Supplementary Data

Copy number gain at Xp22.31 includes complex duplication rearrangements and recurrent triplications

Pengfei Liu, Ayelet Erez, Sandesh C. Sreenath Nagamani, Weimin Bi, Claudia M. B. Carvalho, Alexandra D. Simmons, Joanna Wiszniewska, Ping Fang, Patricia A. Eng, M. Lance Cooper, V. Reid Sutton, Elizabeth R. Roeder, John B. Bodensteiner, Mauricio R. Delgado, Siddharth K. Prakash, John W. Belmont, Pawel Stankiewicz, Jonathan S. Berg, Marwan Shinawi, Ankita Patel, Sau Wai Cheung, James R. Lupski

Table S1. Detailed clinical data specifically for each patient.

	Recurrent Duplications														Recurrent Triplications		
BAB numbers	2829	2830	2831	2835	2843	2844	2845	2937	3081[a]	3082[a]	2814[b]	2815[b]	2846[c]	2862	2822	2817	2828
Sex	M	M	M	F	F	M	M	M	M	M	M	M	F	M	M	M	F
Inheritance	NA	NA	NA	NA	Mat	NA	Mat	Mat	NA	NA	Mat	Mat	NA	NA	Mat	Mat	DN
XCI result	NA	NA	NA	Rdm	Skw	NA	NA	NA	NA	NA	NA	NA	NI	NA	NA	NA	Rdm
Additional CNV	+	-	+	+	-	+	-	-	-	-	-	-	-	+	-	-	+
Birth defects	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-
GER	-	-	+	-	-	+	-	-	+	-	-	-	-	-	-	+	-
Delay- gross motor	-	-	-	+	+	+	-	+	+	+	+	+	-	-	+	+	+
Delay- Speech	-	-	-	+	+	+	-	+	+	+	+	+	-	+	+	+	+
Developmental regression	-	-	-	-	+	-	-	-	-	-	-	+	-	-	-	-	-
Seizures	-	-	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-
Hypotonia	-	-	-	+	-	-	-	+	+	+	-	-	-	-	-	+	+
Autism Spectrum	-	+	-	-	-	-	+	+	+	+	+	+	-	-	-	-	-
ADHD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	-
CHD	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Macrocephaly >95%	+	-	-	-	-	-	+	-	-	-	+	+	-	-	-	-	+
Microcephaly <5%	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-
Short stature	+	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-
Facial dysmorphism	-	-	-	+	-	-	+	-	-	-	-	-	-	-	+	-	-
Retinal or eye abnormalities	-	-	-	+	-	-	-	-	-	-	-	-	-	-	+	-	-
MRI/CT brain abnormalities	-	-	-	+	-	-	+	-	-	-	-	+	-	-	-	-	-
EEG abnormalities	-	-	+	+	-	-	-	+	-	-	-	-	-	-	-	-	-

a. BAB3081 and BAB3082 are twins. b. BAB2814 and BAB2815 are brothers. c. BAB2846 had maternal alcohol and drug exposure along with anti-seizure medication history during pregnancy.

Abbreviations: “+”, feature is present; “-”, feature is not present or not known; GER, gastro-esophageal reflux; ADHD, attention-deficit hyperactivity disorder; CHD, congenital heart disease; NA, not available; Mat, maternal; DN, *de novo*; Rdm, random XCI; Skw, skewed XCI; NI, non-informative.

Table S2. X-inactivation studies in 23 patients and five mothers with Xp22.31 duplications

	Recurrent dup		Recurrent trip		Nonrecurrent dup	
	Mother(s)	Patient(s)	Mother(s)	Patient(s)	Mother(s)	Patient(s)
Random	1	12	1	1	0	4
Skewed	0	4	0	0	1	0
Non-informative	0	1	2	0	0	1

Skewed XCI is defined as the inactivation ratio greater than 80:20. The "nonrecurrent dup" category includes both the subjects with simple nonrecurrent duplication and the subjects with complex rearrangements. The four patients with recurrent duplication who showed skewed XCI are BAB2840, BAB2843, BAB2863 and BAB2936. It is unclear which X chromosome is preferentially inactivated due to lack of parental DNA samples. The one carrier mother of the nonrecurrent duplication who showed skewed XCI is the mother of BAB2824. In this subject, the X chromosome with duplication is preferentially activated. Abbreviations: dup, duplication; trip, triplication.

Table S3. Additional chromosomal CNV findings in the subjects with the recurrent Xp22.31 duplication or triplication.

