation abnormalities. *European Journal of Human Genetics* **2017**.
 84. Choufani, S.; Shuman, C.; Weksberg, R., Beckwith-Wiedemann syndrome. *Am J Med Genet C Semin Med Genet* **2010**, 154C, (3), 343-54.
 85. Eggermann, T.; Begemann, M.; Spengler, S.; Schroder, C.; Kordass, U.; Binder, G., Genetic and epigenetic findings in Silver-Russell syndrome. *Pediatr Endocrinol Rev* **2010**, 8, (2), 86-93.
 86. Hansen, J. C., Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annu Rev Biophys Biomol Struct* **2002**, 31, 361-92.
 87. Carruthers, L. M.; Hansen, J. C., The core histone N termini function independently of linker histones during chromatin condensation. *J Biol Chem* **2000**, 275, (47), 37285-90.
 88. Martin, C.; Zhang, Y., The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol* **2005**, 6, (11), 838-49.
 89. Guenther, M. G.; Jenner, R. G.; Chevalier, B.; Nakamura, T.; Croce, C. M.; Canaani, E.; Young, R. A., Global and Hox-specific roles for the MLL1 methyltransferase. *Proc Natl Acad Sci U S A* **2005**, 102, (24), 8603-8.
 90. Heintzman, N. D.; Hon, G. C.; Hawkins, R. D.; Kheradpour, P.; Stark, A.; Harp, L. F.; Ye, Z.; Lee, L. K.; Stuart, R. K.; Ching, C. W.; Ching, K. A.; Antosiewicz-Bourget, J. E.; Liu, H.; Zhang, X.; Green, R. D.; Lobanenkov, V. V.; Stewart, R.; Thomson, J. A.; Crawford, G. E.; Kellis, M.; Ren, B., Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **2009**, 459, (7243), 108-12.
 91. Denissov, S.; Hofemeister, H.; Marks, H.; Kranz, A.; Ciotta, G.; Singh, S.; Anastassiadis, K.; Stunnenberg, H. G.; Stewart, A. F., Mll2 is required for H3K4

- trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development* **2014**, 141, (3), 526-37.
92. Ng, H. H.; Robert, F.; Young, R. A.; Struhl, K., Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **2003**, 11, (3), 709-19.
 93. Krogan, N. J.; Kim, M.; Tong, A.; Golshani, A.; Cagney, G.; Canadien, V.; Richards, D. P.; Beattie, B. K.; Emili, A.; Boone, C.; Shilatifard, A.; Buratowski, S.; Greenblatt, J., Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* **2003**, 23, (12), 4207-18.
 94. Xiao, T.; Hall, H.; Kizer, K. O.; Shibata, Y.; Hall, M. C.; Borchers, C. H.; Strahl, B. D., Phosphorylation of RNA polymerase II CTD regulates H3 methylation in yeast. *Genes Dev* **2003**, 17, (5), 654-63.
 95. Li, B.; Howe, L.; Anderson, S.; Yates, J. R., 3rd; Workman, J. L., The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J Biol Chem* **2003**, 278, (11), 8897-903.
 96. Cao, R.; Zhang, Y., The functions of E(Z)/EZH2-mediated methylation of lysine 27 in histone H3. *Curr Opin Genet Dev* **2004**, 14, (2), 155-64.
 97. Ringrose, L.; Paro, R., Epigenetic regulation of cellular memory by the Polycomb and Trithorax group proteins. *Annu Rev Genet* **2004**, 38, 413-43.
 98. Cao, R.; Zhang, Y., SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex. *Mol Cell* **2004**, 15, (1), 57-67.
 99. Pasini, D.; Bracken, A. P.; Jensen, M. R.; Lazzerini Denchi, E.; Helin, K., Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO J* **2004**, 23, (20), 4061-71.
 100. Shi, Y.; Lan, F.; Matson, C.; Mulligan, P.; Whetstine, J. R.; Cole, P. A.; Casero, R. A., Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **2004**, 119, (7), 941-53.
 101. Mosammaparast, N.; Shi, Y., Reversal of histone methylation: biochemical and molecular mechanisms of histone demethylases. *Annu Rev Biochem* **2010**, 79, 155-79.
