Original Article

**MLL2 mutation detection in 86 patients with Kabuki syndrome: a genotype–phenotype study**


Recently, pathogenic variants in the **MLL2** gene were identified as the most common cause of Kabuki (Niikawa–Kuroki) syndrome (MIM#147920). To further elucidate the genotype–phenotype correlation, we studied a large cohort of 86 clinically defined patients with Kabuki syndrome (KS) for mutations in **MLL2**. All patients were assessed using a standardized phenotype list and all were scored using a newly developed clinical score list for KS (MLL2-Kabuki score 0–10). Sequencing of the full coding region and intron–exon boundaries of **MLL2** identified a total of 45 likely pathogenic mutations (52%): 31 nonsense, 10 missense and four splice-site mutations, 34 of which were novel. In five additional patients, novel, i.e. non-dbSNP132 variants of clinically unknown relevance, were identified. Patients with likely pathogenic nonsense or missense **MLL2** mutations were usually more severely affected (median ‘MLL2-Kabuki score’ of 6) as compared to the patients without **MLL2** mutations (median ‘MLL2-Kabuki score’ of 5), a significant difference (p < 0.0014). Several typical facial features such as large dysplastic ears, arched eyebrows with sparse lateral third, blue sclerae, a flat nasal root with a broad nasal root, and a thin upper and a full lower lip were observed more often in mutation positive patients.

**Conflict of interest**

Nothing to declare.
Kabuki (Niikawa–Kuroki) syndrome (KS; OMIM 147920), is a multiple malformation syndrome characterized by mild to moderate intellectual disability and a distinctive facial phenotype that includes arched eyebrows with a sparse or dispersed lateral one-third, long palpebral fissures, eversion of the lower eyelid, a depressed nasal tip and large prominent ears (1–3). Other commonly found features include non-specific skeletal anomalies, cardiac and renal anomalies, fetal fingertip pads and postnatal growth deficiency leading to short stature in adult life. Notably, the severity of some of the phenotypic features alters with age; e.g. most KS facial dysmorphic features becomes less prominent at an older age, with the exception of the long palpebral fissures which persist.

The KS is an autosomal dominant disorder and pathogenic mutations in the MLL2 gene have been identified in a large number of affected individuals (myeloid/lymphoid or mixed lineage leukemia 2, OMIM 602113.4–9). MLL2 has 54 coding exons and encodes a histone–lysine N–methyltransferase of 5537 amino acids that belongs to the Trithorax group of proteins (10). A number of different syndromes with congenital malformations are associated with haploinsufficiency of genes encoding methyltransferases including NSD1 [OMIM 606681, Sotos (11) and Weaver syndromes (12)]. EZH2 [OMIM 601573, Weaver syndrome (13, 14)] EHM1 [OMIM 607001, Kleefstra syndrome (15)] SETBP1 [OMIM 611060, Schinzel-Giedion syndrome (16)] and DNMT1 [OMIM 126375, Hereditary Sensory Neuropathy type IE (17)]. MLL2 is also known to be part of ASC-2 (ASCOM, 18), a protein complex functioning as a transcriptional regulator (19–21).

Here, all 54 coding exons and exon–intron boundaries of MLL2 were investigated by Sanger sequencing in a series of 86 patients with a detailed clinical description consistent with KS. We also developed a systematic KS phenotype score (‘MLL2-Kabuki score’, Table 1) to facilitate accurate genotype–phenotype studies.

Patients and methods

Patient selection and phenotype criteria

We investigated 86 patients with the clinical diagnosis of KS, three of whom were familial cases (mildly affected mothers) with the remainder representing sporadic KS cases. Informed consent was obtained from all patients/families after ethics approval. All patient phenotypes were recorded with a standardized clinical questionnaire by the respective counseling clinical geneticist, i.e. a single person per patient (Table S1) which formed the basis of a novel clinical KS phenotype score (‘MLL2-Kabuki score’, 0–10, Table 1), which was developed by B. W. v. B. and B. B. A. d. V. prior to mutation testing in such a way that each score point distinguished between the presence or absence of a clinical feature in approximately half of all individuals included.

Sequencing of the protein-coding part of MLL2

All coding 54 exons and splice junctions of MLL2 were amplified by polymerase chain reaction and subsequently sequenced by capillary (Sanger) sequencing. The identified mutations were tested for a de novo origin when parental DNA was available.