BAB numbers	Type	Maximum size	Minimum size	Minimum coordinate	Additional CNV directly contributing to phenotype
2829	Gain	6.43 Mb	134 kb	Chr3:110,195,502-110,329,344	No
2831	Loss	3.91 Mb	626 kb	Chr8:305,404-931,391	No
2835	Loss	94.2 kb	86.2 kb	Chr8:9,945,886-10,032,091	No
2841	Gain	2.50 Mb	181 kb	Chr4:188,678,418-188,859,354	No
2844	Loss; Gain	929 kb; 1.20 Mb	630 kb; 934 kb	Chr14:83,957,837-84,587,455 Chr21:27,204,228-28,137,956	No
2846	Gain	47.1 Mb	38.4 Mb	Chr9:360,184-38,713,752	Yes
2850	Gain	657 kb	402 kb	Chr7:152,619,353-153,021,406	No
2851	Loss	244 kb	187 kb	Chr12:13,577,342-13,764,730	No
2863	Loss; Gain	8.23 Mb; 19.94 Mb	297 kb; 2.34 Mb	Chr2:110,050,644-110,347,650 Chr16:45,675,606-48,023,182	No
2864	Loss	3.52 Mb	535 kb	chr7:75,048,098-75,583,095	No
2828	Gain	6.70 Mb	3.99 Mb	chr7:149,041,960-153,029,489	Yes

Table S4. MLPA probe sequences.

Name	Type[a]	Sequence[b]
KAL1-104-L	XR	GGGTTCCCTAAGGGTTGGATACCTGAACCAATGCAAAGGCCATCCAGTGA
KAL1-104-R	XR	AGGGCTGCCATGGTTCCTCAACGTTGGAAGTTCTAGATTGGATCTTGCTGGCAC
ACTA2-108-L	AR	GGGTTCCCTAAGGGTTGGATTAGTGCGGTAGGACAGAGCCTGGATGTTCTAC
ACTA2-108-R	AR	CATGGCCTAGTTTCTTGTTCAGCAGGGACACAGTCTAGATTGGATCTTGCTGGCAC
HDHD1A-112-L	XT	GGGTTCCCTAAGGGTTGGATGCTGTCTAATGTTCTCCGGGGTCACTGAAATGTG
HDHD1A-112-R	XT	ATCGGCCTAAACTCCTCCAGCCAAAGAAGCACCTTTCTAGATTGGATCTTGCTGGCAC
PNPLA4-116-L	XT	GGGTTCCCTAAGGGTTGGATCTAAGCGCTATGAGGTAGGTGGATTGTAGGGAGCTT
PNPLA4-116-R	XT	GAATTGATATAGGCTGTGGGAAGCTCCTGGGTGAGAATCTAGATTGGATCTTGCTGGCAC
TERT-120-L	AR	GGGTTCCCTAAGGGTTGGAGTCTACCATTCAGCCGGCAAACACCTCAGAAACATACCA
TERT-120-R	AR	GATGAGACCTCTGAACAAACAATCCCTGCTCCCCACTCTTCTAGATTGGATCTTGCTGGCAC
VCX3Ap-124-L	XT	GGGTTCCCTAAGGGTTGGAGTCCTGAGAGCTTCCAAATCCTCCAGACAGGTGAAAGTGTT
VCX3Ap-124-R	XT	CCTCCAGTTTGCATCCACCACCACATGTTTTATGCCTCCTCTCTAGATTGGATCTTGCTGGCAC
B2M-128-L	AR	GGGTTCCCTAAGGGTTGGACAAGATGGTTACCAAGACTGTTGAGGACGCCAGAGATCTTGAG
B2M-128-R	AR	CACTTTCTAAGTACCTGGCAATACACTAAGCGCGCTCACCTTTTCTAGATTGGATCTTGCTGGCAC
VCX3Ap-132-L	XT	GGGTTCCCTAAGGGTTGGACTCCTTTTTTCTGGACTAGGGATCTTGAAGCATGGAGTTGTTGCA
VCX3Ap-132-R	XT	CTATCTTGTAACCATAAAGCCTGAGGACGAAAGCCTGGAGCCTAGTCTAGATTGGATCTTGCTGGCAC
STS-136-L	XT	GGGTTCCCTAAGGGTTGGACGCATGTGGCAAAGCTCACCATCTTCACTACAAACACGCCTGAGAGT
STS-136-R	XT	GGCACTGGGGAAACATAACTCCATCTACACCTTGGATTTGGACTGATTCTAGATTGGATCTTGCTGGCAC
GPD1-140-L	AR	GGGTTCCCTAAGGGTTGGACGAGATTGGTTTGGAGGTCCTCTCGGGGAGTTTCGGAGGTATAAAGGAA
GPD1-140-R	AR	TGCCTGGGAGATATGAGAAGCGGGCTGGTTTGAGAAGTGTGAAGAGTAGTCTAGATTGGATCTTGCTGGCAC

a. Abbreviations: XT, X chromosomal test probe; XR, X chromosomal reference probe; AR, autosomal reference probe.

b. The underlined sequences are primer-binding sequences.

