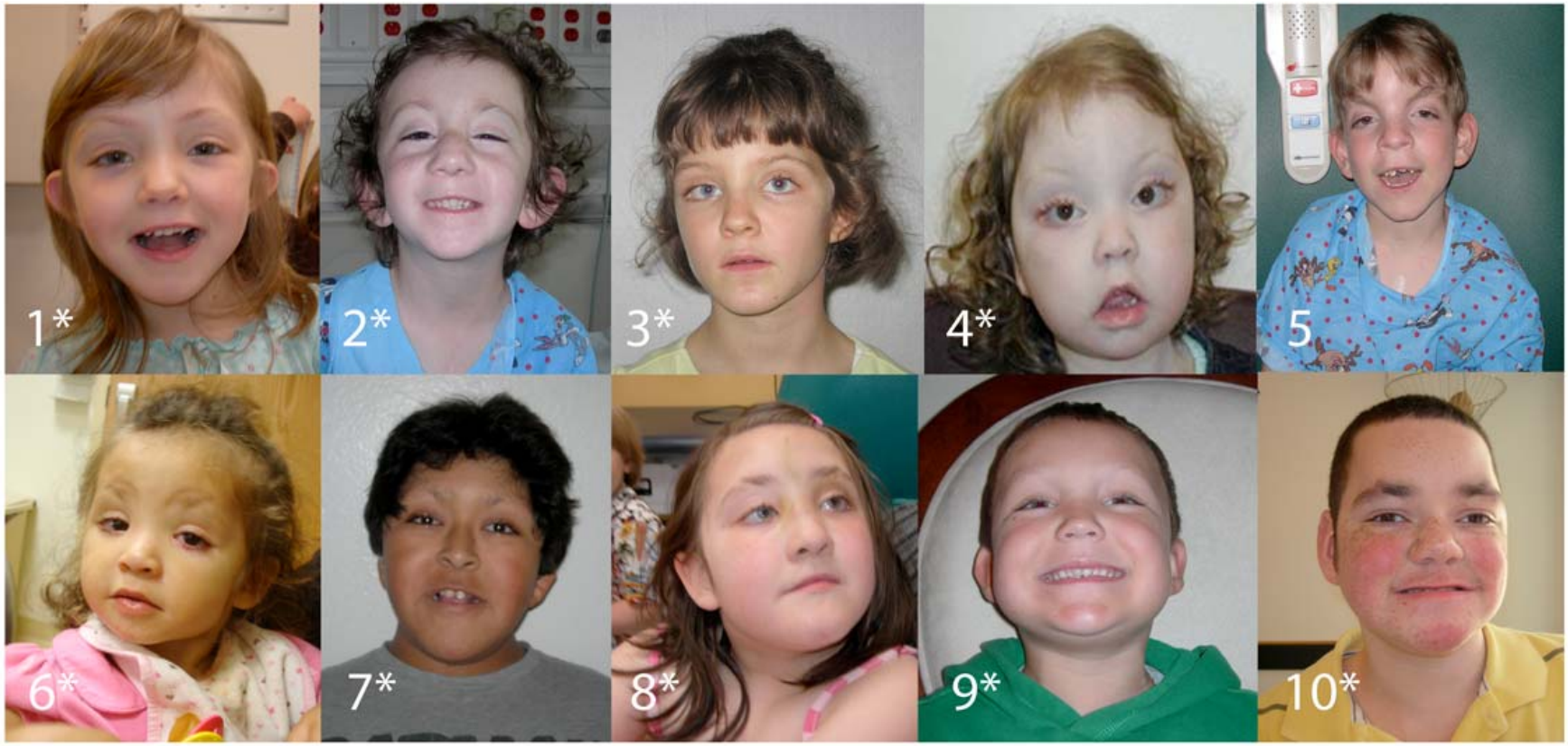
Exome sequencing identifies *MLL2* mutations as a cause of Kabuki syndrome

