

ORIGINAL ARTICLE

Dissecting *KMT2D* missense mutations in Kabuki syndrome patients

Dario Cocciadiferro^{1,2,†}, Bartolomeo Augello¹, Pasquelena De Nittis³, Jiyuan Zhang⁴, Barbara Mandriani⁵, Natascia Malerba^{1,2}, Gabriella M. Squeo¹, Alessandro Romano⁶, Barbara Piccinni⁷, Tiziano Verri⁷, Lucia Micale¹, Laura Pasqualucci⁴ and Giuseppe Merla^{1,*}

¹Division of Medical Genetics, IRCCS Casa Sollievo della Sofferenza Hospital, San Giovanni Rotondo, Italy,

²PhD Program in Experimental and Regenerative Medicine, Faculty of Medicine, University of Foggia, Italy,

³Center for Integrative Genomics, University of Lausanne, CH-1015 Lausanne, Switzerland, ⁴Department of Pathology and Cell Biology, Institute for Cancer Genetics, Columbia University, New York, NY, USA,

⁵Telethon Institute of Genetics and Medicine, TIGEM, Pozzuoli, Naples, Italy, ⁶Institute of Experimental Neurology, Division of Neuroscience, San Raffaele Scientific Institute, Milan, Italy and ⁷Department of Biological and Environmental Sciences and Technologies, University of Salento, Lecce, Italy

*To whom correspondence should be addressed at: Division of Medical Genetics, IRCCS Casa Sollievo della Sofferenza, viale Cappuccini, 71013, San Giovanni Rotondo, Foggia, Italy. Tel: +39 0882416350; Fax: +39 0882411616; Email: g.merla@operapadrepio.it

Abstract

Kabuki syndrome is a rare autosomal dominant condition characterized by facial features, various organs malformations, postnatal growth deficiency and intellectual disability. The discovery of frequent germline mutations in the histone methyltransferase *KMT2D* and the demethylase *KDM6A* revealed a causative role for histone modifiers in this disease. However, the role of missense mutations has remained unexplored. Here, we expanded the mutation spectrum of *KMT2D* and *KDM6A* in KS by identifying 37 new *KMT2D* sequence variants. Moreover, we functionally dissected 14 *KMT2D* missense variants, by investigating their impact on the protein enzymatic activity and the binding of members of the WRAD complex. We demonstrate impaired H3K4 methyltransferase activity in 9 of the 14 mutant alleles and show that this reduced activity is due in part to disruption of protein complex formation. These findings have relevant implications for diagnostic and counseling purposes in this disease.

Introduction

Kabuki syndrome (KS, MIM #147920, MIM #300867) is a rare autosomal dominant condition 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 intellectual disability and postnatal growth deficiency (1,2). Additional findings include congenital heart defects, genitourinary anomalies, cleft lip and/or palate, susceptibility to infections, gastrointestinal abnormalities,

[†]Present address: Laboratory of Medical Genetics, Bambino Gesù Children's Hospital, IRCCS, Rome, Italy.

Received: April 24, 2018. Revised: May 30, 2018. Accepted: June 21, 2018

© The Author(s) 2018. Published by Oxford University Press. All rights reserved.

For permissions, please email: journals.permissions@oup.com

ophthalmologic defects, ptosis and strabismus, dental anomalies (including widely spaced teeth and hypodontia), ear pits, visceral abnormalities and premature thelarche (3). In 2010, whole-exome sequencing successfully identified heterozygous loss of function mutations in the *KMT2D* gene (MIM #602113, NM_003482.3, also known as *MLL2* and *MLL4*) as the major cause of KS (4). *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 (5–7). 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. In mammalian cells, *KMT2D* functions as a major histone H3K4 mono-/di-/tri-methyltransferase (8–15) that is required for the epigenetic control of active chromatin states in a tissue-specific manner (16).

In addition to *KMT2D* mutations, a subset of KS individuals has been identified with either point mutations or microdeletions encompassing the X-linked gene, *KDM6A* (MIM #300128, NM_021140.3, also known as *UTX*) (17–19), which encodes for a Histone H3 lysine-27 demethylase. *KDM6A* plays a crucial role in chromatin remodeling (20,21) and interacts with *KMT2D* in a conserved SET1-like complex (16).

More than 600 *KMT2D* mutations have been identified so far in KS patients; roughly 84% of them are truncating events (22) including non-sense, indels, small duplications, splice-site, and frameshift mutations, while the remaining 16% are represented by missense variants whose pathogenicity has not been investigated. As a consequence, the interpretation of missense mutations has remained a challenging problem in genetic diagnosis and counseling.

Here we surveyed a cohort of 505 KS patients to expand the mutation spectrum of *KMT2D* and *KDM6A*, and investigated the functional impact of missense mutations in the pathogenesis of the disease.

Results

Mutation screening of *KMT2D* and *KDM6A*

To expand the spectrum of mutations targeting *KMT2D* and *KDM6A* in KS, we integrated our mutation database with mutational screening of 202 newly diagnosed KS patients for a total of 505 cases, by using Sanger sequencing and/or multiplex ligation-dependent probe amplification (MLPA) of the *KMT2D* coding sequence, followed by *KDM6A* analysis in patients resulted as *KMT2D*-negative. We identified a total of 208 *KMT2D* variants distributed in 196/505 (39%) patients, including 37 that have not been described before (Table 1). These included 54 non-sense mutations (26%), 59 frameshift mutations (28%), 69 missense mutations (33%), 13 splice site variants (6%), 12 indels (6%) and 1 gross deletion (Table 1). Missense mutations were distributed across the entire length of the *KMT2D* gene (Fig. 1A) and were represented by 8 missense variants (11%) localized within the PHD 1–6 domains, one change (1%) in the Coiled Coil/Poly Q region, 25 variants (36%) localized within the C-terminal of the protein (amino acid 4507–5537) and 36 (52%) variants localized outside of known domains and/or in uncharacterized portions of the protein. Among the *KMT2D* variants identified, 66 occurred *de novo* and 32 were inherited from an apparently asymptomatic parent, whereas for the remaining 110 variants we had no access to parental DNA.

Moreover, we identified 14 *KDM6A* variants; 12 of those were predicted to be pathogenic based on publicly available algorithms (see Materials and Methods) and 7 were never described previously. Seven of the variants were *de novo*, while for the remaining ones the inheritance was unknown. Thus, 210/505 (41%) KS patients in our cohort carried genetic alterations in one of these two genes; the underlying event in the remaining 295 cases remains unknown (see Discussion).

Pathogenic assessment of missense mutations by bioinformatics tools

Analogous to previous studies, most *KMT2D* variants in KS are inactivating truncating events that render the protein functionally defective due to the loss of the catalytic SET activity. However, ~30% (69/208) of the mutations found in our data set, and 100 of 621 (16.1%) mutations from a recently published mutational analysis review (22), are in-frame amino acid changes that affect various residues along the *KMT2D* protein. To begin to elucidate the functional consequences of *KMT2D* missense mutations in KS, we first applied published bioinformatics tools to the 58 unique missense variants identified in the 69 KS patients (see Materials and Methods). These prediction algorithms classified 16/58 variants (identified in 20/69 patients) as pathogenic or likely pathogenic, while 24/58 variants (30/69 patients) were scored as likely benign, and 18/58 (19/69 patients) as variants of uncertain significance (VOUS, Table 2).

To further assess how *KMT2D* missense variants may affect protein function/activity, a combination of structure-based methods was employed. Since *KMT2D* is a multi-domain protein too large to be accurately modeled using computational methods, we aimed to predict and analyze the individual structures of single *KMT2D* domains. Three-dimensional models were obtained for the ZF-7, FYR and SET domains (QMEAN scores of 0.63, 0.68 and 0.73, respectively) of the human *KMT2D* (see Materials and Methods for details). Moreover, a crystal structure at 1.4 Å resolution of the WIN domain was retrieved from the Protein Data Bank. Eight missense variants were analyzed to test their effects on protein stability ($\Delta\Delta G$) and change in total charge (ΔCharge). The analysis showed that all eight variants target highly conserved regions/residues within these domains (Supplementary Material, Fig. S1) and are expected to significantly alter their structure (Supplementary Material, Fig. S2). In particular, with one exception (p.H5059P, which falls on the N-terminal residue of the predicted ZF-PHD7 and can be poorly analyzed), all variants were predicted to alter the total charge of the domain and/or the protein stability (Supplementary Material, Fig. S3). These data suggest that the missense variants located in the ZF-PHD7, FYR, WIN and SET domains affect the normal structure of the *KMT2D* protein and may thus potentially alter/impair the protein function.

Missense variants impair *KMT2D* methyltransferase activity

To experimentally test the functional impact of KS-associated *KMT2D* missense variants, we generated FLAG-tagged versions of 14 representative *KMT2D* mutant alleles that harbor amino acid changes in both functional N-terminal and C-terminal domains of the protein including PHD 4–5–6–7, FYRN, WIN and SET domains, using the FUSION-*KMT2D* construct as template (see Materials and Methods) (Fig. 1B).