Table S5. Primer sequences and PCR assays.

Breakpoint	Primer-1	Primer-2
BAB2824[a]	TTGGGACAAGCAGTTTGGAGTACATA	GGTCACCATGCCAGATACCGTCA
BAB2852[a]	GTGTAGATGTGGGAGATCCGCTGTG	AGCCCAGCTGAAGCCCTTTAGACA
BAB3087[a]	CATCACTTACCTGCCATTTCTCCCTCT	GGACCCTAGGGCATATCTGAAGTTG
BAB3090[a]	CAATGGATTTCTGGTAGGCAAACAAC	CGGGTGATGAGAATCTGAATTTGA
BAB3093[a]	ATACTAATCCAGGCAGGGCAGGC	AAGGTAGAGGAAAGTCGGCAAGGTCAG
BAB3096[a]	ATGGCCCAGTGTCTAAGCTATTTCTC	ACCCATGGCCTTTGATCCTAATCTTGA
BAB2833[a]	GTTGGAGACAGATGCAACCTCTGTG	CTTGTTTCTGTCAACCTTCTTTGCC
BAB3088[a]	TTCACCAATGCTGACGCTTCTTC	CCTCCACACCCAACTGTATTTCCTTA
BAB3091[a]	CACCACCCACACCTCATGTAATATTTT	AGCTTTGCACTGGTGTTGGGAGAA
BAB3094[a]	CATGCAATCGCCATAGTACCTGATCT	AAGCCCTTCCGGTTCAAGTCCTAGT
BAB3095_ I[a]	GGCCTGGGCTCTGTGAACAGTG	TCCTGCCATACACAAATGAAATTCC
BAB3095_ II[a]	TGACAATTCCCTTCGCAAAGTCTC	GATTTTGCCTCCTTCACAGTGGTC
BAB2861[a]	TGACGCAATCCTTATCCTCCTGT	GGGCATACACACATTCTTAGCTTTTC
BAB3084[a]	CCTGACTAATACTTTATTCTACCCATGCGA	TCCCAAGCTCCTCCTCGTCTTC
BAB3089[a]	CCCAAACAATCACAGGACAAGAAACA	AGCCTTGCAAATAAACAGGCTTTGAG
BAB2822-1.7K_to_RU2[a, b]	CAGACATATTCTGAGCAGGCAACTG	GCGTTAGTATTGTTCAAGCATGGTTC
RU2-1st PCR[c]	GAAAGGACCTGTGTATTTCTAtCGC	CCAGGTGGAGGAACCACaGAG
RU2-2nd PCR[c]	CTTTTTGTCTCCTCAGTTGCGAtT	ATGGAAGAACTACCGAGTGTGTtGAC
BAC-1st PCR[d]	GAAAGGACCTGTGTATTTCTAtCGC	CGAGGTGGAAGAACCACaGAG
BAC-2nd PCR[d]	CTTTTTGTCTCCTCAGTTGCGAtT	GTGGAAGAACCACTGACTGTGaGGAT

a: Long-range PCR was performed for these primer pairs. PCR condition included one cycle at 98°C/ 30 sec; 36 cycles at 94°C/ 60 sec, 65°C/ 20 sec, and 68°C/ 20 min; and one cycle at 68°C/ 10 min. 1M of betaine was added to facilitate the amplification. The regular *TaKaRa* PCR buffer (RR002M) was used.

b: PCR products have multiple bands. Bands were gel purified and selected for sequencing.

c: Two-step PCR was used to amplify breakpoints within the RU2 elements. The condition of the first PCR included one cycle at 94°C / 1 min; five cycles at 94°C / 30 sec, (60.5-0.5 per cycle) °C / 1 min, and 72°C / 2 min; 31 cycles at 94°C / 30 sec, 58°C / 1 min sec, and 72°C / 2 min; and one cycle at 72°C / 7 min. The Mg^{2+} concentration was 2.5 mM. The condition of the second PCR included one cycle of 94°C / 1 min; 50 cycles of 94°C / 30 sec, 64°C / 1 min, and 72°C / 100 sec; and one cycle at 72°C / 7 min. The Mg^{2+} concentration was 3 mM. *TaKaRa* GC buffer I (RR02AG) was used. The primer sequence in lower case indicates a mismatch with respect to the reference sequence.
d: Two-step PCR was used to amplify the RU2 element from the BAC DNA, RP11-527B14. PCR condition is the same as that described in c.