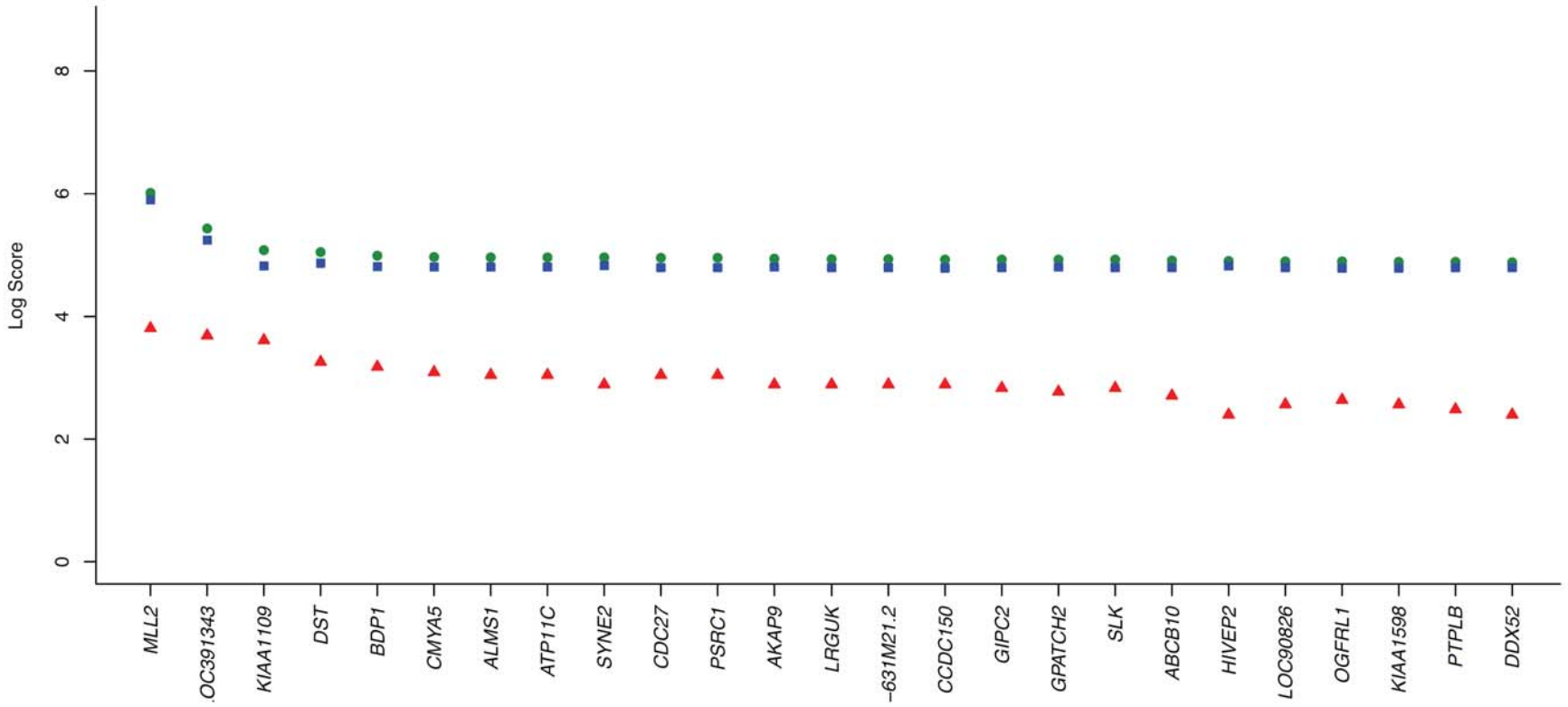
Sarah B. Ng[1*], Abigail W. Bigham[2*], Kati J. Buckingham[2], Mark C. Hannibal[2,3], Margaret McMillin[2], Heidi Gildersleeve[2], Anita E. Beck[2,3], Holly K. Tabor[2,3], Greg M. Cooper[1], Heather C. Mefford[2], Choli Lee[1], Emily H. Turner[1], Josh D. Smith[1], Mark J. Rieder[1], Koh-ichiro Yoshiura[4], Naomichi Matsumoto[5], Tohru Ohta[6], Norio Niikawa[6], Deborah A. Nickerson[1], Michael J. Bamshad[1,2,3†], Jay Shendure[1†]

Departments of [1]Genome Sciences and [2]Pediatrics, University of Washington, Seattle, Washington, USA. [3]Seattle Children's Hospital, Seattle, Washington, USA. [4]Department of Human Genetics, Nagasaki University Graduate School of Biomedical Sciences, Nagasaki, Japan. [5]Department of Human Genetics, Yokohama City University Graduate School of Medicine, Yokohama, Japan. [6]Research Institute of Personalized Health Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan

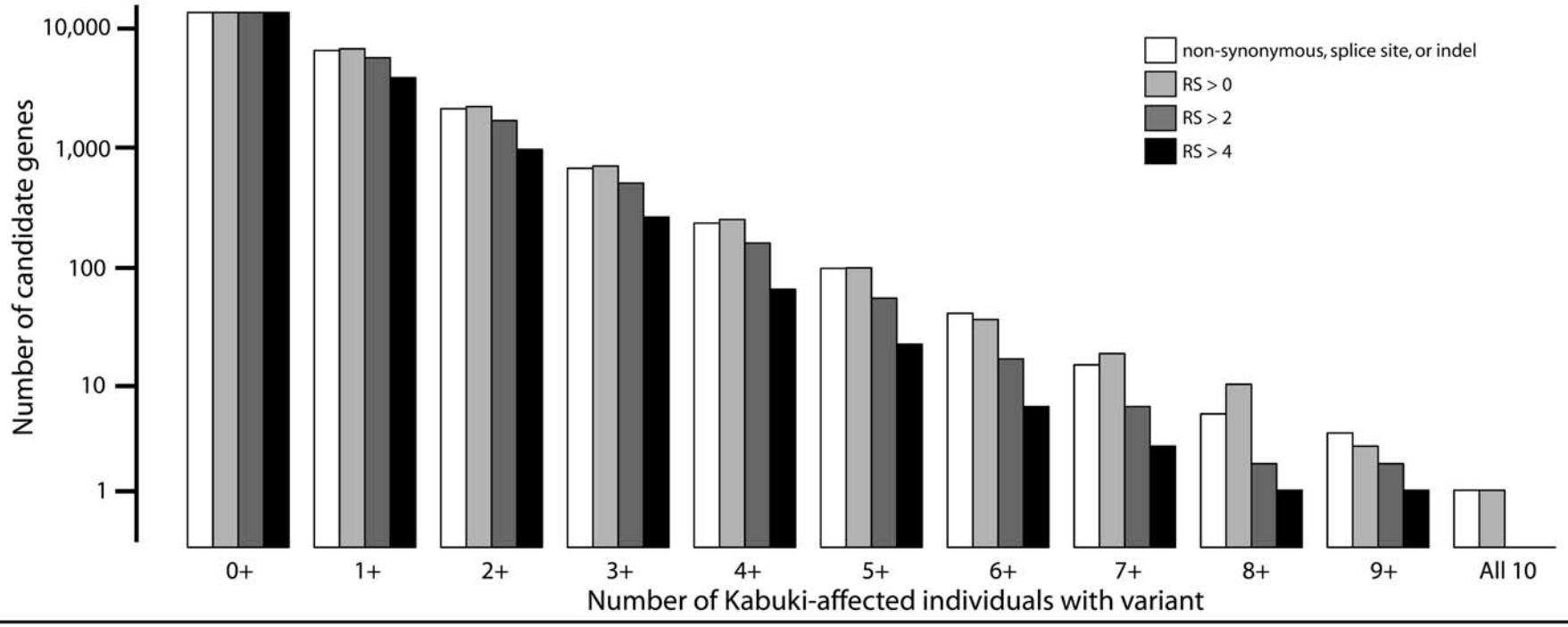
Supplementary Figures 1-3 and Supplementary Tables 1-3



Supplementary Figure 1. Photographs of the facial characteristics used to determine the subjective ranking of Kabuki phenotypes. The phenotype ranking of the ten children with Kabuki syndrome is listed here from 1-10 based on similarity to the canonical phenotype of Kabuki syndrome. Asterisks indicate cases in which *MLL2* mutations were identified. Informed consent was obtained for publication of each of the facial photos shown.



Supplementary Figure 2. Ranking of candidate genes identified by exome sequencing based on shared variants and functional annotation. Plot of the log of the top twenty-five candidate gene scores (green circles) in 10 Kabuki cases ranked by sum of case score (red triangles) and variant score (blue squares).



MLL2 rank :	3,732	3,732	1,232	399	136	47	14	6	3	NA	NA
Total (RS > 0) :	14,407	7,176	2,360	754	269	106	39	20	11	3	1

Supplementary Figure 3. Constraint scores enrich candidate gene pool for variants causing Kabuki syndrome. Number of candidate genes in which at least the given number of individuals with Kabuki syndrome has a rare variant that is functionally defined or with RS scores as indicated. Total number of candidate genes identified at RS > 0 and rank of *MLL2* among those genes are also given.