Table 1. KMT2D and KDM6A variants identified in our cohort

ID	Inheritance	Exon/intron	Variant	AA change	Reference	ACMG classification
KMT2D						
Non-sense						
KB49	NA	ex 5	c.669T>G	p.(Tyr223*)	(46)	P
KB343	NA	ex 8	c.1016G>A	p.(Trp339*)	This study	P
KB35	NA	ex 10	c.1921G>T	p.(Glu641*)	(46)	P
KB33	NA	ex 16	c.4419G>A	p.(Trp1473*)	(46)	P
KB63	NA	ex 19	c.4895delC	p.(Ser1632*)	(46)	P
KB317	NA	ex 22	c.5212G>T	p.(Glu1738*)	(19)	P
KB336	De novo	ex 22	c.5269C>T	p.(Arg1757*)	(18)	P
KB262	NA	ex 26	c.5674C>T	p.(Gln1892*)	(19)	P
KB429	NA	ex 26	c.5707C>T	p.(Arg1903*)	(18,22,73,74)	P
KB26	NA	ex 31	c.6295C>T	p.(Arg2099*)	(4,22,46)	P
KB502	De novo	ex 31	c.7228C>T	p.(Arg2410*)	(9,73,75,76)	P
KB66	NA	ex 31	c.7246C>T	p.(Gln2416*)	(46)	P
KB59	NA	ex 31	c.7903C>T	p.(Arg2635*)	(22,46)	P
KB153	De novo	ex 31	c.7903C>T	p.(Arg2635*)	(22,46)	P
KB226	De novo	ex 31	c.7903C>T	p.(Arg2635*)	(22,46)	P
KB338	De novo	ex 31	c.7933C>T	p.(Arg2645*)	(77)	P
KB198	De novo	ex 31	c.7936G>T	p.(Glu2646*)	(19)	P
KB352	NA	ex 32	c.8227C>T	p.(Gln2743*)	This study	P
KB323	NA	ex 33	c.8311C>T	p.(Arg2771*)	(76,77)	P
KB289	NA	ex 34	c.8743C>T	p.(Arg2915*)	(22,77-79)	P
KB422	De novo	ex 34	c.9396C>A	p.(Cys3132*)	This study	P
KB186	De novo	ex 34	c.9961C>T	p.(Arg3321*)	(4,75,76,79)	P
KB56	De novo	ex 34	c.10135C>T	p.(Gln3379*)	(46)	P
KB168	De novo	ex 39	c.10750C>T	p.(Gln3584*)	(19)	P
KB46	De novo	ex 39	c.10841C>G	p.(Ser3614*)	(46)	P
KB41	NA	ex 39	c.11119C>T	p.(Arg3707*)	(46)	P
KB44	NA	ex 39	c.11119C>T	p.(Arg3707*)	(46)	P
KB42	De novo	ex 39	c.11269C>T	p.(Gln3757*)	(22,46)	P
KB25	NA	ex 39	c.11434 C>T	p.(Gln3812*)	(46)	P
KB244	De novo	ex 39	c.11674 C>T	p.(Gln3892*)	(76)	P
KB178	NA	ex 39	c.11704C>T	p.(Gln3902*)	(19)	P
KB425	De novo	ex 39	c.11731C>T	p.(Gln3911*)	This study	P
KB461 ^a	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*)	(19)	P
KB358	NA	ex 39	c.11944C>T	p.(Arg3982*)	(18,22,77)	P
KB40	NA	ex 39	c.12274C>T	p.(Gln4092*)	(18,46)	P
KB114	De novo	ex 39	c.12274C>T	p.(Gln4092*)	(18,46)	P
KB65	NA	ex 39	c.12076C>T	p.(Gln4026*)	(46)	P
KB333	NA	ex 39	c.12703C>T	p.(Gln4235*)	(4)	P
KB410	NA	39	c.12760C>T	p.(Gln4254*)	(22)	P

(continued)

Table 1. Continued

ID	Inheritance	Exon/intron	Variant	AA change	Reference	ACMG classification
KB82	De novo	ex 39	c.12844C>T	p.(Arg4282*)	(19)	P
KB350	De novo	ex 39	c.12844C>T	p.(Arg4282*)	(19)	P
KB189	De novo	ex 39	c.12955A>T	p.(Arg4319*)	(19,22)	P
KB183	De novo	ex 39	c.13450C>T	p.(Arg4484*)	(9,22,74,77)	P
KB450	NA	ex 39	c.13450C>T	p.(Arg4484*)	(9,22,74,77)	P
KB175	De novo	ex 39	c.13507C>T	p.(Gln4503*)	(19)	P
KB73	De novo	ex 40	c.13666A>T	p.(Lys4556*)	(46)	P
KB83	NA	ex 48	c.15022G>T	p.(Glu5008*)	(19)	P
KB377	NA	ex 48	c.15061C>T	p.(Arg5021*)	(18,76)	P
KB45	NA	ex 48	c.15079C>T	p.(Arg5027*)	(22,46,77)	P
KB72	NA	ex 48	c.15079C>T	p.(Arg5027*)	(22,46,77)	P
KB362	NA	ex 50	c.16018C>T	p.(Arg5340*)	(77)	P
KB130	NA	ex 52	c.16360C>T	p.(Arg5454*)	(4,22,75,77)	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)	(45)	P
KB8	De novo	ex 5	c.588delC	p.(Cys197Alafs*11)	(46)	P
KB58	NA	ex 6	c.705delA	p.(Glu237Serfs*24)	(46)	P
KB57	NA	ex 8	c.1035_1036delCT	p.(Cys346Serfs*17)	(46)	P
KB89	NA	ex 10	c.1345_1346delICT	p.(Leu449Valfs*5)	(46,74)	P
KB156	De novo	ex 10	c.1503dupT	p.(Pro502Serfs*7)	(19)	P
KB116	NA	ex 10	c.1634delT	p.(Leu545Argfs*385)	(76)	P
KB349	NA	ex 10	c.1634delT	p.(Leu545Argfs*385)	(76)	P
KB545	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)	(46)	P
KB203	NA	ex 11	c.3161_3171delCGTTGAGTCCC	p.(Pro1054Hisfs*10)	(18,19,43)	P
KB309	NA	ex 11	c.3730delG	p.(Val1244Serfs*86)	(19)	P
KB142	De novo	ex 13	c.4021delG	p.(Val1341Leufs*35)	(19)	P
KB311	NA	ex 14	c.4135_4136delAAT	p.(Met1379Valfs*52)	(19,22,80)	P
KB524	NA	14	c.4135_4136delAAT	p.(Met1379Valfs*52)	(19,22,80)	P
KB188	De novo	ex 16	c.4454delC	p.(Pro1485Leufs*21)	(19)	P
KB159	NA	ex 19	c.4896_4905delAGATCCCCTT	p.(Asp1633Alafs*86)	(19)	P
KB443 ^a	De novo	ex 25	c.5575delG	p.(Asp1859Thrfs*17)	This study	P
KB3	NA	ex 26	c.5652dup	p.(Lys1885Glnfs*18)	(45)	P
KB84	NA	ex 26	c.5779delC	p.(Gln1927Lysfs120*)	(46)	P
KB146	De novo	ex 27	c.5857delC	p.(Leu1953Trpfs*94)	(19)	P
KB208	NA	ex 28	c.5954delC	p.(Thr1985Lysfs*62)	(19)	P
KB221	De novo	ex 29	c.6149_6150delIGA	p.(Arg2050Lysfs*6)	(19)	P
KB525	NA	30	c.6212_6213delIAC	p.(His2071Profs*10)	This study	P
KB152	De novo	ex 31	c.6583delA	p.(Thr2195Profs*69)	(19)	P

(continued)

Table 1. Continued

ID	Inheritance	Exon/intron	Variant	AA change	Reference	ACMG classification
KB267	NA	ex 31	c.6594delC	p.(Tyr2199)Leufs*65)	(75)	P
KB79	De novo	ex 31	c.6595delT	p.(Tyr2199)Leufs*65)	(4,20,22,45,46,75-77,81)	P
KB102	De novo	ex 31	c.6595delT	p.(Tyr2199)Leufs*65)	(4,20,22,45,46,75-77,81)	P
KB342	NA	ex 31	c.6595delT	p.(Tyr2199)Leufs*65)	(4,20,22,45,46,75-77,81)	P
KB67	De novo	ex 31	c.6638_6641delGCGC	p.(Gly2213)Alafs*50)	(46)	P
KB176	NA	ex 31	c.6738delA	p.(Lys2246)Asnfs*18)	(19)	P
KB253	NA	ex 31	c.6794delG	p.(Gly2265)Gluufs*21)	(19,22)	P
KB278	NA	ex 31	c.7481dupT	p.(Ala2496)Serfs*10)	(19,79)	P
KB313	De novo	ex 32	c.8196delG	p.(Ser2733)Valfs*24)	(19)	P
KB80	NA	ex 33	c.8273delG	p.(Gly2758)Alafs*29)	(46)	P
KB243	De novo	ex 34	c.8430_8431insAA	p.(Gln2811)Asnfs*41)	(19)	P
KB182	NA	ex 34	c.9203delA	p.(Gln3068)Glyfs*3)	(19)	P
KB30	NA	ex 38	c.10606delC	p.(Arg3536)Alafs*122)	(46)	P
KB101	De novo	ex 39	c.11066_11078delCT GGATCCCTGGC	p.(Ala3689)Valfs*56)	(46)	P
KB504	De novo	ex 39	c.11093dupG	p.(Phe3699)Leufs*14)	This study	P
KB495	De novo	ex 39	c.11715delG	p.(Gln3905)Hisfs*74)	This study	P
KB172	NA	ex 39	c.12647delC	p.(Pro4216)Leufs*62)	(19)	P
KB192	De novo	ex 39	c.12966delA	p.(Gln4322)Hisfs*62)	(19)	P
KB54	NA	ex 39	c.13129dupT	p.(Trp4377)Leufs*33)	(46)	P
KB121	De novo	ex 39	c.13277dupT	p.(Ala4428)Serfs*59)	(19)	P
KB540	NA	41	c.13780delG	p.(Ala4594)Profs*23)	(22)	P
KB123	De novo	ex 42	c.13884dupC	p.(Thr4629)Hisf*18)	(19)	P
KB481	NA	ex 42	c.13895dupC	p.(Ser4633)Leufs*14)	This study	P
KB197	De novo	ex 47	c.14592dupG	p.(Pro4865)Alafs*48)	(19)	P
KB125	NA	ex 48	c.15031delG	p.(Glu5011)Serfs*40)	(19)	P
KB16	De novo	ex 48	c.15374dupT	p.(Phe5126)Leufs*12)	(45)	P
KB535	NA	50	c.16043_16044delAC	p.(His5348)Leufs*14)	This study	P
KB64	NA	ex 53	c.16438_16441delAACT	p.(Asn5480)Val*6)	(75)	P
KB533	NA	ex 53	c.16469_16470delAA	p.(Lys5490)Argfs*21)	(46)	P
KB533	NA	53	c.16469_16470delAA	p.(Lys5490)Argfs*21)	(46)	P
Missense						
KB21 ^a	inherited M	ex 3	c.346T>C	p.(Ser116)Pro)	(19)	LB
KB21 ^{a,b}	inherited M	ex 4	c.510G>C	p.Gln170His	(19,45)	P
KB256	NA	ex 5	c.626C>T	p.(Thr209)Ile)	(19)	VOUS
KB458	NA	ex 8	c.1076G>C	p.(Arg359)Pro)	This study	VOUS
KB269	inherited M	ex 10	c.1940C>A	p.(Pro647)Gln)	(19,45,77)	LB
KB126	NA	ex 10	c.2074C>A	p.(Pro692)Thr)	(77)	LB
KB374 ^c	inherited M	ex 10	c.2074C>A	p.(Pro692)Thr)	(77)	LB
KB487	NA	ex 10	c.2654C>T	p.(Pro885)Leu)	This study	VOUS
KB370 ^a	Inherited P-M	ex 11	c.2837C>G	p.(Ala946)Gly)	This study	LB
KB215	Inherited M	ex 11	c.3392C>T	p.(Pro1131)Leu)	(19)	LB
KB341	Inherited M	ex 11	c.3392C>T	p.(Pro1131)Leu)	(19)	LB
KB222	Inherited P	ex 11	c.3572C>T	p.(Pro1191)Leu)	(19)	LB