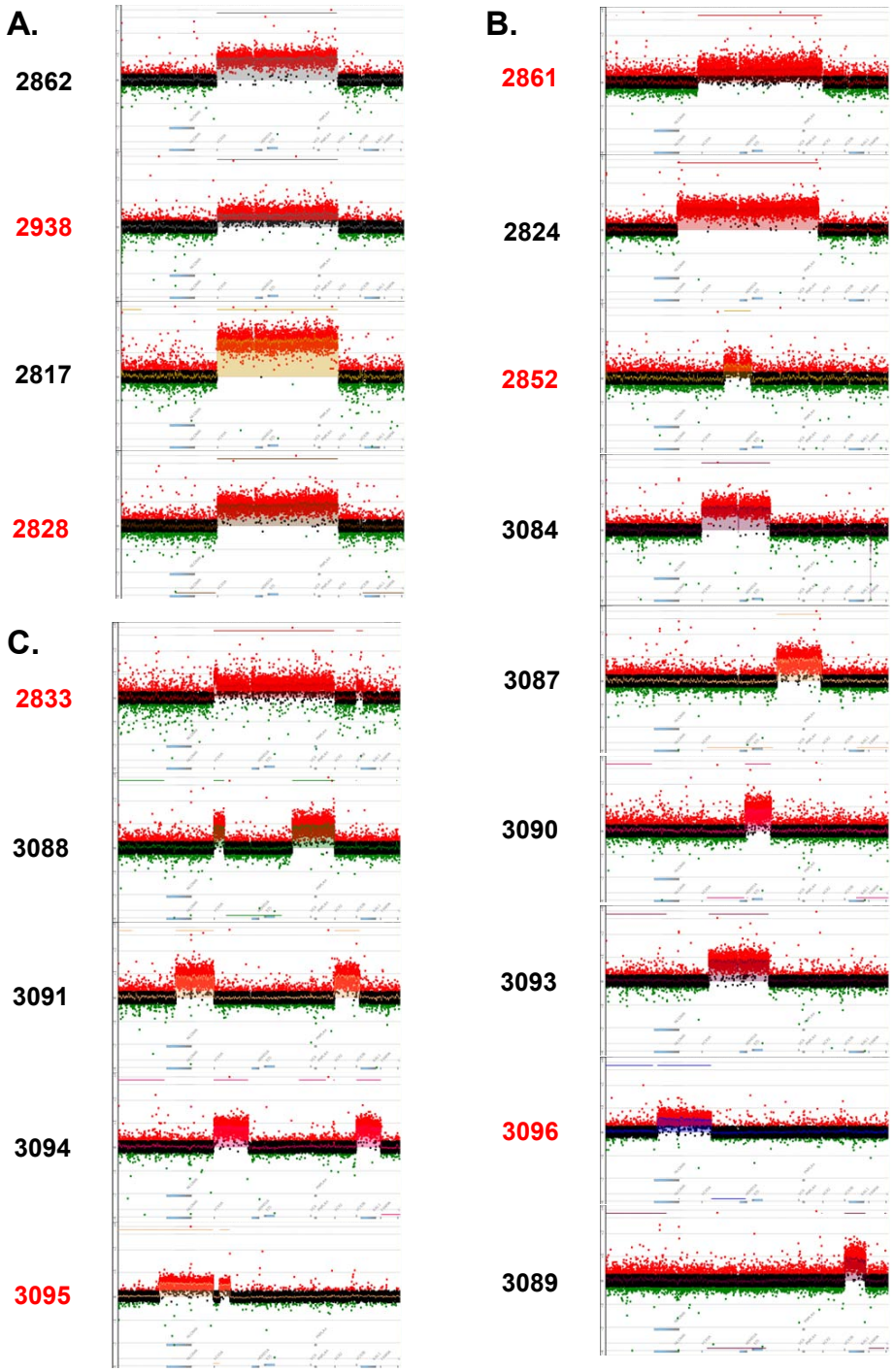

Figure S1. aCGH results of representative cases. The X-axis indicates genomic coordinates based on the reference haploid human genome sequence. The Y-axis indicates $\log_2$ ratio of subject copy number/control copy number; red dots indicate gain

