The American Journal of Human Genetics

Supplemental Data

Chromothripsis in Healthy Individuals Affects Multiple Protein-Coding Genes and Can Result in Severe Congenital Abnormalities in Offspring

Mirjam S. de Pagter, Markus J. van Roosmalen, Annette F. Baas, Ivo Renkens, Karen J. Duran, Ellen van Binsbergen, Masoumeh Tavakoli-Yaraki, Ron Hochstenbach, Lars T. van der Veken, Edwin Cuppen, and Wigard P. Kloosterman

Supplemental Figures

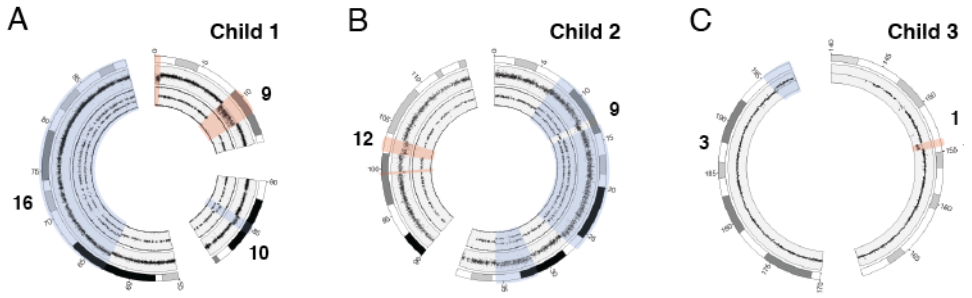

Figure S1. Circos plots displaying CNVs identified in child 1 (A), 2 (B) and 3 (C). CNVs were detected using Illumina SNParrays (CytoSNP-850K for case 1 and HumanCNV370 arrays for case 2) or custom Agilent 105k microarrays (Amadid 019015; case 3). Deletions and duplications were detected using Nexus software and are highlighted in red and blue, respectively. (A) Circos plot displaying two deletions (0.5 and 3 Mb) on chromosome 9, a 1.1 Mb duplication on chromosome 10 and a duplication of ~27 Mb on chromosome 16 in child 1. (B) In child 2, three duplications on chromosome 9 (4.5, 12 and 3 Mb) and two deletions on chromosome 12 (150 kb and 2 Mb) were detected by SNP array analysis. Previously published data, deposited in the NCBI Gene Expression Omnibus under accession number GSE37906[8]. (C) A 0.7 Mb deletion on chromosome 1 and a 2.3 Mb duplication on chromosome 3 were detected in child 3.