Supplementary Table 1. Clinical characteristics used to determine the subjective phenotype ranking of the 10 children with Kabuki syndrome.

	Cardiovascular	Spleen/Liver abnormality	Kidney abnormality	Kidney dysfunction	Hearing loss	Preauricular pits/tags	Cleft palate	High arched palate	Hypotonia	Developmental delay
1	ASD/VSD, aortic coarctation, bicuspid valves, dysrhythmia	np	X	np	X	X	X	X	X	X
2	aortic coarctation, bicuspid valves	-	X	X	X	X	X	X	X	X
3	ASD/VSD, aortic coarctation, bicuspid valves	-	-	-	X	np	np	X	X	X
4	VSD, aortic coarctation	X	np	np	X	-	X	X	X	X
5	ASD, VSD	np	np	np	np	np	X	X	X	X
6	np	np	np	np	-	np	np	X	X	X
7	np	-	X	-	X	np	np	X	X	X
8	np	X	X	np	np	X	X	X	X	X
9	np	np	np	np	np	np	np	np	np	X
10	ASD, VSD, dysrhythmia	X	X	X	X	np	np	X	X	X

X denotes abnormality present
np denotes abnormality not present
- denotes data not found in medical record

Supplementary Table 2. Lists of certain candidate genes as identified in Table 1 and 2.

Gene names given for cells in Table 1 where number of genes < 20

A. Subset analysis (any *x* of 10)	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
NS/SS/I	12,042	8,722	7,084	6,049	5,289	4,581	3,940	3,244	2,486	1,459
Not in dbSNP129 or 1000 genomes	7,419	2,697	1,057	488	288	192	128	88	60	34
Not in control exomes	7,827	2,865	1,025	399	184	90	50	22	7 MUC2, LIN37, MUC6, AHNAK2, UGT2B10, MUC16, MUC5B	2 MUC16, AHNAK2
Not in either	6,935	2,227	701	242	104	44	16 AHNAK2, MUC2, EVPL, LOC100128468, MUC6, MLL2, PCDHGB1, DSPP, LOC391343, MYCBP2, PCDHGA3, PCDHGA2, MUC16, PCDHGA1, LYST, PKHD1L1	6 MUC2, MUC6, DSPP, MUC16, AHNAK2, LYST	3 MUC16, AHNAK2, MUC2	1 MUC16
Is loss-of-function (nonsense / frameshift indel)	753	49	7 FLJ34443, GJB4, MLL3, MLL2, SQLE, LOC391343, ZNF598	3 MLL2, ZNF598, FLJ34443	2 MLL2, ZNF598	2 MLL2, ZNF598	1 MLL2	0	0	0

Supplementary Table 2 continued. Lists of certain candidate genes as identified in Table 1 and 2.

Gene names given for cells in Table 2 where number of genes < 20

B. Sequential analysis	**1**	**+ 2**	**+ 3**	**+ 4**	**+ 5**	**+ 6**	**+ 7**	**+ 8**	**+ 9**	**+ 10**
NS/SS/I	5,282	3,850	3,250	2,354	2,028	1,899	1,772	1,686	1,600	1,459
Not in dbSNP129 or 1000 genomes	687	214	145	84	63	54	42	40	39	34
Not in control exomes	675	134	50	26	13 OBSCN, LIN37, MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, PCDHGA4, MUC2, MUC6, UGT2B10, AHNAK2, PKHD1L1	13 OBSCN, LIN37, MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, PCDHGA4, MUC2, MUC6, UGT2B10, AHNAK2, PKHD1L1	8 OBSCN, LIN37, MUC16, MUC2, MUC6, UGT2B10, AHNAK2, PKHD1L1	5 LIN37, MUC16, MUC2, MUC6, AHNAK2	4 MUC16, MUC2, MUC6, AHNAK2	2 MUC16, AHNAK2
Not in either	467	89	34	18 SRRM2, ATG2B, VPS13A, FNDC1, MLL2, MUC4, UBR1, MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, LYST, A26C1A, MUC2, MUC6, AHNAK2, PKHD1L1	9 MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, MUC2, MUC6, AHNAK2, PKHD1L1	8 MUC16, PCDHGB1, PCDHGA2, PCDHGA3, PCDHGA1, MUC2, MUC6, AHNAK2	4 MUC16, MUC2, MUC6, AHNAK2	4 MUC16, MUC2, MUC6, AHNAK2	3 MUC16, MUC2, AHNAK2	1 MUC16
Is loss-of-function (nonsense / frameshift indel)	25	1 MLL2	1 MLL2	1 MLL2	0	0	0	0	0	0

Supplementary Table 3. Annotation of all *MLL2* mutations found in 53 Kabuki cases screened.