(continued)

Table 1. Continued

ID	Inheritance	Exon/intron	Variant	AA change	Reference	ACMG classification
KB32	Inherited P	ex 11	c.3773G>A	p.(Arg1258Gln)	(46)	LB
KB307	De novo	ex 14	c.4171G>A	p.(Glu1391Lys)	(19,22)	LP
KB28 ^a	Inherited M	ex 15	c.4249A>G	p.(Met1417Val)	(46)	LB
KB28 ^a	Inherited M	ex 15	c.4252C>A	p.(Leu1418Met)	(46)	LB
KB174	Inherited M	ex 15	c.4283T>C	p.(Ile1428Thr)	(19)	LB
KB138	NA	ex 16	c.4427C>G	p.(Ser1476Cys)	(19)	VOUS
KB34	Inherited P	ex 16	c.4565A>G	p.(Gln1522Arg)	(46)	LB
KB535	NA	25	c.5549G>A	p.(Gly1850Asp)	This study	LB
KB119	Inherited M	ex 31	c.6638G>A	p.(Gly2213Asp)	(19)	LB
KB204 ^c	Inherited M	ex 31	c.6638G>A	p.(Gly2213Asp)	(19)	LB
KB330 ^a	NA	ex 31	c.6733C>G	p.(Leu2245Val)	This study	VOUS
KB326 ^c	Inherited M	ex 31	c.6811C>T	p.(Pro2271Ser)	(19)	LB
KB107 ^a	NA	ex 31	c.6970C>A	p.(Pro2324Thr)	(19)	VOUS
KB430	NA	ex 31	c.7378C>T	p.(Arg2460Cys)	(77)	LB
KB122	Inherited M	ex 31	c.7829T>C	p.(Leu2610Pro)	(19,73)	LB
KB287	Inherited M	ex 31	c.7829T>C	p.(Leu2610Pro)	(19,73)	LB
KB27	NA	ex 34	c.8521C>A	p.(Pro2841Thr)	(46)	VOUS
KB330 ^a	NA	ex 34	c.8774C>T	p.(Ala2925Val)	This study	VOUS
KB443 ^a	Inherited M	ex 34	c.9971G>T	p.(Gly3324Val)	This study	LB
KB326 ^c	Inherited P	ex 34	c.10192A>G	p.(Met3398Val)	(19)	LB
KB357	Inherited P	ex 34	c.10192A>G	p.(Met3398Val)	(19)	LB
KB297	NA	ex 35	c.10256A>G	p.(Asp3419Gly)	(77)	LB
KB378	NA	ex 35	c.10256A>G	p.(Asp3419Gly)	(77)	LB
KB292	De novo	ex 37	c.10499G>T	p.(Gly3500Val)	(19)	LP
KB86 ^a	NA	ex 39	c.10966C>T	p.(Arg3656Cys)	(19)	VOUS
KB374 ^c	Inherited P	ex 39	c.11380C>T	p.(Pro3794Ser)	This study	LB
KB293	Inherited P	ex 39	c.11794C>G	p.(Gln3932Glu)	(19)	LB
KB204 ^c	Inherited P	ex 39	c.12070A>G	p.(Lys4024Glu)	(19)	LB
KB385	NA	ex 39	c.12302A>C	p.(Gln4101Pro)	This study	VOUS
KB247	NA	ex 39	c.12485G>A	p.(Arg4162Gln)	(19)	VOUS
KB107 ^a	NA	ex 39	c.12488C>T	p.(Pro4163Leu)	(19)	VOUS
KB512	Inherited M	ex 39	c.13256C>T	p.(Pro4419Leu)	(19)	LB
KB86 ^a	NA	45	c.14381A>G	p.(Lys4794Arg)	This study	VOUS
KB38 ^a	De novo	ex 48	c.14893G>A	p.(Ala4965Thr)	(19)	VOUS
KB38 ^a	De novo	ex 48	c.15084C>G	p.(Asp5028Glu)	(46)	LP
KB154	NA	ex 48	c.15088C>T	p.(Arg5030Cys)	(18,45)	VOUS
KB185	De novo	ex 48	c.15088C>T	p.(Arg5030Cys)	(18,45)	VOUS
KB423	Inherited P	ex 48	c.15089G>A	p.(Arg5030His)	This study	LB
KB38 ^a	De novo	ex 48	c.15100T>G	p.(Phe5034Val)	(46)	LP
KB76	De novo	ex 48	c.15176A>C	p.(His5059Pro)	(46)	LP
KB129	NA	ex 48	c.15292A>C	p.(Thr5098Pro)	(19)	VOUS
KB462	De novo	ex 48	c.15310T>C	p.(Cys5104Arg)	This study	LP
KB171	Inherited M	ex 48	c.15565G>A	p.(Gly5189Arg)	(18,22,46)	VOUS
KB264	NA	ex 48	c.15640C>T	p.(Arg5214Cys)	(22,45,75,76)	LP

(continued)

Table 1. Continued

ID	Inheritance	Exon/intron	Variant	AA change	Reference	ACMG classification
KB376	NA	ex 48	c.15640C>T	p.(Arg5214Cys)	(22,45,75,76)	LP
KB24	De novo	ex 48	c.15641G>A	p.(Arg5214His)	(4)	LP
KB219	NA	ex 48	c.15641G>A	p.(Arg5214His)	(4,75)	LP
KB408	De novo	ex 48	c.15641G>A	p.(Arg5214His)	(4,75)	LP
KB109	De novo	ex 48	c.15649T>C	p.(Trp5217Arg)	(19)	LP
KB17	De novo	ex 50	c.16019G>A	p.(Arg5340Gln)	(22,46)	LP
KB169	De novo	ex 51	c.16273G>A	p.(Glu5425Lys)	(19,22,78)	P
KB90	NA	ex 51	c.16295G>A	p.(Arg5432Gln)	(22,82,83)	LP
KB467	NA	ex 51	c.16295G>A	p.(Arg5432Gln)	(22,82,83)	LP
KB480	NA	ex 52	c.16385A>G	p.(Arg5432Gln)	This study	VOUS
KB177	De novo	ex 52	c.16412G>T	p.(Arg5471Met)	(19)	LP
KB489	NA	ex 53	c.16498C>T	p.(Arg5500Trp)	(78)	P
KB120	De novo	ex 54	c.16528T>G	p.(Tyr5510Asp)	(19)	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)	(19)	VOUS
KB370 ^d	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 ^{a,d}	NA	ex 39	c.11220_11222dup	p.(Gln3745dup)	This study	LB
KB281	Inherited M	ex 39	c.11714_11716dup	p.(Gln3905dup)	(19)	LB
KB71	Inherited M	ex 39	c.11819_11836dup	p.(Leu3940_Gln3945dup)	(46)	LB
KB227	Inherited P	ex 39	c.11843_11860del	p.(L3948_Q3953del)	(19)	LB
KB228	Inherited P	ex 39	c.11854_11874dup	p.(Q3952_Q3958dup)	(19)	LB
KB77	NA	ex 48	c.15163_15168dup	p.(Asp5055_Leu5056dup)	(18,46)	VOUS
KB403	De novo	ex 53	c.16489_16491del	p.(Ile5497del)	(22,46,75,76)	LP
KB53	NA	ex 53	c.16489_16491del	p.(Ile5497del)	(22,46,75,76)	LP
Splice site						
KB286	De novo	int 2-3	c.177-2A>C	r.177_400del224; p.S59Rfs*86	(19)	P
KB31	NA	int 3-4	c.400 + 1 G>A	r.177_400del224; p.Ser59Argfs*86	(46)	P
KB20	De novo	int 3-5	c.401-3 A>G	r.400_401insAG; p.Gly134Glnfs*75	(46)	P
KB442	NA	int 6-7	c.840-6delC	r.?	This study	VOUS
KB519	NA	int 13-14	c.4132-2A>G	r.?	This study	P
KB529	NA	int 16-17	c.4584-6C>G	r.?	This study	VOUS
KB210	De novo	int 17-18	c.4693 + 1G>A	r.4681_4693del13; p.Val1561Argfssplice*11	(48)	P
KB29	NA	int 42-43	c.13999 + 5 G>A	r.13840_13999del160; p.Asn4614Ilefs*5	(46,77)	P
KB290	De novo	int 47-48	c.14643 + 1G>A	r.14644_14875del232; p.Glu4882Profs*36	(19)	P
KB195	De novo	int 47-48	c.14644-3C>G	r.14644_14875del232; p.Gln4882Profs*36	(19)	P
KB7	De novo	int 44-45	c.14252-6_14252-5insGAAA	r.14252_14382del131; p.Val4751_Glu1fssplice*22	(45)	P
KB360	NA	int 47-48	c.14644-2A>T	r.?	(77)	P
KB496	NA	int 53-54	c.16520_16521 + 1delAGG	r.?	This study	P