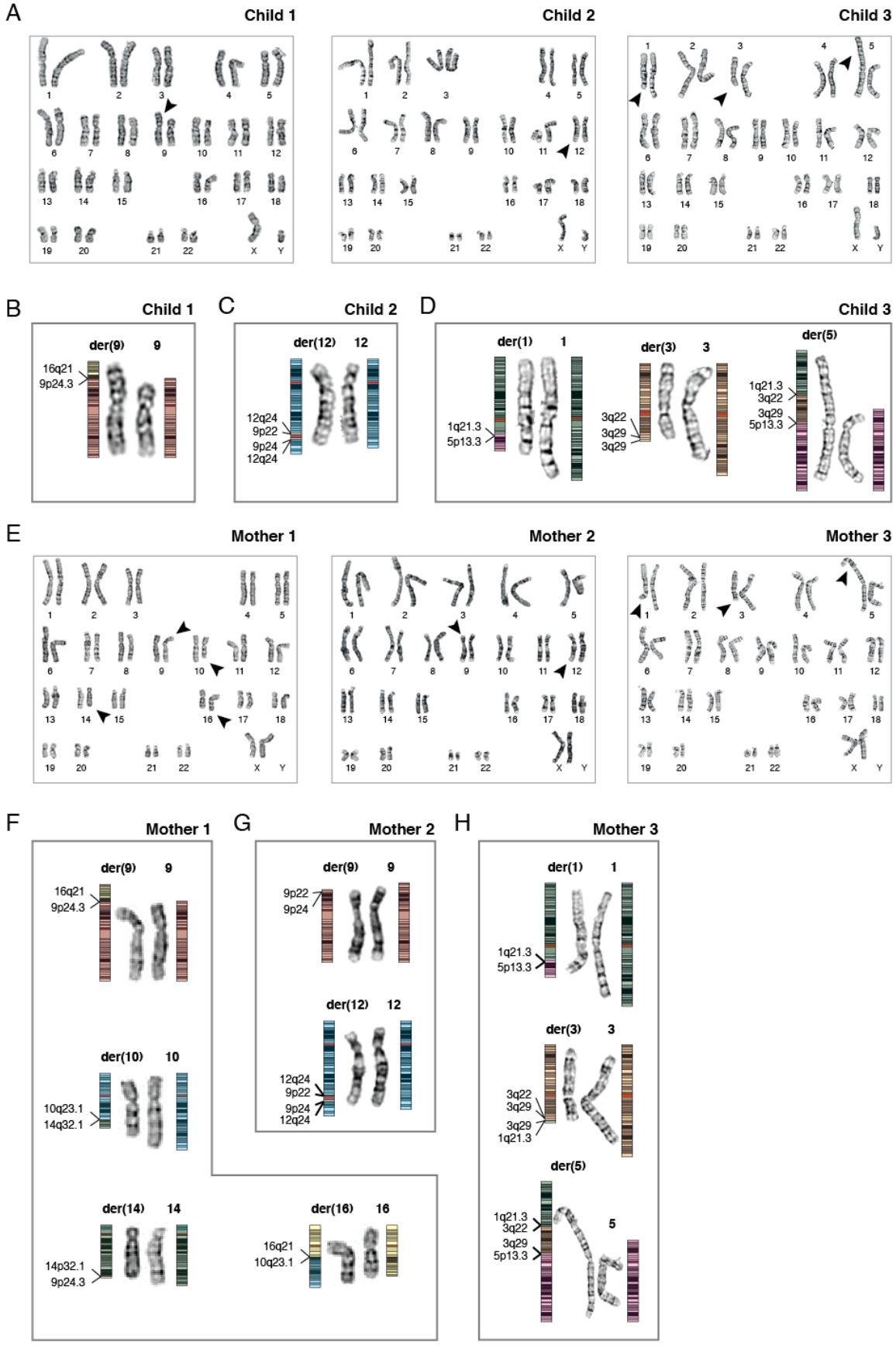
A
Child 1
Child 2
Child 3
B
Child 1
der(9) 9
16q21
9p24.3
C
Child 2
der(12) 12
12q24
9p22
9p24
12q24
D
Child 3
der(1) 1
1q21.3
5p13.3
der(3) 3
3q22
3q29
3q29
der(5)
1q21.3
3q22
3q29
5p13.3
E
Mother 1
Mother 2
Mother 3
F
Mother 1
der(9) 9
16q21
9p24.3
der(10) 10
10q23.1
14q32.1
der(14) 14
14p32.1
9p24.3
der(16) 16
16q21
10q23.1
G
Mother 2
der(9) 9
9p22
9p24
der(12) 12
12q24
9p22
9p24
12q24
H
Mother 3
der(1) 1
1q21.3
5p13.3
der(3) 3
3q22
3q29
3q29
1q21.3
der(5)
1q21.3
3q22
3q29
5p13.3
5

Figure S2. Karyograms of the children and mothers. (A) Karyogram of child 1, 2 and 3. (B-D) Partial karyograms and schematic diagrams of derivative chromosomes detected in child 1 (B), 2 (C) and 3 (D). (E) Karyogram of mother 1, 2 and 3. (F-H) Partial karyograms and schematic diagrams of the derivative chromosomes detected in mother 1 (F), 2 (G) and 3 (H).