Pathogenicity assessment

All nonsense and frameshift mutations were classified as pathogenic. All missense mutations that occurred de novo were classified as causative. For all missense variants, four in silico prediction tools [SIFT (22), AGVGD (23), Mutation Taster, and MutPred (24), Table S3] were used to classify these variants. In order for a variant to be deemed pathogenic, it had to be predicted by at least two of these programs.
Potential splice site mutations were analysed by in silico prediction tools [Splice Site Prediction by Neural Network (25); NetGene2 (26); Human Splicing Finder (27), Table S2], if at least two of those indicated an effect on splicing, the variants were classified as likely pathogenic.

Results

Variants identified in the MLL2 gene

In the 86 samples sequenced, we identified 45 (in 52% of the cases) mutations, 34 of which were novel, that are most likely pathogenic based on assessment by four in silico prediction tools. Of these mutations, 13 were nonsense, 18 were frameshift mutations leading to a premature stop codon (Table S2), 10 were missense (Table S3) and four were splice-site mutations (Table S4). The parental origin of 16 mutations was investigated of which 15 were de novo (10 nonsense, 4 frameshift, and 1 missense). The MLL2 p.(Pro647Glut) mutation was predicted to be pathogenic by only one of the four algorithms. However, since it was found to be de novo it led us to categorize it as pathogenic, underscoring the difficulties of novel variant interpretation in this gene. In addition, two missense and three potential splice-site variants of unknown significance were detected (Table S5), as well as novel benign polymorphisms in five cases (Table S5). No MLL2 coding variants were detected in 31 patients (36%). The identified synonymous changes or known single nucleotide polymorphism (SNP) variants are not reported.

In total, of the 238 pathogenic mutations reported here and previously ([4–9], Fig. 1], 173 are nonsense and frameshifts (72.7%), 40 are missense (16.8%) and 21 affected splice sites (8.8%; Fig. 1, Figure S1, Table S10). The mutations of all published studies are compared in Table S10. It is very interesting to note that 82.5% (33/40) of the missense mutations are found in and around the 4th to 6th PHD (plant homeodomain finger) domain and at the C terminus part of the protein at the FYRN (FY-rich domain N-terminal region), FYRC (FY-rich domain, C-terminal region), SET (Su(var)3-9, Enhancer-of-zeste, Trithorax) and PostSET domains.

We observed more than one variant in nine cases, but of these, only one per patient was classified as pathogenic by the in silico prediction tools; parental samples were not available for testing for these patients to determine parent of origin or cis/trans configuration, several of the second hits are now categorized as benign polymorphisms. In two patients, additional chromosomal aberrations were identified by DNA microarray-based copy number variant (CNV) detection. The first patient (patient 41) has an p.(Ser5498Phe) mutation in MLL2 and a de novo interstitial duplication on chromosome 4p16.3 (chr4:954,359-1,798,665, hg18, with the last non-duplicated marker SNP_A-2301516, first duplicated marker SNP_A-2190948; last duplicated marker SNP_A-2241290 first non-duplicated marker SNP_A-1913509). In a second case (patient 40), a heterozygous de novo MLL2 mutation (c.16294C>T, p.(Arg5432Trp)) was detected; and a mosaic approximately 1 Mb deletion on chromosome 1q43q44 was identified by pre- and postnatal SNP microarray and FISH analysis (Fig. 2).

Genotype–phenotype correlation

A genotype–phenotype correlation was performed (Fisher’s exact test) by comparing patients with pathogenic MLL2 mutations to those without. If data were unavailable for specific phenotypes, they were categorized as unknown and the patient was excluded from the specific phenotype analysis. The results are summarized in Table S6. Table S7 shows a comparison of the clinical features of patients with nonsense MLL2 mutations versus those of patients without an MLL2 mutation (Figure S2).

Both tables are ranked by the frequency with which each feature is reported. It is evident that the primary features used by the different clinical teams in order to make the diagnosis of KS, namely long palpebral fissures, persistent fetal finger tip pads, everted lower eyelids and brachy/clinodactyly, are listed first however these showed no significant differences in frequency between the patients with and without MLL2 mutations (Table S6). Of the remaining features, six showed a difference between the two groups with a p value below 0.05 (flat nasal tip p = 0.001, intellectual disability

541
c.5865_5867delinsCCCCC

c.8721C>G,
c.12753 12754del
c.12414dup*, c.12441del

c.9494del, c.9581del,
c.5886del3, c.3329_3333del

Fig. 1. Schematic representation of MLL2 mutations identified to date. Left side: the structure of MLL2 gene including all 54 exons (orange rectangles), the 3’ untranslated region (yellow rectangle) and introns (orange horizontal line); right side: the MLL2 protein domains: PHD, plant homeodomain finger; HMG-box, high mobility group; CC, coiled coil; LXXLL domain, FYRN, FY-rich domain, N-terminal region; FYRC, FY-rich (missense), and in black (small duplication). The variants identified in this study are highlighted by *.
Table 1. Kabuki phenotype scoring list considering 23 individual phenotypic features.