(continued)

Table 1. Continued

ID	Inheritance	Exon/intron	Variant	AA change	Reference	ACMG classification
Gross deletion						
KB43 ^e	NA	ex 48–51	c.15785-238_16168delins	p.?	(19)	P
KDM6A						
Non-sense						
KB215 ^f	De novo	ex 6	c.514C>T	p.(Arg172*)	(19,22,84)	P
KB341	De novo	ex 6	c.514C>T	p.(Arg172*)	(22,84)	P
Frameshift						
KB39	NA	ex 16	c.1846_1849del	p.(Thr616tyrfs*8)	(19)	P
KB141	NA	ex 17	c.2118_2119ins	p.(G707Hfs*13)	This study	P
KB434	NA	ex 17	c.2515_2518del	p.(Asn839Valfs*27)	(85)	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)	(19)	P
KB380	De novo	ex 26	c.3743A>G	p.(Gln1248Arg)	This study	P
Gross deletions						
KB11	NA	ex 1–2		p.?	This study	P
KB50	De novo	ex 5–9		p.?	(17)	P
Splice site						
KB314	De novo	int 11–12	c.975-1G>A	r.876_1320del; p.Cys293IlefsX26	This study	P
KB127	De novo	int 22–23	c.3384+3_3384+6del	r.3210_3284del; p.Asn1070_Lys1094del	(19)	P

^aDetected together with pathogenic mutation.

^bFalls in the last base of exon, predicted to disrupt the donor splice site.

^cPatients with compound heterozygous variants.

^dDetected together with other variant.

^eIdentified by MLPA.

^fDetected together with missense in KMT2D.

P, pathogenic; LP, likely pathogenic; LB, likely benign; VOUS variant of unknown significance.

Table 2. Bioinformatics analysis of KMT2D missense variants

Code	Variant	AA change	ACMG Variant Class	Prediction softwares											
				Polyphen		Align GVGD		Provean		SIFT		UMD predictor		Mutation Taster	
				Score	Result	Score	Result	Score	Result	Score	Result	Score	Result	Score	Result
KB21 ^a	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 ^{a,b}	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	102.71	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 ^a	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 ^a	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 ^a	c.4249A>G	p.(Met1417Val)	LB	PD	0.476	20.52	Class C15	00: 19	N	0.4097	T	60	PPM	DC	0.559
KB174	c.4252C>A	p.(Leu1418Met)	LB	PD	1.000	14.30	Class C0	-1.71	N	0.004	D	69	PP	DC	0.999
KB34	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	111.67	Class C65	0.05	N	0.1771	T	99	P	PM	0.885
KB535 ^a	c.4565A>G	p.(Gln1522Arg)	LB	PD	0.997	42.81	Class C35	-3.43	D	0.021	D	100	P	DC	0.999
KB119, KB204 ^a	c.5549G>A	p.(Gly1850Asp)	LB	PD	0.535	93.77	Class C65	-0.72	N	0.103	T	87	P	PM	0.997
KB330 ^a	c.6638G>A	p.(Gly2213Asp)	LB	PD	0.454	93.77	Class C65	-0.70	N	0.004	D	42	PM	DC	0.917
KB326 ^a	c.6811C>T	p.(Leu2245Val)	VOUS	PD	0.69	30.92	Class C25	-0.46	N	0.046	D	69	PP	PM	0.996
KB107 ^a	c.6970C>A	p.(Pro2271Ser)	LB	B	0.024	73.35	Class C65	-2.00	N	0.004	D	75	P	DC	0.850
KB430	c.7378C>T	p.(Arg2460Cys)	LB	PD	0.985	37.56	Class C35	-2.47	N	0.003	D	81	P	PM	0.998
KB122, KB287	c.7829T>C	p.(Leu2610Pro)	LB	PD	1.000	179.53	Class C65	-2.22	N	0.023	D	96	P	DC	0.999
KB27	c.8521C>A	p.(Pro2841Thr)	VOUS	B	0.076	97.78	Class C65	-1.20	N	0.0819	T	72	PP	DC	0.998
KB330 ^a	c.8774C>T	p.(Arg2925Val)	VOUS	B	0.009	37.56	Class C35	-1.45	N	0.010	D	69	PP	DC	0.997
KB443 ^a	c.9971G>T	p.(Gly3324Val)	LB	PD	0.988	109.55	Class C65	-2.24	N	0.011	D	78	P	DC	0.999
KB326 ^a , KB357	c.10192A>G	p.(Met3398Val)	LB	B	0.000	20.52	Class C15	-2.84	D	0.002	D	41	PM	PM	0.958
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	109.55	Class C65	-8.18	D	0.000	D	100	P	DC	0.999
KB86 ^a	c.110966C>T	p.(Arg3656Cys)	VOUS	PD	1.000	179.53	Class C65	-1.87	N	0.001	D	84	P	DC	0.999
KB374 ^a	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 ^a	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 ^a	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
KB512	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 ^a	c.14893G>A	p.(Ala4965Thr)	VOUS	PD	0.941	58.02	Class C55	-1.17	N	0.0847	T	75	P	DC	0.997

(continued)

Table 2. Continued

Code	Variant	AA change	ACMG Variant Class	Prediction softwares											
				Polyphen		Align GVGD		Provean		SIFT		UMD predictor		Mutation Taster	
				Score	Result	Score	Result	Score	Result	Score	Result	Score	Result	Score	Result
KB38 ^a	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	179.53	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 ^a	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	179.53	Class C65	-11.51	D	0.000	D	96	P	DC	0.999
KB171	c.15565G>A	p.(Gly5189Arg)	VOUS	PD	1.000	125.13	Class C65	-7.68	D	0.000	D	100	P	DC	0.999
KB264, KB376	c.15640C>T	p.(Arg5214Cys)	LP	PD	1.000	179.53	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	101.29	Class C65	-13.43	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	101.29	Class C65	-7.44	D	0.000	D	99	P	DC	0.999
KB120	c.16528T>G	p.(Tyr5510Asp)	LP	PD	1.000	159.94	Class C65	-9.52	D	0.000	D	96	P	DC	0.999

Bold values indicate *de novo* variants.

^aDetected together with other variant.

^bVariant that falls in the last base of the exon.

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.

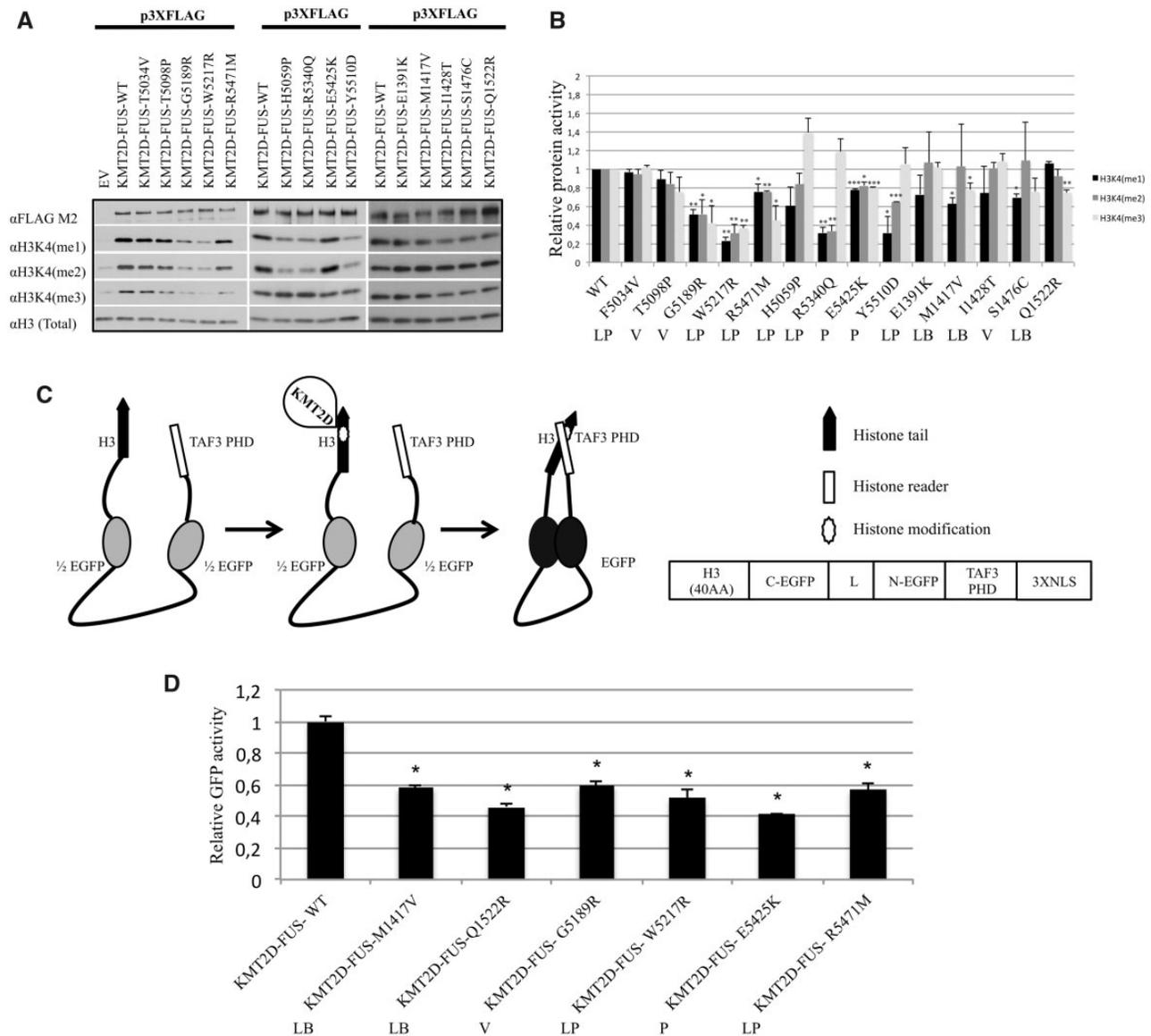


Figure 2. KMT2D missense variants are associated with defective methyltransferase activity and diminished H3K4 methylation. (A) H3K4 Lysine methyltransferase activity of mutated FLAG-KMT2D proteins on HeLa nucleosomes. (B) Relative H3K4 methyltransferase activity of semi-purified FLAG-KMT2D proteins on HeLa nucleosomes (mean \pm SD; n =two independent biological replicates). Data are expressed as fold differences compared with the activity of the wild-type control, arbitrarily set as 1. Pathogenic variants are indicated with P, likely pathogenic with LP, likely benign with LB, and VOUS with V. Student's T-test is indicated as * P <0.05; ** P <0.01; *** P <0.001. (C) Schematic representation of the epigenetic reporter allele encoding the two halves of GFP separated by a flexible linker region with a histone tail (H3) and a histone reader (TAF3 PHD) at the N and C termini, respectively (23). Modification of the histone tail by (KMT2D-mediated) trimethylation leads to reconstitution of the GFP structure and function, which can be measured by fluorescence. (D) The H3K4me3 indicator demonstrated a decreased H3K4me3 activity for the 6 tested missense variants compared with the WT. Pathogenic variants are indicated with P, likely pathogenic with LP, likely benign with LB, and VOUS with V. Student's t-test is indicated as * P <0.001.