 102. Klose, R. J.; Yamane, K.; Bae, Y.; Zhang, D.; Erdjument-Bromage, H.; Tempst, P.; Wong, J.; Zhang, Y., The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36. *Nature* **2006**, 442, (7100), 312-6.
 103. Bjornsson, H. T.; Benjamin, J. S.; Zhang, L.; Weissman, J.; Gerber, E. E.; Chen, Y. C.; Vaurio, R. G.; Potter, M. C.; Hansen, K. D.; Dietz, H. C., Histone deacetylase inhibition rescues structural and functional brain deficits in a mouse model of Kabuki syndrome. *Sci Transl Med* **2014**, 6, (256), 256ra135.
 104. Mirabella, A. C.; Foster, B. M.; Bartke, T., Chromatin deregulation in disease. *Chromosoma* **2016**, 125, (1), 75-93.
 105. van Bokhoven, H.; Kramer, J. M., Disruption of the epigenetic code: an emerging mechanism in mental retardation. *Neurobiol Dis* **2010**, 39, (1), 3-12.
 106. Graff, J.; Mansuy, I. M., Epigenetic dysregulation in cognitive disorders. *Eur J Neurosci* **2009**, 30, (1), 1-8.
 107. Kleefstra, T.; Schenck, A.; Kramer, J. M.; van Bokhoven, H., The genetics of cognitive epigenetics. *Neuropharmacology* **2014**, 80, 83-94.
 108. Wapenaar, H.; Dekker, F. J., Histone acetyltransferases: challenges in targeting bi-substrate enzymes. *Clin Epigenetics* **2016**, 8, 59.
 109. Roelfsema, J. H.; White, S. J.; Ariyurek, Y.; Bartholdi, D.; Niedrist, D.; Papadia, F.; Bacino, C. A.; den Dunnen, J. T.; van Ommen, G. J.; Breuning, M. H.; Hennekam, R.

- C.; Peters, D. J., Genetic heterogeneity in Rubinstein-Taybi syndrome: mutations in both the CBP and EP300 genes cause disease. *Am J Hum Genet* **2005**, *76*, (4), 572-80.
110. Petrij, F.; Giles, R. H.; Dauwerse, H. G.; Saris, J. J.; Hennekam, R. C.; Masuno, M.; Tommerup, N.; van Ommen, G. J.; Goodman, R. H.; Peters, D. J.; et al., Rubinstein-Taybi syndrome caused by mutations in the transcriptional co-activator CBP. *Nature* **1995**, *376*, (6538), 348-51.
 111. Spinner, N. B.; Conlin, L. K., Mosaicism and clinical genetics. *Am J Med Genet C Semin Med Genet* **2014**, *166C*, (4), 397-405.
 112. Banka, S.; Howard, E.; Bunstone, S.; Chandler, K. E.; Kerr, B.; Lachlan, K.; McKee, S.; Mehta, S. G.; Tavares, A. L.; Tolmie, J.; Donnai, D., MLL2 mosaic mutations and intragenic deletion-duplications in patients with Kabuki syndrome. *Clin Genet* **2013**, *83*, (5), 467-71.
 113. Lepri, F. R.; Cocciadiferro, D.; Augello, B.; Alfieri, P.; Pes, V.; Vancini, A.; Caciolo, C.; Squeo, G. M.; Malerba, N.; Adipietro, I.; Novelli, A.; Sotgiu, S.; Gherardi, R.; Digilio, M. C.; Dallapiccola, B.; Merla, G., Clinical and Neurobehavioral Features of Three Novel Kabuki Syndrome Patients with Mosaic KMT2D Mutations and a Review of Literature. *Int J Mol Sci* **2017**, *19*, (1).
 114. Frischmeyer, P. A.; Dietz, H. C., Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet* **1999**, *8*, (10), 1893-900.
 115. Mendell, J. T.; Dietz, H. C., When the message goes awry: disease-producing mutations that influence mRNA content and performance. *Cell* **2001**, *107*, (4), 411-4.
 116. Perez, B.; Rodriguez-Pombo, P.; Ugarte, M.; Desviat, L. R., Readthrough strategies for therapeutic suppression of nonsense mutations in inherited metabolic disease. *Mol Syndromol* **2012**, *3*, (5), 230-6.