with respect to CGH control. Duplication in male/female has a log ration of 1/ 0.58; triplication in male/female has a log ratio of 1.58/1. Each result is shown using the same scale. Black subject numbers indicate male and red indicate female. (A) Recurrent duplications and recurrent triplications. BAB2862 shows an example of a male with recurrent duplication. BAB2938 is an example of a female with recurrent duplication. BAB2817 is an example of a male with recurrent triplication. BAB2828 is an example of a female with recurrent triplication. (B) Array CGH results for subjects carrying nonrecurrent duplications. (C) Array CGH results for subjects carrying complex rearrangements.

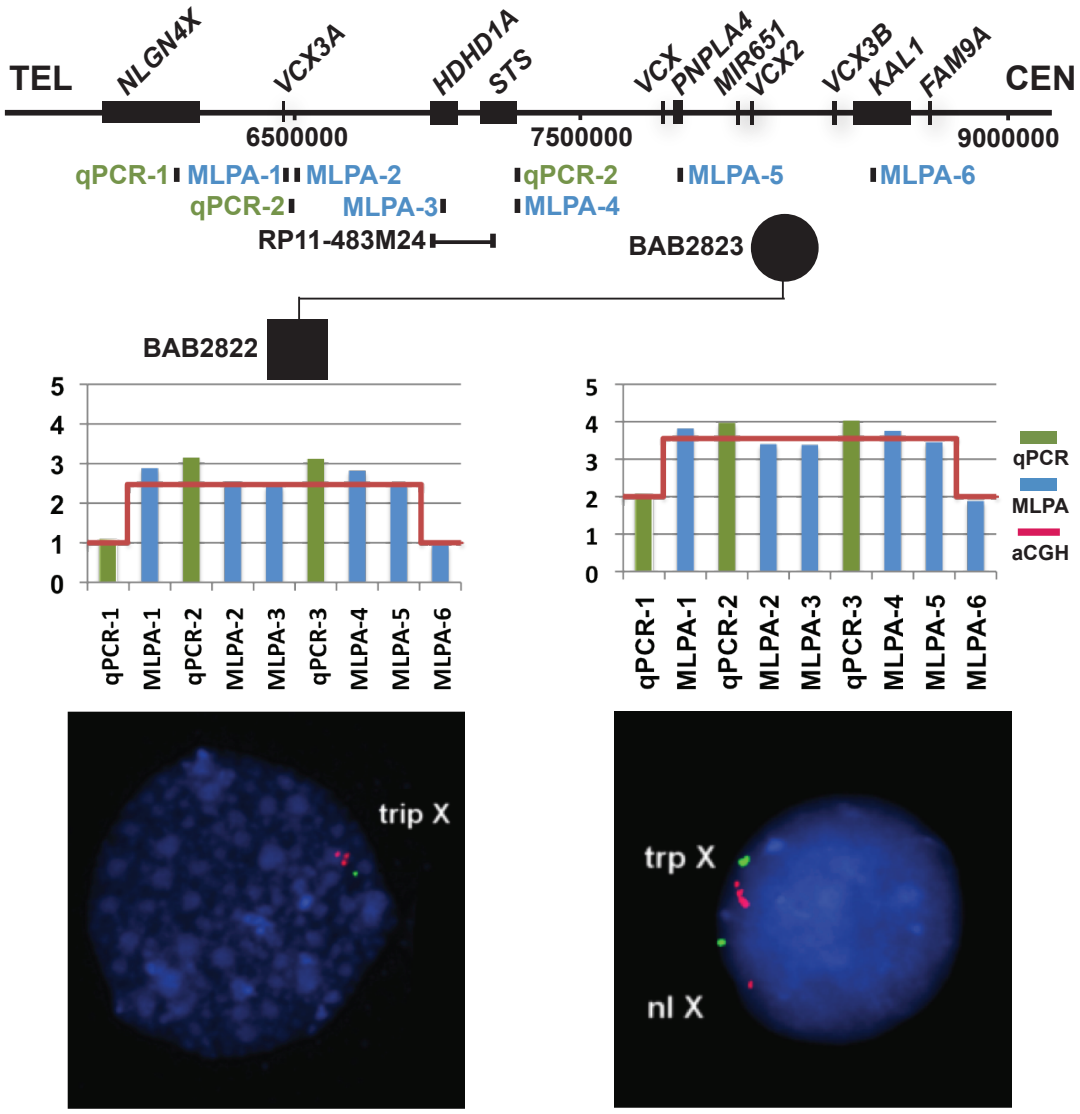

Figure S2. Ascertainment of triplications in subject BAB2822 and the mother. The figure legends are the same as Figure 2 in the Results section.