<table>
<thead>
<tr>
<th>Clinical feature details</th>
<th>Clinical feature details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facial features</td>
<td>Possible Score</td>
</tr>
<tr>
<td>Long palpebral fissures; everted lower eyelids; large dysplastic ears; arched eyebrows, sparse lateral one third; flat nasal tip; abnormal dentition; high/cleft palate; strabismus; blue sclera; micrognathia; ptosis; broad nasal root; oligodontia; thin upper and full lower lip; lipnodules</td>
<td>0–5 points (0–3 features = 1 point; 4–6 = 2 point; 7–9 = 3 point; 10–12 = 4 point; 13–15 = 5 point)</td>
</tr>
<tr>
<td>Limb/extremity features</td>
<td>Up to 1 point (0–1 features = 0 point; 2–4 = 1 point)</td>
</tr>
<tr>
<td>Microcephaly</td>
<td>1 point</td>
</tr>
<tr>
<td>Short stature</td>
<td>1 point</td>
</tr>
<tr>
<td>Heart</td>
<td>1 point</td>
</tr>
<tr>
<td>Kidney</td>
<td>1 point</td>
</tr>
<tr>
<td>Sum</td>
<td>0–10</td>
</tr>
</tbody>
</table>

Fig. 2. Representative facial images of patients with Kabuki syndrome with MLL2 mutations. This figure shows facial images of Kabuki patients, in which we identified MLL2 mutations. Shown are patients 11, 12, 23, 25, 26, 27, 35 and 40 (carrying the following mutations, respectively: p.(Gln4401*), p.(Arg4484*), p.(Phe2244Ilefs*11), p.(Ser2438Ilefs*11), p.(Phe2494Serfs*49), p.(His2914Profs*6), p.(Arg5030Cys), and p.(Arg5432Trp). Patient 40 is from Faas et al. postnatal identification of Kabuki syndrome in a girl with a prenatally diagnosed mosaic 1q43q44 submicroscopic deletion (in preparation).

This distinction marks the difference between the classically described KS and the more wide description of ‘Kabuki-like’ syndrome, i.e. commonly assigned to several patients. The proportion of patients with intellectual disability differed between the patients with and without MLL2 mutations (p = 0.004). Intellectual disability appears to also be an excellent positive discriminator for the presence of MLL2 mutations, as all patients, with MLL2 mutations (with exception of the affected mother of a patient) had intellectual disability.

Finally, we applied a new ‘MLL2-Kabuki score’ for which all patients had been scored prior to mutation status assessment. This score was designed in such a way that each score point distinguished between the presence or absence of a clinical feature in approximately half of all individuals included (Tables S1, S8). Patients with likely pathogenic nonsense or missense MLL2 mutations were usually more severely affected (mean ‘MLL2-Kabuki score’ of 6.1) as compared to the patients without MLL2 mutations (mean
Makrythanasis et al.

‘MLL2-Kabuki score’ of 4.5). The distribution of the ‘MLL2-Kabuki score’ for MLL2 mutation positive cases was significantly different from MLL2 mutation negative cases (tested by Mann–Whitney U test; p < 0.0014). The latter remained significant even when corrected for cases for which many (>50%) of facial or total features were not assessed (p < 0.001).

Discussion

In this study, a predominance of loss of function MLL2 mutations were identified in KS patients consistent with previous reports, further confirming haploinsufficiency of the MLL2 gene as the molecular mechanism in KS. In addition, 33 of the mutations reported in our series are novel. Phenotypic studies reported here are in concordance with previous studies concerning the presence or absence of some of the most characteristic KS facial features in MLL2 mutation positive and negative patients (7, 9). Interestingly, these differences were observed both when comparing all patients with likely pathogenic MLL2 mutations or only those with nonsense mutations vs patients without MLL2 mutations, suggesting that (i) most of these mutations are pathogenic, and that (ii) the pathogenic mechanism of haploinsufficiency of MLL2 is the cause of KS in these patients. We thus suspect that the pathogenic missense mutations found in patients with KS are loss of function mutations. A more detailed delineation of the facial phenotype allowed more accurate differentiation in both groups of patients. Among the other clinical features developmental abnormalities, short stature and cardiac anomalies have also been found to differ in frequency between patients with and without MLL2 mutations (7, 9). In order to see whether one could predict MLL2 mutation status in the entire cohort, we considered the 10 phenotypic features (based on Table S6: flat nasal tip, intellectual disability, arched eyebrows, broad nasal root, thin upper and full lower lip, short stature, lax joints, blue sclerae, frequent infections, and large dysplastic ears) that best distinguish patients with and without MLL2 point mutations. Mutation positive cases have on average 7/10 of these features, while MLL2 mutation negative cases have on average only 4/10 of these features. Whether one can use the combination of these phenotypic features prior to future genetic testing of MLL2 remains to be seen.