 117. Welch, E. M.; Barton, E. R.; Zhuo, J.; Tomizawa, Y.; Friesen, W. J.; Trifillis, P.; Paushkin, S.; Patel, M.; Trotta, C. R.; Hwang, S.; Wilde, R. G.; Karp, G.; Takasugi, J.; Chen, G.; Jones, S.; Ren, H.; Moon, Y. C.; Corson, D.; Turpoff, A. A.; Campbell, J. A.; Conn, M. M.; Khan, A.; Almstead, N. G.; Hedrick, J.; Mollin, A.; Risher, N.; Weetall, M.; Yeh, S.; Branstrom, A. A.; Colacino, J. M.; Babiak, J.; Ju, W. D.; Hirawat, S.; Northcutt, V. J.; Miller, L. L.; Spatrnick, P.; He, F.; Kawana, M.; Feng, H.; Jacobson, A.; Peltz, S. W.; Sweeney, H. L., PTC124 targets genetic disorders caused by nonsense mutations. *Nature* **2007**, *447*, (7140), 87-91.
 118. Hirawat, S.; Welch, E. M.; Elfring, G. L.; Northcutt, V. J.; Paushkin, S.; Hwang, S.; Leonard, E. M.; Almstead, N. G.; Ju, W.; Peltz, S. W.; Miller, L. L., Safety, tolerability, and pharmacokinetics of PTC124, a nonaminoglycoside nonsense mutation suppressor, following single- and multiple-dose administration to healthy male and female adult volunteers. *J Clin Pharmacol* **2007**, *47*, (4), 430-44.
 119. Du, M.; Liu, X.; Welch, E. M.; Hirawat, S.; Peltz, S. W.; Bedwell, D. M., PTC124 is an orally bioavailable compound that promotes suppression of the human CFTR-G542X nonsense allele in a CF mouse model. *Proc Natl Acad Sci U S A* **2008**, *105*, (6), 2064-9.
 120. Lee, H. L.; Dougherty, J. P., Pharmaceutical therapies to recode nonsense mutations in inherited diseases. *Pharmacol Ther* **2012**, *136*, (2), 227-66.
 121. Linde, L.; Kerem, B., Introducing sense into nonsense in treatments of human genetic diseases. *Trends Genet* **2008**, *24*, (11), 552-63.
 122. Recht, M. I.; Douthwaite, S.; Puglisi, J. D., Basis for prokaryotic specificity of action of aminoglycoside antibiotics. *EMBO J* **1999**, *18*, (11), 3133-8.

123. Helip-Wooley, A.; Park, M. A.; Lemons, R. M.; Thoene, J. G., Expression of CTNS alleles: subcellular localization and aminoglycoside correction in vitro. *Mol Genet Metab* **2002**, 75, (2), 128-33.
124. Lewis, B. P.; Shih, I. H.; Jones-Rhoades, M. W.; Bartel, D. P.; Burge, C. B., Prediction of mammalian microRNA targets. *Cell* **2003**, 115, (7), 787-98.
125. Rehwinkel, J.; Raes, J.; Izaurrealde, E., Nonsense-mediated mRNA decay: Target genes and functional diversification of effectors. *Trends Biochem Sci* **2006**, 31, (11), 639-46.
126. Nagy, E.; Maquat, L. E., A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. *Trends Biochem Sci* **1998**, 23, (6), 198-9.
127. Brogna, S.; Wen, J., Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol* **2009**, 16, (2), 107-13.
128. Amrani, N.; Ganesan, R.; Kervestin, S.; Mangus, D. A.; Ghosh, S.; Jacobson, A., A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* **2004**, 432, (7013), 112-8.
129. Le Hir, H.; Moore, M. J.; Maquat, L. E., Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. *Genes Dev* **2000**, 14, (9), 1098-108.
130. Kashima, I.; Yamashita, A.; Izumi, N.; Kataoka, N.; Morishita, R.; Hoshino, S.; Ohno, M.; Dreyfuss, G.; Ohno, S., Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev* **2006**, 20, (3), 355-67.
131. Ohnishi, T.; Yamashita, A.; Kashima, I.; Schell, T.; Anders, K. R.; Grimson, A.; Hachiya, T.; Hentze, M. W.; Anderson, P.; Ohno, S., Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. *Mol Cell* **2003**, 12, (5), 1187-200.
132. Unterholzner, L.; Izaurrealde, E., SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol Cell* **2004**, 16, (4), 587-96.