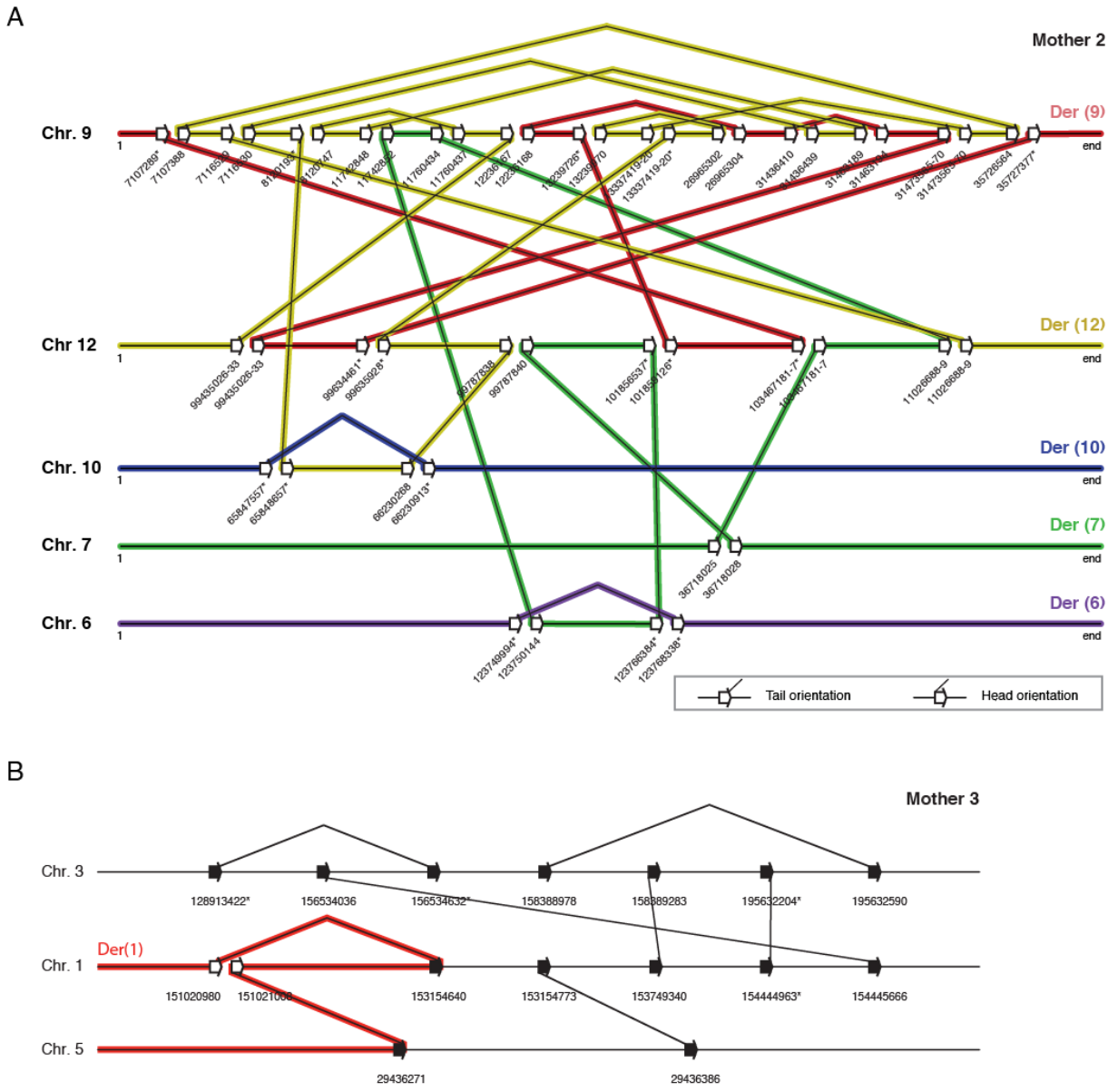

Figure S3. Schematic diagram showing the genomic positions and orientations of breakpoint junctions in mother 2 (A) and 3 (B). Sets of adjacent white arrows indicate a double-strand break (DSB), connecting lines between two arrows indicate breakpoint junctions. (A) Mate-pair sequencing revealed 23 breakpoint junctions involving five chromosomes in mother 2. Colored lines indicate predicted derivative chromosomes. The rearrangements gave rise to five derivative chromosomes der(6) (purple), der(7) (green), der(9) (red), der(10) (blue) and der(12) (yellow). Previously published data, deposited in the European Nucleotide Archive under accession number ERP001035 [8]. (B) Mate-pair sequencing revealed at least 8 breakpoint junctions involving three chromosomes in mother 3. Due to the repetitive nature of the 3q29 region involved in the rearrangements we were unable to fully reconstruct the derivative chromosomes for this individual.