with the WDR5 protein (26), we used transient transfection/co-immunoprecipitation assays to assess whether p.R5340Q alters the ability of KMT2D [expressed as an Myc-tagged green fluorescent protein (GFP) fusion] to physically bind a Flag-tagged WDR5 protein.

As shown in Figure 3B, western blot analysis of FLAG-immunoprecipitates using anti-Myc antibodies showed efficient co-immunoprecipitation of the wild-type KMT2D, but not of the R5340 mutant, with WDR5. This was not due to differences in expression levels/protein stability between the two KMT2D proteins, as documented by western blot analysis of EGFP in the total cell lysates, indicating that this variant completely abrogates the interaction with WDR5 (Fig. 3B).

Discussion

Here we report the mutation pattern of KMT2D and KDM6A in a cohort of 505 patients clinically diagnosed as KS-affected, and document the functional role of a subset of KMT2D missense mutations in impairing the protein enzymatic activity.

Overall, the mutation detection rate for KMT2D and KDM6A in our cohort (41%) was lower than that reported in a recent survey (22). This does not seem to be due to differences in sensitivity as the same Sanger Sequencing approach was used, and may be derived from the involvement of other causative genes, epigenetics mechanisms or yet unrecognized promoter or deep

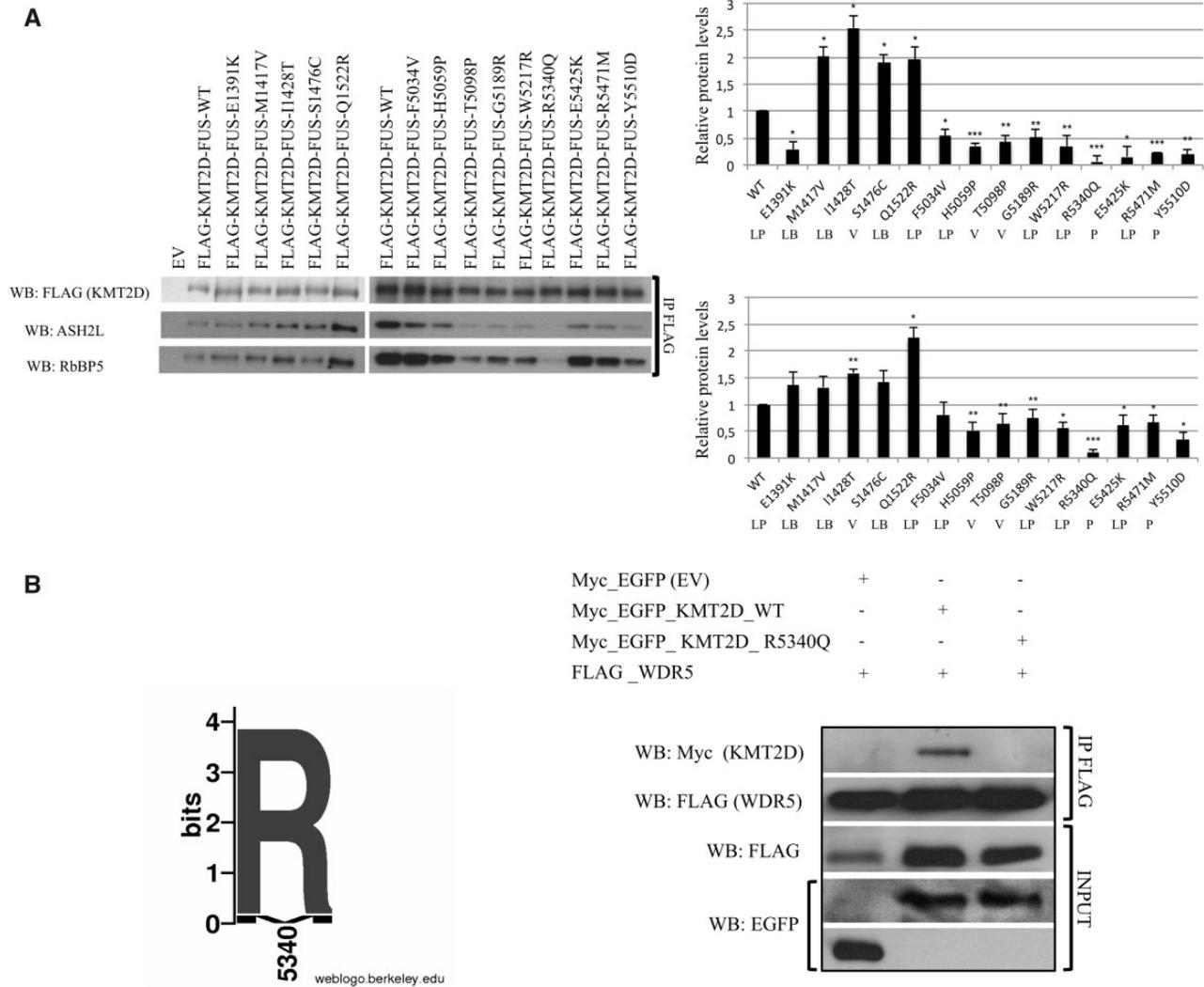


Figure 3. Interaction of KMT2D missense mutants with ASH2L, RbBP5 and WDR5. (A) On the left, immunoblot analysis of KMT2D, ASH2L and RbBP5 in HEK293T cell line transfected with wild-type KMT2D or KMT2D missense mutants followed by immunoprecipitation with anti-Flag antibody. On the right, relative interaction of the indicated KMT2D missense mutants with ASH2L (top) and RbBP5 (bottom) (mean \pm SD, n=2 independent replicates). Pathogenic variants are indicated with P, likely pathogenic with LP, likely benign with LB, and VOUS with V. Student's T-test is indicated as *P<0.05, **P<0.01 and ***P<0.001. (B) On the left, the predicted amino acid tolerance for KMT2D p.5340: the presence of Methionine (M) is very slightly tolerated. 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. EV, empty vector.

intronic variants affecting normal splicing. Moreover, we cannot exclude that some patients might have been clinically misdiagnosed and present conditions partially overlapping with KS. In view of multiple studies highlighting the clinical and molecular overlap of KS (27–31), these samples should be interrogated by using NGS targeted-genes panels focused on components of the histone methylation machinery that are associated to diseases in differential diagnosis with KS.

Our screening revealed 58 different missense variants distributed across the entire length of the KMT2D gene. Indeed, 9 of the 14 representative mutations that were tested in our study had variably reduced histone methylation activity *in vitro*, including H3K4 mono-, di- and/or tri-methylation. This effect is consistent with the known function of KMT2D as a major mono/dimethyltransferase, which is enriched at enhancer regions and is required for enhancer activation during cell differentiation, as reported in brown adipose tissue and skeletal

muscle development (11,32), or in mature B cells (33,34). Even though recent publications focus on the role of KMT2D as major monomethyltransferase, at least during carcinogenesis (35), KMT2D is also required during oogenesis and early development for bulk histone H3 lysine 4 trimethylation (36). The ability to modulate H3K4(me3), at active promoters/transcription start sites is important for the regulation of actively transcribed genes (14,15), and may reflect in part the formation of promoter-enhancer loops. As a matter of fact, six of the nine KS missense variants with reduced histone methylation activity showed a flawed H3K4(me3) activity in two different assays, emphasizing the role of KMT2D in modulating histone H3K4(me3) *in vitro*.

Of note, all nine functionally defective missense mutations were localized within the FYRN, WIN and SET domains, and may contribute to the KS phenotypes by interfering with its enzymatic activity in a manner analogous to C-terminal

truncating mutations. In contrast, amino acid changes in the N-terminal domain of the protein had minimal effects on H3K4 methylation. Indeed, according to ACMG, most of the amino acid changes in the N-terminal domain of the protein are classified as Likely Benign or Variants Of Unknown Significance (Table 2), suggesting that they may act through different mechanisms or represent neutral events.

The entire KMT2D C-terminal domain (aa 4507–5537) is involved in the interaction with the protein of the WRAD complex (25), albeit the specific amino acid interaction region is not fully characterized yet. Our study showed that missense variants localized in the C-terminal domains impaired the interaction of KMT2D with other components of its multi-subunit complex, including WDR5, RbBP5 and ASH2L, whereas most of the variants localized within the PHD 4–6 do not affect the interaction. The same C-terminal missense variants also showed altered protein domain structure and free energy interaction levels, suggesting that structural changes occurring in KMT2D may reduce the interaction with the proteins of the WRAD complex.

These data corroborate the hypothesis that the reduced methyltransferase activity is a direct consequence of the lack of multi-protein WRAD complex formation. Consistently, the missense variant R5340Q, localized within the WIN domain, fully abrogated the interaction with WDR5, resulting in a significantly reduced methyltransferase activity.