133. Kuzmiak, H. A.; Maquat, L. E., Applying nonsense-mediated mRNA decay research to the clinic: progress and challenges. *Trends Mol Med* **2006**, 12, (7), 306-16.
134. Priolo, M.; Micale, L.; Augello, B.; Fusco, C.; Zucchetti, F.; Prontera, P.; Paduano, V.; Biamino, E.; Selicorni, A.; Mammi, C.; Lagana, C.; Zelante, L.; Merla, G., Absence of deletion and duplication of MLL2 and KDM6A genes in a large cohort of patients with Kabuki syndrome. *Mol Genet Metab* **2012**, 107, (3), 627-9.
135. Adzhubei, I. A.; Schmidt, S.; Peshkin, L.; Ramensky, V. E.; Gerasimova, A.; Bork, P.; Kondrashov, A. S.; Sunyaev, S. R., A method and server for predicting damaging missense mutations. *Nat Methods* **2010**, 7, (4), 248-9.
136. Tavtigian, S. V.; Deffenbaugh, A. M.; Yin, L.; Judkins, T.; Scholl, T.; Samollow, P. B.; de Silva, D.; Zharkikh, A.; Thomas, A., Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J Med Genet* **2006**, 43, (4), 295-305.
137. Choi, Y.; Sims, G. E.; Murphy, S.; Miller, J. R.; Chan, A. P., Predicting the functional effect of amino acid substitutions and indels. *PLoS One* **2012**, 7, (10), e46688.
138. Kumar, P.; Henikoff, S.; Ng, P. C., Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* **2009**, 4, (7), 1073-81.
139. Frederic, M. Y.; Lalande, M.; Boileau, C.; Hamroun, D.; Claustres, M.; Beroud, C.; Collod-Beroud, G., UMD-predictor, a new prediction tool for nucleotide substitution pathogenicity -- application to four genes: FBN1, FBN2, TGFBR1, and TGFBR2. *Hum Mutat* **2009**, 30, (6), 952-9.

140. Schwarz, J. M.; Cooper, D. N.; Schuelke, M.; Seelow, D., MutationTaster2: mutation prediction for the deep-sequencing age. *Nat Methods* **2014**, 11, (4), 361-2.
141. Desmet, F. O.; Hamroun, D.; Lalande, M.; Collod-Beroud, G.; Claustres, M.; Beroud, C., Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* **2009**, 37, (9), e67.
142. Reese, M. G.; Eeckman, F. H.; Kulp, D.; Haussler, D., Improved splice site detection in Genie. *J Comput Biol* **1997**, 4, (3), 311-23.
143. Hebsgaard, S. M.; Korning, P. G.; Tolstrup, N.; Engelbrecht, J.; Rouze, P.; Brunak, S., Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. *Nucleic Acids Res* **1996**, 24, (17), 3439-52.
144. Kurowski, M. A.; Bujnicki, J. M., GeneSilico protein structure prediction meta-server. *Nucleic Acids Res* **2003**, 31, (13), 3305-7.
145. Kallberg, M.; Margaryan, G.; Wang, S.; Ma, J.; Xu, J., RaptorX server: a resource for template-based protein structure modeling. *Methods Mol Biol* **2014**, 1137, 17-27.
146. Cheng, J.; Li, J.; Wang, Z.; Eickholt, J.; Deng, X., The MULTICOM toolbox for protein structure prediction. *BMC Bioinformatics* **2012**, 13, 65.
147. Webb, B.; Sali, A., Comparative Protein Structure Modeling Using MODELLER. *Curr Protoc Protein Sci* **2016**, 86, 2 9 1-2 9 37.
148. Rodrigues, J. P.; Levitt, M.; Chopra, G., KoBaMIN: a knowledge-based minimization web server for protein structure refinement. *Nucleic Acids Res* **2012**, 40, (Web Server issue), W323-8.
149. Benkert, P.; Kunzli, M.; Schwede, T., QMEAN server for protein model quality estimation. *Nucleic Acids Res* **2009**, 37, (Web Server issue), W510-4.
150. Lauck, F.; Smith, C. A.; Friedland, G. F.; Humphris, E. L.; Kortemme, T., RosettaBackrub--a web server for flexible backbone protein structure modeling and design. *Nucleic Acids Res* **2010**, 38, (Web Server issue), W569-75.