Our finding that a significant number of patients with KS or a Kabuki-like phenotype do not carry an MLL2 mutation, suggests that there is at least one other gene (or functional genomic element) causing a syndrome similar to MLL2 mutation-positive KS. During the preparation of this manuscript, a new X-linked gene KDM6A was described (28), CNVs of which explain the Kabuki phenotype in individual additional cases. Of note, Banka et al. (29) report on novel mutation types including mosaic small deletions (indels) and (mosaic) CNVs. This study emphasizes to broaden the diagnostic detection to all possible types of genetic variants. Next to the fact that we might have missed MLL2 mutations that escaped Sanger sequencing detection one could discuss why our detection rate of MLL2 point mutations is slightly lower compared to other studies (Table S10). Despite the fact that the majority of all cases are positive for a number of typical Kabuki syndrome features [e.g. long palpebral fissures (98%), everted lower eyelids (83%), thin upper and full lower lip (80%), persistent fetal pads (88%), ID (89%) and hypotonia (80%)Table S6], it is possible that our phenotypic inclusion criteria were slightly broader compared to previous studies. While this allows to explore the phenotypic boundaries of the syndrome, this is the identification of less typical patients, this might also result in a lower detection rate.

Our study extends the extensive MLL2 mutation spectrum in KS by adding 33 novel mutations. Based upon a new clinical score, we have shown that MLL2 mutation positive patients have a more severe and typical Kabuki phenotype than the MLL2 mutation negative group (36% of KS patients reported here).

Supporting Information

The following Supporting information is available for this article:

Fig. S1. Graphic representation of the 238 pathogenic mutations of MLL2 compared to respective mutation type identified in patients with Kabuki syndrome. The enclosed table (Table S10) shows the absolute numbers for each category while the percentages inside the graph show the relative frequencies.

Table S1. Detailed overview of clinical features of each individual case presented in this study
Table S2. List of the nonsense and frameshift mutations identified in this study
Table S3. List of pathogenic missense mutations identified in this study
Table S4. List of pathogenic splice-site mutations identified in this study
Table S5. Variants of unknown significance and novel benign variants
Table S6. Table comparing the clinical features identified in the patients with and without MLL2 mutations. The features are sorted by their frequency in all the patients in descending order. The first column describes the sign, the second the total number of patients for which a clinical description was provided and the third column the percentage of patients that presented each sign. The next columns show how many patients with and without mutations had (phen +) or did not have (phen −) the interrogated phenotype. The last column shows the p value (calculated by the Fisher’s exact test) of the difference among the two groups.

Table S7. Table comparing the clinical features identified in the patients with nonsense and without MLL2 mutations. The table follows the structure and sorting of Table S5
Table S8. Clinical scoring list to quantify Kabuki phenotypes
Table S9. Summary of all the mutations reported as pathogenic in the literature and the mutations identified in this study
Table S10. Comparison and summary of all the mutations reported as pathogenic in the literature

Additional Supporting information may be found in the online version of this article.
MLL2 mutation detection in 86 patients with Kabuki syndrome

Acknowledgements

This study was financially supported by the Netherlands Organization for Health Research and Development (ZonMW grants 917-66-36 and 911-08-025 to J. A. V.; 917-86-319 to B. B. A. d. V., and 916-12-095 to A. H.), the EU-funded TECHGENE project (Health-F5-2009-223143 to J. A. V.), the AnEUUploidy project (LSHG-CT-2006-37627), the SNF and Gebert grants (to S. E. A.), the Australian NHMRC (Overseas Training Fellowship-512123 to T. R); a grant from the Bodossaki Foundation (to P. M.) and the Swedish Medical Research Council (B.-M. A., A. N.).

References

12. Douglas J, Hanks S et al. NSD1 mutations are the major cause of Sotos syndrome and occur in some cases of Weaver syndrome but are rare in other overgrowth phenotypes. Am J Hum Genet 2003: 72 (1): 132–143.

545