The classification and interpretation of KMT2D missense variants is a significant challenge in molecular diagnostics and genetic counseling. A number of prediction tools have been implemented in the last years, mainly as support for NGS data. However, *in silico* analyses should not 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 (37) recommended that these predictions should be used as support to additional findings, including functional evidence that can prove the pathogenicity of the variants found. Recently, 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 samples with pathogenic mutations from those with benign variants, and therefore enabling the functional assessment of genetic variants of unknown clinical significance (31,38–41).

Although additional studies will be required to dissect the precise mechanisms underlying the pathogenic role of the missense variants in KS, our data indicate that a subset of them may influence the disease status by affecting: (i) the methyltransferase activity of the protein and (ii) the interaction with the WRAD complex.

The biochemical approaches presented here may provide the basis for the development of diagnostic assays that could be of support to *in silico* prediction tools, with the advantage of addressing the functional effect of the variant.

Conclusions

This study expands the number of KMT2D and KDM6A mutations that are associated with KS and provides evidence that a number of naturally occurring missense mutations in KMT2D effectively impact KMT2D interaction and H3K4 methylation activity. These data have direct relevance for diagnostic and counseling purposes.

Materials and Methods

Patients and samples preparation

Our study cohort comprised a total of 505 index patients clinically diagnosed as affected by KS (Table 1), including 202 new cases that were referred to our institution between 2014 and 2017, and 303 patients from previous studies (19,42–48).

Patients were enrolled after obtaining appropriate informed consent by the physicians in charge, and approval by the local ethics committees. KS patients were included following the inclusions criteria as reported (46). Genomic DNA was extracted from fresh and/or frozen peripheral blood leukocytes of patients and their available family members using an automated DNA extractor and commercial DNA extraction Kits (Qiagen, Germany). Total RNA was extracted from peripheral blood leukocytes using TRIzol reagent (ThermoFisher Scientific, USA) and reverse transcribed using the QuantiTect Transcription kit (Qiagen), according to the manufacturer's instructions.

Sequencing and MLPA of KMT2D and KDM6A

Mutation screening of all 54 coding exons of the KMT2D (MIM #602113, NM_003482.3) gene and 29 coding exons of the KDM6A (MIM #300128, NM_021140.3) gene was performed by PCR amplification and direct sequencing as reported (46). MLPA analysis was performed as in (47).

In silico analysis of KMT2D and KDM6A variants

The putative causal and functional effect of missense variants was estimated by using the following *in silico* prediction tools: Polyphen-2 version 2.2.2 (<http://genetics.bwh.harvard.edu/pph>; date last accessed 5 July 2018) (49), Align GVGD (<http://agvgd.hci.utah.edu>; date last accessed 5 July 2018) (50), PROVEAN v1.1 (<http://provean.jcvi.org/index.php>; date last accessed 5 July 2018) (51), SIFT v1.03 (<http://sift.jcvi.org/>; date last accessed 5 July 2018) (52), UMD-predictor (<http://www.umd.be/>; date last accessed 5 July 2018) (53) and Mutation Taster (<http://www.mutationtaster.org/>; date last accessed 5 July 2018) (54) 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/>; date last accessed 5 July 2018) (55), NNSPLICE (http://www.fruitfly.org/seq_tools/splice.html; date last accessed 5 July 2018) (56) and NetGene2 (<http://www.cbs.dtu.dk/services/NetGene2>; date last accessed 5 July 2018) (57). Frequency variants were checked on dbSNP (<http://www.ncbi.nlm.nih.gov/projects/SNP/>; date last accessed 5 July 2018), 1000 Genomes Project (<http://www.internationalgenome.org/>; date last accessed 5 July 2018), EVS (<http://evs.gs.washington.edu/EVS/>; date last accessed 5 July 2018) and ExAC (<http://exac.broadinstitute.org/>; date last accessed 5 July 2018).

Protein modeling

Three-dimensional models of the ZF-PHD7 (residues 5059–5137), FYRN (residues 5180–5327) and SET (residues 5398–5537) domains of the human KMT2D protein were obtained employing a combination of threading and homology methods. For each domain, different models were generated using publicly available web servers (58–61). Final KMT2D model domains were obtained using the Modeller program and, as templates, the server models (62). A total of 200 unique models were generated by Modeller for each KMT2D domain. Qmean server (63)

was used to assess the quality of the predicted models and KoBaMIN web server (64) was used to carry out the protein energy minimization. The crystal structure of the WIN domain bound to WDR5 was obtained from the RCSB Protein Data Bank (PDB code: 3UVK; residues 5337–5347) (65). The effects of missense variants on KMT2D domains structure and stability were predicted by the Rosetta Backrub server (66) and FoldX program (67). Total charge of domains was calculated using the PROPKA program (68) and the electrostatic surface potentials were computed using the Adaptive Poisson-Boltzmann Solver (APBS) software (69). Models' inspection and model figures were performed using UCSF Chimera (version 1.11) (70).

Plasmids, cell lines 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 (25). We sub-cloned this partial KMT2D sequence into the p3XFlag-CMV14 vector (Sigma) using standard procedures. Fourteen representative KMT2D variants identified in our KS cohort that harbor amino acid changes in both N-terminal and C-terminal domains of the protein, including PHD 4–5–6–7, FYRN, WIN and SET domains, have been selected for functional assays: p.E1391K (Likely Pathogenic), p.M1417V (Likely Benign), p.I1428T (Likely Benign), p.S1476C (VOUS), p.Q1522R (Likely Benign), p.F5034V (Likely pathogenic), p.H5059P (Likely Pathogenic), p.T5098P (VOUS), p.G5189R (VOUS), p.W5217R (Likely pathogenic), p.R5340Q (Likely Pathogenic), p.E5425K (Pathogenic), p.R5471M (Likely Pathogenic) and p.Y5510D (Pathogenic). Expression plasmids harboring patient-derived KMT2D missense variants were generated by site-directed mutagenesis according to standard methods.

HEK 293 and 293T cells were cultured in Dulbecco's Modified Eagle Medium with 10% FBS, penicillin (100 U/ml) and streptomycin (100 µg/ml) (Life Technologies). The KMT2D C-terminal ORF was assembled into the pcDNA3-Myc-EGFP vector (71) by PCR site directed amplification using human cDNA and the pFlag-CMV2 FUSION-KMT2D vector as templates respectively. The WDR5 full length ORF was assembled into the p3XFlag-CMV14 vector (Sigma) by RT-PCR amplification reactions using cDNA from HEK 293T cells as template. HEK 293T cells were transiently transfected using the polyethylenimine method, following the published protocols (72). Cells were harvested 48 h after transfection and used for protein extraction and histone methyltransferase (HMT) assay.

In vitro HMT assay and epigenetic reporter allele

Partially purified FLAG-KMT2D wild-type and mutant derivative proteins were extracted from transfected HEK 293T cells by co-IP buffer (50 mM Tris, pH 7.5, 250 mM NaCl, 1% Triton X-100, 1 mM EDTA), followed by overnight incubation with EZview™ Red Anti-FLAG Affinity Gel (Sigma) at 4°C and final elution in BC100 buffer (20 mM Tris, pH 7.5, 10% glycerol, 0.2 mM EDTA, 1% Triton X-100, 100 mM NaCl) containing FLAG peptide (Sigma). KMT2D protein amounts were quantified by Coomassie staining and immunoblot analysis using mouse monoclonal FLAG antibodies. Enzymatic activity against HeLa nucleosomes was measured following a published method (25). 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 two times independently.

The epigenetic reporter allele, a kind gift from Dr Hans T. Bjornsson (McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University, Baltimore), was used as described (23).

Co-immunoprecipitation assay and western blotting analysis

Co-immunoprecipitations were performed using Dynabeads magnetic beads (ThermoFisher Scientific) and EZview™ Red Anti-FLAG Affinity Gel (Sigma) following the 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) (ThermoFisher Scientific) 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), 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), mouse monoclonal anti Myc (Roche), rabbit polyclonal anti Ash2 (Bethyl) and rabbit polyclonal anti RbBP5 (Bethyl). Horseradish peroxidase conjugated anti-mouse (Santa Cruz) or anti-rabbit (Santa Cruz) antibodies, and the ECL chemiluminescence system (GE Healthcare) were used for detection. ImageJ software (<http://imagej.nih.gov/ij/>; date last accessed 5 July 2018) was used to quantify band signal intensity. Values are expressed as fold differences relative to the wild-type protein sample, set at 1, after normalization for the loading control.

Supplementary Material

Supplementary Material is available at HMG online.

Acknowledgements

We acknowledge the family that agreed to participate and made this study possible. We also acknowledge Professor Hans T. Bjornsson for providing Epigenetic reporter allele vector. We are grateful to the Genomic Disorder Biobank and Telethon Network of Genetic Biobanks (Telethon Italy Grant GTB12001G) for biobanking biospecimens.

Conflict of Interest statement: The authors declare no conflicts of interest with the exception of G.M. who is a consultant for Takeda Pharmaceutical Company.