151. Schymkowitz, J.; Borg, J.; Stricher, F.; Nys, R.; Rousseau, F.; Serrano, L., The FoldX web server: an online force field. *Nucleic Acids Res* **2005**, 33, (Web Server issue), W382-8.
152. Dhar, S. S.; Lee, S. H.; Kan, P. Y.; Voigt, P.; Ma, L.; Shi, X.; Reinberg, D.; Lee, M. G., Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev* **2012**, 26, (24), 2749-62.
153. Thomas, M.; Lu, J. J.; Ge, Q.; Zhang, C.; Chen, J.; Klivanov, A. M., Full deacylation of polyethylenimine dramatically boosts its gene delivery efficiency and specificity to mouse lung. *Proc Natl Acad Sci U S A* **2005**, 102, (16), 5679-84.
154. Sermet-Gaudelus, I.; Renouil, M.; Fajac, A.; Bidou, L.; Parbaille, B.; Pierrot, S.; Davy, N.; Bismuth, E.; Reinert, P.; Lenoir, G.; Lesure, J. F.; Rousset, J. P.; Edelman, A., In vitro prediction of stop-codon suppression by intravenous gentamicin in patients with cystic fibrosis: a pilot study. *BMC Med* **2007**, 5, 5.
155. Bidou, L.; Hatin, I.; Perez, N.; Allamand, V.; Panthier, J. J.; Rousset, J. P., Premature stop codons involved in muscular dystrophies show a broad spectrum of readthrough efficiencies in response to gentamicin treatment. *Gene Ther* **2004**, 11, (7), 619-27.
156. Andersen, M. S.; Menazzi, S.; Brun, P.; Cocah, C.; Merla, G.; Solari, A., [Clinical diagnosis of Kabuki syndrome: phenotype and associated abnormalities in two new cases]. *Arch Argent Pediatr* **2014**, 112, (1), 26-32.
157. Cappuccio, G.; Rossi, A.; Fontana, P.; Acampora, E.; Avolio, V.; Merla, G.; Zelante, L.; Secinaro, A.; Andria, G.; Melis, D., Bronchial isomerism in a Kabuki syndrome patient with a novel mutation in MLL2 gene. *BMC Med Genet* **2014**, 15, 15.

158. Ejarque, I.; Uliana, V.; Forzano, F.; Marciano, C.; Merla, G.; Zelante, L.; Di Maria, E.; Faravelli, F., Is Hardikar syndrome distinct from Kabuki (Niikawa-Kuroki) syndrome? *Clin Genet* **2011**, *80*, (5), 493-6.
159. Makrythanasis, P.; van Bon, B. W.; Steehouwer, M.; Rodriguez-Santiago, B.; Simpson, M.; Dias, P.; Anderlid, B. M.; Arts, P.; Bhat, M.; Augello, B.; Biamino, E.; Bongers, E. M.; Del Campo, M.; Cordeiro, I.; Cueto-Gonzalez, A. M.; Cusco, I.; Deshpande, C.; Frysira, E.; Izatt, L.; Flores, R.; Galan, E.; Gener, B.; Gilissen, C.; Granneman, S. M.; Hoyer, J.; Yntema, H. G.; Kets, C. M.; Koolen, D. A.; Marcelis, C.; Medeira, A.; Micale, L.; Mohammed, S.; de Munnik, S. A.; Nordgren, A.; Psoni, S.; Reardon, W.; Revencu, N.; Roscioli, T.; Ruitkamp-Versteeg, M.; Santos, H. G.; Schoumans, J.; Schuurs-Hoeijmakers, J. H.; Silengo, M. C.; Toledo, L.; Vendrell, T.; van der Burgt, I.; van Lier, B.; Zweier, C.; Reymond, A.; Trembath, R. C.; Perez-Jurado, L.; Dupont, J.; de Vries, B. B.; Brunner, H. G.; Veltman, J. A.; Merla, G.; Antonarakis, S. E.; Hoischen, A., MLL2 mutation detection in 86 patients with Kabuki syndrome: a genotype-phenotype study. *Clin Genet* **2013**, *84*, (6), 539-45.