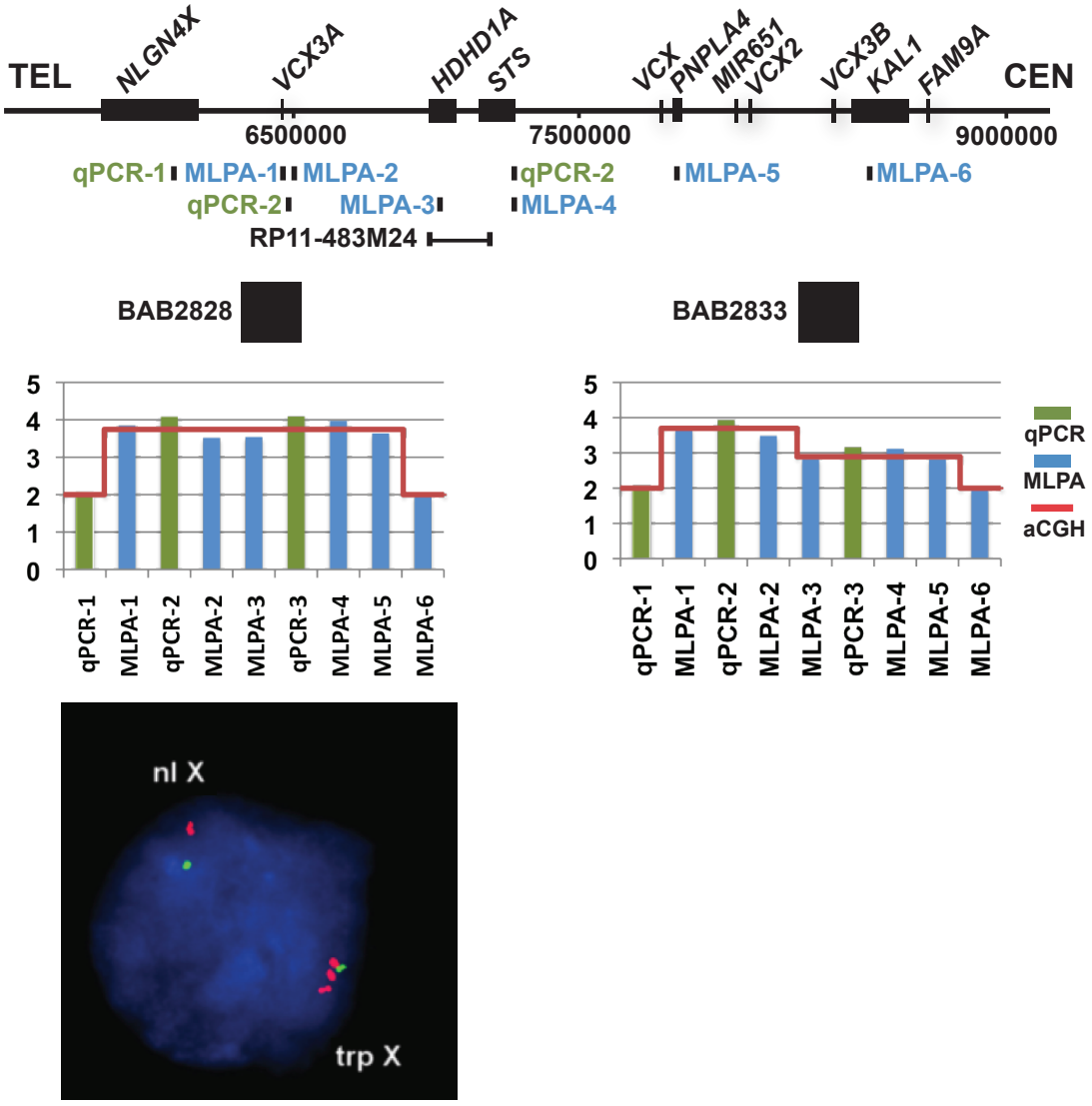

Figure S3. Ascertainment of triplications in subject BAB2828 and BAB2833. The figure legends are the same as Figure 2 in the Results section.