Funding

N.I.H. Grant RO1-CA172492 (to L.P.), in part; Lymphoma Research Foundation post-doctoral fellowship to J.Z. Telethon - Italy (Grant no. GGP13231), Italian Ministry of Health, Jérôme Lejeune Foundation, Daunia Plast, Circolo Unione Apricena, Fidapa Apricena, and Associazione Italiana Sindrome Kabuki (AISK) to G.M. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

References

- Kuroki, Y., Suzuki, Y., Chyo, H., Hata, A. and Matsui, I. (1981) 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.*, **99**, 570–573.
- Niikawa, N., Matsuura, N., Fukushima, Y., Ohsawa, T. and Kajii, T. (1981) Kabuki make-up syndrome: a syndrome of mental retardation, unusual facies, large and protruding ears, and postnatal growth deficiency. *J. Pediatr.*, **99**, 565–569.
- Cheon, C.K. and Ko, J.M. (2015) Kabuki syndrome: clinical and molecular characteristics. *Korean J. Pediatr.*, **58**, 317–324.
- 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. et al. (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat. Genet.*, **42**, 790–793.
- Shilatifard, A. (2012) The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu. Rev. Biochem.*, **81**, 65–95.
- Kouzarides, T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
- Li, B., Carey, M. and Workman, J.L. (2007) The role of chromatin during transcription. *Cell*, **128**, 707–719.
- Ang, S.Y., Uebersohn, A., Spencer, C.I., Huang, Y., Lee, J.E., Ge, K. and Bruneau, B.G. (2016) KMT2D regulates specific programs in heart development via histone H3 lysine 4 di-methylation. *Development*, **143**, 810–821.
- Herz, H.M., Mohan, M., Garruss, A.S., Liang, K., Takahashi, Y.H., Mickey, K., Voets, O., Verrijzer, C.P. and Shilatifard, A. (2012) Enhancer-associated H3K4 monomethylation by Trithorax-related, the Drosophila homolog of mammalian Mll3/Mll4. *Genes. Dev.*, **26**, 2604–2620.
- Hu, D., Gao, X., Morgan, M.A., Herz, H.M., Smith, E.R. and Shilatifard, A. (2013) The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol. Cell. Biol.*, **33**, 4745–4754.
- Lee, J.E., Wang, C., Xu, S., Cho, Y.W., Wang, L., Feng, X., Baldridge, A., Sartorelli, V., Zhuang, L., Peng, W. et al. (2013) H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. *Elife (Cambridge)*, **2**, e01503.
- 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. et al. (2013) KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. *Oncotarget*, **4**, 2144–2153.
- 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. et al. (2013) Remodeling of the enhancer landscape during macrophage activation is coupled to enhancer transcription. *Mol. Cell*, **51**, 310–325.
- 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. et al. (2009) Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature*, **459**, 108–112.
- Denissov, S., Hofemeister, H., Marks, H., Kranz, A., Ciotta, G., Singh, S., Anastassiadis, K., Stunnenberg, H.G. and Stewart, A.F. (2014) Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development*, **141**, 526–537.
- Issaeva, I., Zonis, Y., Rozovskaia, T., Orlovsky, K., Croce, C.M., Nakamura, T., Mazo, A., Eisenbach, L. and Canaani, E. (2007) Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. *Mol. Cell. Biol.*, **27**, 1889–1903.
- Lederer, D., Grisart, B., Digilio, M.C., Benoit, V., Crespin, M., Ghariani, S.C., Maystadt, I., Dallapiccola, B. and Verellen-Dumoulin, C. (2012) Deletion of KDM6A, a histone demethylase interacting with MLL2, in three patients with Kabuki syndrome. *Am. J. Hum. Genet.*, **90**, 119–124.
- Miyake, N., Mizuno, S., Okamoto, N., Ohashi, H., Shiina, M., Ogata, K., Tsurusaki, Y., Nakashima, M., Saitsu, H., Niikawa, N. et al. (2013) KDM6A point mutations cause Kabuki syndrome. *Hum. Mutat.*, **34**, 108–110.
- Micale, L., Augello, B., Maffeo, C., Selicorni, A., Zucchetti, F., Fusco, C., De Nittis, P., Pellico, M.T., Mandriani, B., Fischetto, R. et al. (2014) Molecular analysis, pathogenic mechanisms, and readthrough therapy on a large cohort of Kabuki syndrome patients. *Hum. Mutat.*, **35**, 841–850.
- Hong, S., Cho, Y.W., Yu, L.R., Yu, H., Veenstra, T.D. and Ge, K. (2007) Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. *Proc. Natl. Acad. Sci. U.S.A.*, **104**, 18439–18444.
- Lan, F., Bayliss, P.E., Rinn, J.L., Whetstone, J.R., Wang, J.K., Chen, S., Iwase, S., Alpatov, R., Issaeva, I., Canaani, E. et al. (2007) A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature*, **449**, 689–694.
- Bogershausen, N., Gatinois, V., Riehmer, V., Kayserili, H., Becker, J., Thoenes, M., Simsek-Kiper, P.O., Barat-Houari, M., Elcioglu, N.H., Wiczorek, D. et al. (2016) Mutation update for Kabuki syndrome genes KMT2D and KDM6A and further delineation of X-linked Kabuki syndrome subtype 2. *Hum. Mutat.*, **37**, 847–864.
- 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. and Dietz, H.C. (2014) Histone deacetylase inhibition rescues structural and functional brain deficits in a mouse model of Kabuki syndrome. *Sci. Transl. Med.*, **6**, 256ra135.
- Kim, J.H., Sharma, A., Dhar, S.S., Lee, S.H., Gu, B., Chan, C.H., Lin, H.K. and Lee, M.G. (2014) UTX and MLL4 coordinately regulate transcriptional programs for cell proliferation and invasiveness in breast cancer cells. *Cancer Res.*, **74**, 1705–1717.
- Dhar, S.S., Lee, S.H., Kan, P.Y., Voigt, P., Ma, L., Shi, X., Reinberg, D. and Lee, M.G. (2012) Trans-tail regulation of MLL4-catalyzed H3K4 methylation by H4R3 symmetric dimethylation is mediated by a tandem PHD of MLL4. *Genes Dev.*, **26**, 2749–2762.
- Shinsky, S.A., Hu, M., Vought, V.E., Ng, S.B., Bamshad, M.J., Shendure, J. and Cosgrove, M.S. (2014) 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.*, **426**, 2283–2299.
- Au, P.Y.B., You, J., Caluseriu, O., Schwartztruber, J., Majewski, J., Bernier, F.P., Ferguson, M., Valle, D., Parboosingh, J.S., Sobreira, N. et al. (2015) 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.*, **36**, 1009–1014.
- 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. et al. (2016) A de novo frameshift in HNRNPK causing a

- Kabuki-like syndrome with nodular heterotopia. *Clin. Genet.*, **90**, 258–262.
29. Yuan, B., Pehlivan, D., Karaca, E., Patel, N., Charnig, W.L., Gambin, T., Gonzaga-Jauregui, C., Sutton, V.R., Yesil, G., Bozdogan, S.T. et al. (2015) Global transcriptional disturbances underlie Cornelia de Lange syndrome and related phenotypes. *J. Clin. Invest.*, **125**, 636–651.
 30. Parenti, I., Teresa-Rodrigo, M.E., Pozojevic, J., Ruiz Gil, S., Bader, I., Braunholz, D., Bramswig, N.C., Gervasini, C., Larizza, L., Pfeiffer, L. et al. (2017) Mutations in chromatin regulators functionally link Cornelia de Lange syndrome and clinically overlapping phenotypes. *Hum. Genet.*, **136**, 307–320.
 31. Sobreira, N., Brucato, M., Zhang, L., Ladd-Acosta, C., Ongaco, C., Romm, J., Doheny, K.F., Mingroni-Netto, R.C., Bertola, D., Kim, C.A. et al. (2017) Patients with a Kabuki syndrome phenotype demonstrate DNA methylation abnormalities. *Eur. J. Hum. Genet.*, **25**, 1335–1344.
 32. Guenther, M.G., Jenner, R.G., Chevalier, B., Nakamura, T., Croce, C.M., Canaani, E. and Young, R.A. (2005) Global and Hox-specific roles for the MLL1 methyltransferase. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 8603–8608.
 33. Zhang, J., Dominguez-Sola, D., Hussein, S., Lee, J.E., Holmes, A.B., Bansal, M., Vlasevska, S., Mo, T., Tang, H., Basso, K. et al. (2015) Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. *Nat. Med.*, **21**, 1190–1198.
 34. Lindsley, A.W., Saal, H.M., Burrow, T.A., Hopkin, R.J., Shchelochkov, O., Khandelwal, P., Xie, C., Bleesing, J., Filipovich, L., Risma, K. et al. (2016) Defects of B-cell terminal differentiation in patients with type-1 Kabuki syndrome. *J. Allergy Clin. Immunol.*, **137**, 179–187.
 35. Lv, S., Ji, L., Chen, B., Liu, S., Lei, C., Liu, X., Qi, X., Wang, Y., Lai-Han Leung, E., Wang, H. et al. (2018) Histone methyltransferase KMT2D sustains prostate carcinogenesis and metastasis via epigenetically activating LIFR and KLF4. *Oncogene*, **37**, 1354–1368.
 36. Andreu-Vieyra, C.V., Chen, R., Agno, J.E., Glaser, S., Anastassiadis, K., Stewart, A.F. and Matzuk, M.M. (2010) MLL2 is required in oocytes for bulk histone 3 lysine 4 trimethylation and transcriptional silencing. *PLoS Biol.*, **8**, e1000453.
 37. Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E. et al. (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.*, **17**, 405–424.
 38. 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. et al. (2017) CHARGE and Kabuki syndromes: gene-specific DNA methylation signatures identify epigenetic mechanisms linking these clinically overlapping conditions. *Am. J. Hum. Genet.*, **100**, 773–788.
 39. 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. et al. (2015) NSD1 mutations generate a genome-wide DNA methylation signature. *Nat. Commun.*, **6**, 10207.
 40. 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. et al. (2013) Multilocus loss of DNA methylation in individuals with mutations in the histone H3 lysine 4 demethylase KDM5C. *BMC Med. Genomics*, **6**, 1.
 41. Kernohan, K.D., Cigana Schenkel, L., Huang, L., Smith, A., Pare, G., Ainsworth, P., Boycott, K.M., Warman-Chardon, J. and Sadikovic, B. (2016) Identification of a methylation profile for DNMT1-associated autosomal dominant cerebellar ataxia, deafness, and narcolepsy. *Clin. Epigenetics*, **8**, 91.
 42. Andersen, M.S., Menazzi, S., Brun, P., Cocah, C., Merla, G. and Solari, A. (2014) Clinical diagnosis of Kabuki syndrome: phenotype and associated abnormalities in two new cases. *Arch. Argent. Pediatr.*, **112**, 26–32.
 43. Cappuccio, G., Rossi, A., Fontana, P., Acampora, E., Avolio, V., Merla, G., Zelante, L., Secinaro, A., Andria, G. and Melis, D. (2014) Bronchial isomerism in a Kabuki syndrome patient with a novel mutation in MLL2 gene. *BMC Med. Genet.*, **15**, 15.
 44. Ejarque, I., Uliana, V., Forzano, F., Marciano, C., Merla, G., Zelante, L., Di Maria, E. and Faravelli, F. (2011) Is Hardikar syndrome distinct from Kabuki (Niikawa-Kuroki) syndrome? *Clin. Genet.*, **80**, 493–496.
 45. Makrythanasis, P., van Bon, B.W., Steehouwer, M., Rodriguez-Santiago, B., Simpson, M., Dias, P., Anderlid, B.M., Arts, P., Bhat, M., Augello, B. et al. (2013) MLL2 mutation detection in 86 patients with Kabuki syndrome: a genotype-phenotype study. *Clin. Genet.*, **84**, 539–545.
 46. Micale, L., Augello, B., Fusco, C., Selicorni, A., Loviglio, M.N., Silengo, M.C., Reymond, A., Gumiero, B., Zucchetti, F., D'Addetta, E.V. et al. (2011) Mutation spectrum of MLL2 in a cohort of Kabuki syndrome patients. *Orphanet J. Rare Dis.*, **6**, 38.
 47. Priolo, M., Micale, L., Augello, B., Fusco, C., Zucchetti, F., Prontera, P., Paduano, V., Biamino, E., Selicorni, A., Mammi, C. et al. (2012) Absence of deletion and duplication of MLL2 and KDM6A genes in a large cohort of patients with Kabuki syndrome. *Mol. Genet. Metab.*, **107**, 627–629.
 48. Ratbi, I., Fejjal, N., Micale, L., Augello, B., Fusco, C., Lyahyai, J., Merla, G. and Sefiani, A. (2013) Report of the first clinical case of a Moroccan Kabuki patient with a novel MLL2 mutation. *Mol. Syndromol.*, **4**, 152–156.
 49. Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S. and Sunyaev, S.R. (2010) A method and server for predicting damaging missense mutations. *Nat. Methods*, **7**, 248–249.
 50. Tavtigian, S.V., Deffenbaugh, A.M., Yin, L., Judkins, T., Scholl, T., Samollow, P.B., de Silva, D., Zharkikh, A. and Thomas, A. (2005) Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *J. Med. Genet.*, **43**, 295–305.
 51. Choi, Y., Sims, G.E., Murphy, S., Miller, J.R. and Chan, A.P. (2012) Predicting the functional effect of amino acid substitutions and indels. *PLoS One*, **7**, e46688.
 52. Kumar, P., Henikoff, S. and Ng, P.C. (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.*, **4**, 1073–1081.
 53. Frederic, M.Y., Lalande, M., Boileau, C., Hamroun, D., Claustres, M., Beroud, C. and Collod-Beroud, G. (2009) UMD-predictor, a new prediction tool for nucleotide substitution pathogenicity – application to four genes: FBN1, FBN2, TGFB1, and TGFB2. *Hum. Mutat.*, **30**, 952–959.
 54. Schwarz, J.M., Cooper, D.N., Schuelke, M. and Seelow, D. (2014) MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods*, **11**, 361–362.
 55. Desmet, F.O., Hamroun, D., Lalande, M., Collod-Beroud, G., Claustres, M. and Beroud, C. (2009) Human splicing finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.*, **37**, e67.