160. Micale, L.; Augello, B.; Maffeo, C.; Selicorni, A.; Zucchetti, F.; Fusco, C.; De Nittis, P.; Pellico, M. T.; Mandriani, B.; Fischetto, R.; Boccone, L.; Silengo, M.; Biamino, E.; Perria, C.; Sotgiu, S.; Serra, G.; Lapi, E.; Neri, M.; Ferlini, A.; Cavaliere, M. L.; Chiurazzi, P.; Monica, M. D.; Scarano, G.; Faravelli, F.; Ferrari, P.; Mazzanti, L.; Pilotta, A.; Patricelli, M. G.; Bedeschi, M. F.; Benedicenti, F.; Prontera, P.; Toschi, B.; Salviati, L.; Melis, D.; Di Battista, E.; Vancini, A.; Garavelli, L.; Zelante, L.; Merla, G., Molecular analysis, pathogenic mechanisms, and readthrough therapy on a large cohort of Kabuki syndrome patients. *Hum Mutat* **2014**, *35*, (7), 841-50.
161. Ratbi, I.; Fejjal, N.; Micale, L.; Augello, B.; Fusco, C.; Lyahyai, J.; Merla, G.; Sefiani, A., Report of the First Clinical Case of a Moroccan Kabuki Patient with a Novel MLL2 Mutation. *Mol Syndromol* **2013**, *4*, (3), 152-6.
162. Paulussen, A. D.; Stegmann, A. P.; Blok, M. J.; Tserpelis, D.; Pasma-Velter, C.; Detisch, Y.; Smeets, E. E.; Wagemans, A.; Schrandt, J. J.; van den Boogaard, M. J.; van der Smagt, J.; van Haeringen, A.; Stolte-Dijkstra, I.; Kerstjens-Frederikse, W. S.; Mancini, G. M.; Wessels, M. W.; Hennekam, R. C.; Vreeburg, M.; Geraedts, J.; de Ravel, T.; Fryns, J. P.; Smeets, H. J.; Devriendt, K.; Schrandt-Stumpel, C. T., MLL2 mutation spectrum in 45 patients with Kabuki syndrome. *Hum Mutat* **2011**, *32*, (2), E2018-25.
163. Li, Y.; Bogershausen, N.; Alanay, Y.; Simsek Kiper, P. O.; Plume, N.; Keupp, K.; Pohl, E.; Pawlik, B.; Rachwalski, M.; Milz, E.; Thoenes, M.; Albrecht, B.; Prott, E. C.; Lehmkuhler, M.; Demuth, S.; Utine, G. E.; Boduroglu, K.; Frankenbusch, K.; Borck, G.; Gillissen-Kaesbach, G.; Yigit, G.; Wiczorek, D.; Wollnik, B., A mutation screen in patients with Kabuki syndrome. *Hum Genet* **2011**, *130*, (6), 715-24.
164. Hannibal, M. C.; Buckingham, K. J.; Ng, S. B.; Ming, J. E.; Beck, A. E.; McMillin, M. J.; Gildersleeve, H. I.; Bigham, A. W.; Tabor, H. K.; Mefford, H. C.; Cook, J.; Yoshiura, K.; Matsumoto, T.; Matsumoto, N.; Miyake, N.; Tonoki, H.; Naritomi, K.; Kaname, T.; Nagai, T.; Ohashi, H.; Kurosawa, K.; Hou, J. W.; Ohta, T.; Liang, D.; Sudo, A.; Morris, C. A.; Banka, S.; Black, G. C.; Clayton-Smith, J.; Nickerson, D. A.; Zackai, E. H.; Shaikh, T. H.; Donnai, D.; Niikawa, N.; Shendure, J.; Bamshad, M. J., Spectrum of MLL2 (ALR) mutations in 110 cases of Kabuki syndrome. *Am J Med Genet A* **2011**, *155A*, (7), 1511-6.
165. Pasqualucci, L.; Trifonov, V.; Fabbri, G.; Ma, J.; Rossi, D.; Chiarenza, A.; Wells, V. A.; Grunn, A.; Messina, M.; Elliot, O.; Chan, J.; Bhagat, G.; Chadburn, A.; Gaidano, G.; Mullighan, C. G.; Rabadan, R.; Dalla-Favera, R., Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet* **2011**, *43*, (9), 830-7.

166. Banka, S.; Howard, E.; Bunstone, S.; Chandler, K.; Kerr, B.; Lachlan, K.; McKee, S.; Mehta, S.; Tavares, A.; Tolmie, J.; Donnai, D., MLL2 mosaic mutations and intragenic deletion-duplications in patients with Kabuki syndrome. *Clin Genet* **2012**.