Kindred	Indiv	Exome Sequenced	Mutation	Exon	Predicted Amino Acid Change	Confirmed as *de novo*	Position[a]
1		yes	c.G15195A	48	p.W5065X	+	chr12:47706821
2		yes	c.C6010T	28	p.Q2004X	+	chr12:47722238
3		yes	c.C12697T	39	p.Q4233X	+	chr12:47712058
4		yes	c.C8488T	34	p.R2830X	-	chr12:47718918
5		yes	--	--	--	--	--
6		yes	c.11794_11797delCAAC	39	p.Q3932SfsX46	+	chr12:47712958-61
7		yes	c.T15618G	48	p.Y5206X	+	chr12:47706398
8		yes	c.3585_3586insA	11	p.P1196TfsX11	+	chr12:47730053-54
9		yes	c.C6295T	31	p.R2099X	+	chr12:47721525
10		yes	c.6595delT	31	p.Y2199IfsX65	+	chr12:47721225
11		no	c.G15326T	48	p.C5109F	+	chr12:47706690
12		no	c.G15536A	48	p.R5179H	+	chr12:47706480
13		no	c.C11149T	39	p.Q3717X	+	chr12:47713606
14		no	c.15444_15445delTT	48	p.F5149CfsX9	-	chr12:47706571-72
15		no	c.C15217T	48	p.Q5073X	-	chr12:47706799
16		no	c.C9961T	34	p.R3321X	-	chr12:47717445
17		no	c.C14710T	48	p.R4904X	-	chr12:47707306
18		no	c.5875_5891dup17	28	p.E1965GfsX88	-	chr12:47722341-57
19		no	c.G15536A	48	p.R5179H	-	chr12:47706480
20		no	c.C12703T	39	p.Q4235X	-	chr12:47712152
21		no	c.C12241T	39	p.Q4081X	-	chr12:47712514
22		no	c.C13390T	39	p.Q4464X	+	chr12:47711365
23		no	c.G15641A	48	p.R5214H	-	chr12:47706375
24		no	--	--	--	--	--
25	1	no	c.A13580T	40	p.K4527X	*	chr12:47711035
	2	no	c.A13580T	40	p.K4527X	-	chr12:47711035
26		no	c.C16501T	53	p.R5501X	-	chr12:47702113
27	1	no	--	--	--	--	--
	2	no	--	--	--	--	--
28		no	--	--	--	--	--
29		no	c.C10738T	38	p.Q3580X	-	chr12:47714119
30		no	--	--	--	--	--
31		no	c.C16360T	52	p.R5454X	-	chr12:47702382
32		no	--	--	--	--	--
33		no	--	--	--	--	--
34		no	--	--	--	--	--
35		no	c.4956_4957insG	19	p.E1654X	-	chr12:47724800-01
36		no	c.10599_10630del32	38	p.V3534QfsX11	-	chr12:47714227-58
37		no	--	--	--	--	--
38		no	c.C13606T	40	p.R4536X	-	chr12:47711008
39		no	c.G16019T	50	p.R5340L	-	chr12:47704661
40		no	c.C4843T	19	p.R1615X	-	chr12:47724914
41		no	c.C16391T	52	p.T5464M	-	chr12:47702351
42		no	--	--	--	--	--
43		no	--	--	--	--	--
44		no	--	--	--	--	--
45		no	c.C16360T	52	p.R5454X	-	chr12:47702382
46		no	--	--	--	--	--
47		no	--	--	--	--	--
48		no	--	--	--	--	--
49		no	--	--	--	--	--
50		no	c.1324delC	10	p.P442HfsX487	-	chr12:47732409
51		no	--	--	--	--	--
52	1	no	c.C16391T	52	p.T5464M	*	chr12:47702351
	2	no	c.C16391T	52	p.T5464M	-	chr12:47702351
53		no	--	--	--	--	--

-- no mutation identified
\+ confirmed *de novo*
\- no parental samples available
X stop codon
fs frameshift
[a] chromosomal position was determined using March 2006 assembly from UCSC (hg18)
* confirmed as inherited

Kindreds 25, 27 and 52 show dominant transmission of Kabuki syndrome from parent to child. Both affected individuals are listed here.