56. Reese, M.G., Eeckman, F.H., Kulp, D. and Haussler, D. (1997) Improved splice site detection in Genie. *J. Comput. Biol.*, **4**, 311–323.
57. Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouze, P. and Brunak, S. (1996) Splice site prediction in *Arabidopsis thaliana* pre-mRNA by combining local and global sequence information. *Nucleic Acids Res.*, **24**, 3439–3452.
58. Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics*, **9**, 40.
59. Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N. and Sternberg, M.J. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.*, **10**, 845–858.
60. Li, J., Deng, X., Eickholt, J. and Cheng, J. (2013) Designing and benchmarking the MULTICOM protein structure prediction system. *BMC Struct. Biol.*, **13**, 2.
61. Kallberg, M., Wang, H., Wang, S., Peng, J., Wang, Z., Lu, H. and Xu, J. (2012) Template-based protein structure modeling using the RaptorX web server. *Nat. Protoc.*, **7**, 1511–1522.
62. Sali, A. and Blundell, T.L. (1993) Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.*, **234**, 779–815.
63. Benkert, P., Kunzli, M. and Schwede, T. (2009) QMEAN server for protein model quality estimation. *Nucleic Acids Res.*, **37**, W510–W514.
64. Rodrigues, J.P., Levitt, M. and Chopra, G. (2012) KoBaMIN: a knowledge-based minimization web server for protein structure refinement. *Nucleic Acids Res.*, **40**, W323–W328.
65. Zhang, P., Lee, H., Brunzelle, J.S. and Couture, J.F. (2012) The plasticity of WDR5 peptide-binding cleft enables the binding of the SET1 family of histone methyltransferases. *Nucleic Acids Res.*, **40**, 4237–4246.
66. Lauck, F., Smith, C.A., Friedland, G.F., Humphris, E.L. and Kortemme, T. (2010) RosettaBackrub—a web server for flexible backbone protein structure modeling and design. *Nucleic Acids Res.*, **38**, W569–W575.
67. Guerois, R., Nielsen, J.E. and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.*, **320**, 369–387.
68. Rostkowski, M., Olsson, M.H., Sondergaard, C.R. and Jensen, J.H. (2011) Graphical analysis of pH-dependent properties of proteins predicted using PROPKA. *BMC Struct. Biol.*, **11**, 6.
69. Baker, N.A., Sept, D., Joseph, S., Holst, M.J. and McCammon, J.A. (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 10037–10041.
70. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
71. Reymond, A., Meroni, G., Fantozzi, A., Merla, G., Cairo, S., Luzi, L., Riganelli, D., Zanaria, E., Messali, S., Cainarca, S. et al. (2001) The tripartite motif family identifies cell compartments. *EMBO J.*, **20**, 2140–2151.
72. Thomas, M., Lu, J.J., Ge, Q., Zhang, C., Chen, J. and Klivanov, A.M. (2005) Full deacylation of polyethylenimine dramatically boosts its gene delivery efficiency and specificity to mouse lung. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 5679–5684.
73. Pasqualucci, L., Trifonov, V., Fabbri, G., Ma, J., Rossi, D., Chiarenza, A., Wells, V.A., Grunn, A., Messina, M., Elliot, O. et al. (2011) Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat. Genet.*, **43**, 830–837.
74. Dentici, M.L., Di Pede, A., Lepri, F.R., Gnazzo, M., Lombardi, M.H., Auriti, C., Petrocchi, S., Pisaneschi, E., Bellacchio, E., Capolino, R. et al. (2015) Kabuki syndrome: clinical and molecular diagnosis in the first year of life. *Arch. Dis. Child.*, **100**, 158–164.
75. 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. et al. (2011) Spectrum of MLL2 (ALR) mutations in 110 cases of Kabuki syndrome. *Am. J. Med. Genet. A*, **155A**, 1511–1516.
76. Banka, S., Veeramachaneni, R., Reardon, W., Howard, E., Bunstone, S., Ragge, N., Parker, M.J., Crow, Y.J., Kerr, B., Kingston, H. et al. (2012) How genetically heterogeneous is Kabuki syndrome? MLL2 testing in 116 patients, review and analyses of mutation and phenotypic spectrum. *Eur. J. Hum. Genet.*, **20**, 381–388.
77. Paulussen, A.D., Stegmann, A.P., Blok, M.J., Tserpelis, D., Posma-Velter, C., Detisch, Y., Smeets, E.E., Wagemans, A., Schrandt, J.J., van den Boogaard, M.J. et al. (2011) MLL2 mutation spectrum in 45 patients with Kabuki syndrome. *Hum. Mutat.*, **32**, E2018–E2025.
78. 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. and Wu, M.C. (2015) Immunologic assessment and KMT2D mutation detection in Kabuki syndrome. *Clin. Genet.*, **88**, 255–260.
79. 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. and Shaikh, T.H. (2015) Kabuki syndrome genes KMT2D and KDM6A: functional analyses demonstrate critical roles in craniofacial, heart and brain development. *Hum. Mol. Genet.*, **24**, 4443–4453.
80. 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. et al. (2014) Identification of KMT2D and KDM6A mutations by exome sequencing in Korean patients with Kabuki syndrome. *J. Hum. Genet.*, **59**, 321–325.
81. 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. et al. (2015) Speech and language in a genotyped cohort of individuals with Kabuki syndrome. *Am. J. Med. Genet. A*, **167**, 1483–1492.
82. Kokitsu-Nakata, N.M., Petrin, A.L., Heard, J.P., Vendramini-Pittoli, S., Henkle, L.E., dos Santos, D.V., Murray, J.C. and Richieri-Costa, A. (2012) Analysis of MLL2 gene in the first Brazilian family with Kabuki syndrome. *Am. J. Med. Genet. A*, **158A**, 2003–2008.
83. Liu, S., Hong, X., Shen, C., Shi, Q., Wang, J., Xiong, F. and Qiu, Z. (2015) Kabuki syndrome: a Chinese case series and systematic review of the spectrum of mutations. *BMC Med. Genet.*, **16**, 26.
84. Banka, S., Lederer, D., Benoit, V., Jenkins, E., Howard, E., Bunstone, S., Kerr, B., McKee, S., Chris Lloyd, I., Shears, D. et al. (2015) Novel KDM6A (UTX) mutations and a clinical and molecular review of the X-linked Kabuki syndrome (KS2). *Clin. Genet.*, **87**, 252–258.
85. Lederer, D., Shears, D., Benoit, V., Verellen-Dumoulin, C. and Maystadt, I. (2014) A three generation X-linked family with Kabuki syndrome phenotype and a frameshift mutation in KDM6A. *Am. J. Med. Genet. A*, **164**, 1289–1292.