167. Kokitsu-Nakata, N. M.; Petrin, A. L.; Heard, J. P.; Vendramini-Pittoli, S.; Henkle, L. E.; dos Santos, D. V.; Murray, J. C.; Richieri-Costa, A., Analysis of MLL2 gene in the first Brazilian family with Kabuki syndrome. *Am J Med Genet A* **2012**, 158A, (8), 2003-8.
168. Miyake, N.; Koshimizu, E.; Okamoto, N.; Mizuno, S.; Ogata, T.; Nagai, T.; Kosho, T.; Ohashi, H.; Kato, M.; Sasaki, G.; Mabe, H.; Watanabe, Y.; Yoshino, M.; Matsuishi, T.; Takanashi, J. I.; Shotelersuk, V.; Tekin, M.; Ochi, N.; Kubota, M.; Ito, N.; Ihara, K.; Hara, T.; Tonoki, H.; Ohta, T.; Saito, K.; Matsuo, M.; Urano, M.; Enokizono, T.; Sato, A.; Tanaka, H.; Ogawa, A.; Fujita, T.; Hiraki, Y.; Kitanaka, S.; Matsubara, Y.; Makita, T.; Taguri, M.; Nakashima, M.; Tsurusaki, Y.; Saito, H.; Yoshiura, K. I.; Matsumoto, N.; Niikawa, N., MLL2 and KDM6A mutations in patients with Kabuki syndrome. *Am J Med Genet A* **2013**.
169. Banka, S.; Lederer, D.; Benoit, V.; Jenkins, E.; Howard, E.; Bunstone, S.; Kerr, B.; McKee, S.; Chris Lloyd, I.; Shears, D.; Stewart, H.; White, S. M.; Savarirayan, R.; Mancini, G. M.; Beysen, D.; Cohn, R. D.; Grisart, B.; Maystadt, I.; Donnai, D., Novel KDM6A (UTX) mutations and a clinical and molecular review of the X-linked Kabuki syndrome (KS2). *Clin Genet* **2014**.
170. Cheon, C. K.; Sohn, Y. B.; Ko, J. M.; Lee, Y. J.; Song, J. S.; Moon, J. W.; Yang, B. K.; Ha, I. S.; Bae, E. J.; Jin, H. S.; Jeong, S. Y., Identification of KMT2D and KDM6A mutations by exome sequencing in Korean patients with Kabuki syndrome. *J Hum Genet* **2014**, 59, (6), 321-5.
171. Lin, J. L.; Lee, W. I.; Huang, J. L.; Chen, P. K.; Chan, K. C.; Lo, L. J.; You, Y. J.; Shih, Y. F.; Tseng, T. Y.; Wu, M. C., Immunologic assessment and KMT2D mutation detection in Kabuki syndrome. *Clin Genet* **2015**, 88, (3), 255-60.
172. Dentici, M. L.; Di Pede, A.; Lepri, F. R.; Gnazzo, M.; Lombardi, M. H.; Auriti, C.; Petrocchi, S.; Pisaneschi, E.; Bellacchio, E.; Capolino, R.; Braguglia, A.; Angioni, A.; Dotta, A.; Digilio, M. C.; Dallapiccola, B., Kabuki syndrome: clinical and molecular diagnosis in the first year of life. *Arch Dis Child* **2015**, 100, (2), 158-64.
173. Morgan, A. T.; Mei, C.; Da Costa, A.; Fifer, J.; Lederer, D.; Benoit, V.; McMillin, M. J.; Buckingham, K. J.; Bamshad, M. J.; Pope, K.; White, S. M., Speech and language in a genotyped cohort of individuals with Kabuki syndrome. *Am J Med Genet A* **2015**, 167, (7), 1483-92.
174. Liu, S.; Hong, X.; Shen, C.; Shi, Q.; Wang, J.; Xiong, F.; Qiu, Z., Kabuki syndrome: a Chinese case series and systematic review of the spectrum of mutations. *BMC Med Genet* **2015**, 16, 26.
175. Schott, D. A.; Blok, M. J.; Gerver, W. J.; Devriendt, K.; Zimmermann, L. J.; Stumpel, C. T., Growth pattern in Kabuki syndrome with a KMT2D mutation. *Am J Med Genet A* **2016**, 170, (12), 3172-3179.
176. Lederer, D.; Shears, D.; Benoit, V.; Verellen-Dumoulin, C.; Maystadt, I., A three generation X-linked family with Kabuki syndrome phenotype and a frameshift mutation in KDM6A. *Am J Med Genet A* **2014**, 164A, (5), 1289-92.
177. Brichta, L.; Garbes, L.; Jędrzejowska, M.; Grellescheid, S. N.; Holker, I.; Zimmermann, K.; Wirth, B., Nonsense-mediated messenger RNA decay of survival motor neuron 1 causes spinal muscular atrophy. *Hum Genet* **2008**, 123, (2), 141-53.

178. Manuvakhova, M.; Keeling, K.; Bedwell, D. M., Aminoglycoside antibiotics mediate context-dependent suppression of termination codons in a mammalian translation system. *RNA* **2000**, 6, (7), 1044-55.
179. Floquet, C.; Hatin, I.; Rousset, J. P.; Bidou, L., Statistical analysis of readthrough levels for nonsense mutations in mammalian cells reveals a major determinant of response to gentamicin. *PLoS Genet* **2012**, 8, (3), e1002608.
180. Banka, S.; Veeramachaneni, R.; Reardon, W.; Howard, E.; Bunstone, S.; Ragge, N.; Parker, M. J.; Crow, Y. J.; Kerr, B.; Kingston, H.; Metcalfe, K.; Chandler, K.; Magee, A.; Stewart, F.; McConnell, V. P.; Donnelly, D. E.; Berland, S.; Houge, G.; Morton, J. E.; Oley, C.; Revencu, N.; Park, S. M.; Davies, S. J.; Fry, A. E.; Lynch, S. A.; Gill, H.; Schweiger, S.; Lam, W. W.; Tolmie, J.; Mohammed, S. N.; Hobson, E.; Smith, A.; Blyth, M.; Bennett, C.; Vasudevan, P. C.; Garcia-Minaur, S.; Henderson, A.; Goodship, J.; Wright, M. J.; Fisher, R.; Gibbons, R.; Price, S. M.; D, C. d. S.; Temple, I. K.; Collins, A. L.; Lachlan, K.; Elmslie, F.; McEntagart, M.; Castle, B.; Clayton-Smith, J.; Black, G. C.; Donnai, D., How genetically heterogeneous is Kabuki syndrome?: MLL2 testing in 116 patients, review and analyses of mutation and phenotypic spectrum. *Eur J Hum Genet* **2012**, 20, (4), 381-8.
181. Kim, J. H.; Sharma, A.; Dhar, S. S.; Lee, S. H.; Gu, B.; Chan, C. H.; Lin, H. K.; Lee, M. G., UTX and MLL4 coordinately regulate transcriptional programs for cell proliferation and invasiveness in breast cancer cells. *Cancer Res* **2014**, 74, (6), 1705-17.
182. Shinsky, S. A.; Hu, M.; Vought, V. E.; Ng, S. B.; Bamshad, M. J.; Shendure, J.; Cosgrove, M. S., A non-active-site SET domain surface crucial for the interaction of MLL1 and the RbBP5/Ash2L heterodimer within MLL family core complexes. *J Mol Biol* **2014**, 426, (12), 2283-99.
183. Lange, L.; Pagnamenta, A. T.; Lise, S.; Clasper, S.; Stewart, H.; Akha, E. S.; Quaghebeur, G.; Knight, S. J.; Keays, D. A.; Taylor, J. C.; Kini, U., A de novo frameshift in HNRNPK causing a Kabuki-like syndrome with nodular heterotopia. *Clin Genet* **2016**, 90, (3), 258-62.
184. Au, P. Y. B.; You, J.; Caluseriu, O.; Schwartzentruber, J.; Majewski, J.; Bernier, F. P.; Ferguson, M.; Valle, D.; Parboosingh, J. S.; Sobreira, N.; Innes, A. M.; Kline, A. D., GeneMatcher aids in the identification of a new malformation syndrome with intellectual disability, unique facial dysmorphisms, and skeletal and connective tissue abnormalities caused by de novo variants in HNRNPK. *Hum Mutat* **2015**, 36, (10), 1009-1014.