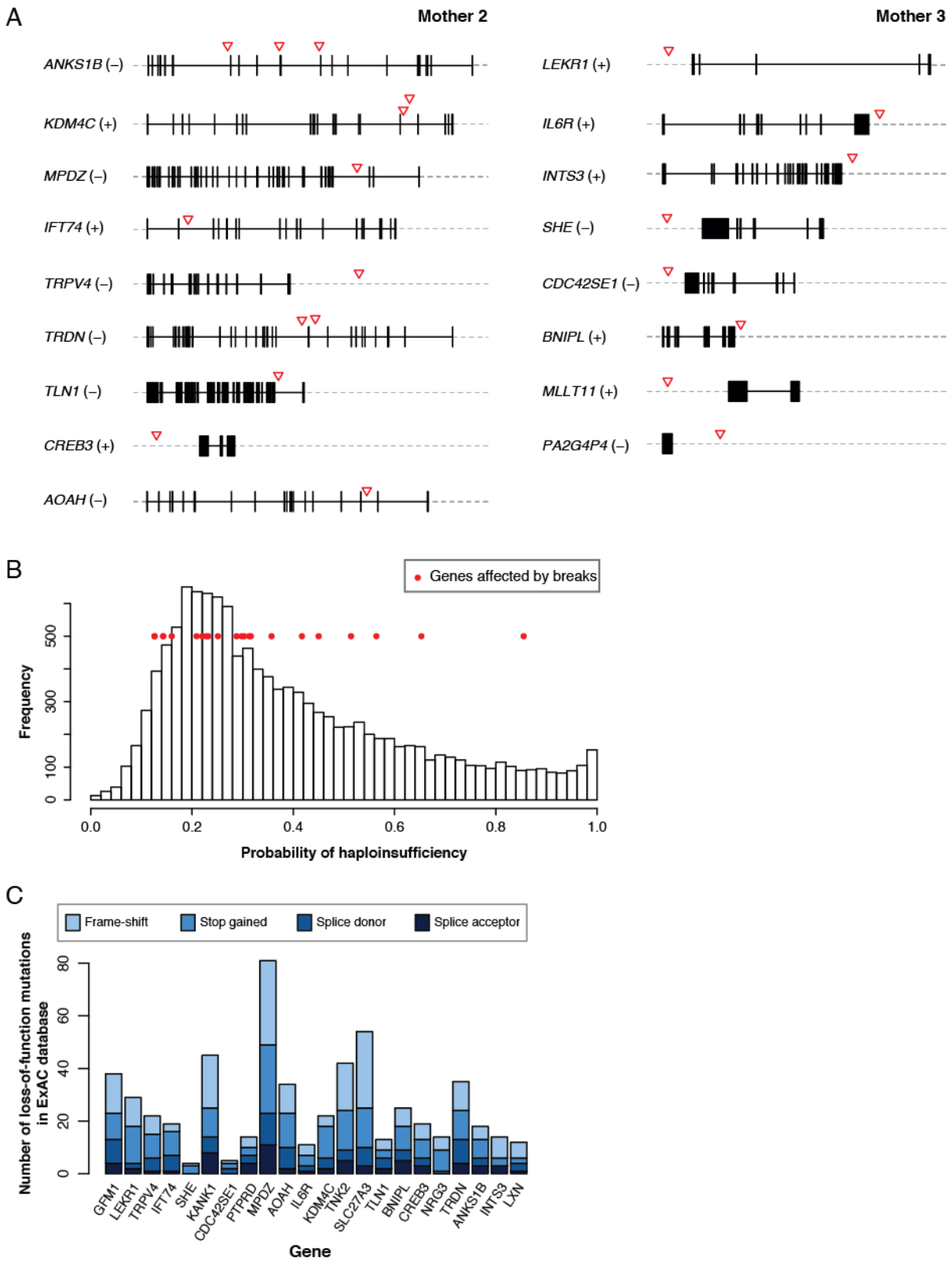

Figure S4. Chromothripsis directly affects protein-coding genes. (A) Gene structure and indication of breakpoint positions (triangles). (B) Predicted probability of haploinsufficiency for all genes affected by

chromothripsis breakpoints (red dots) relative to 12,218 genes in the genome (histogram)[17]. (C) Presence of loss-of-function mutations in genes affected by chromothripsis breakpoints. Loss-of-function mutations are derived from the exome aggregation consortium.

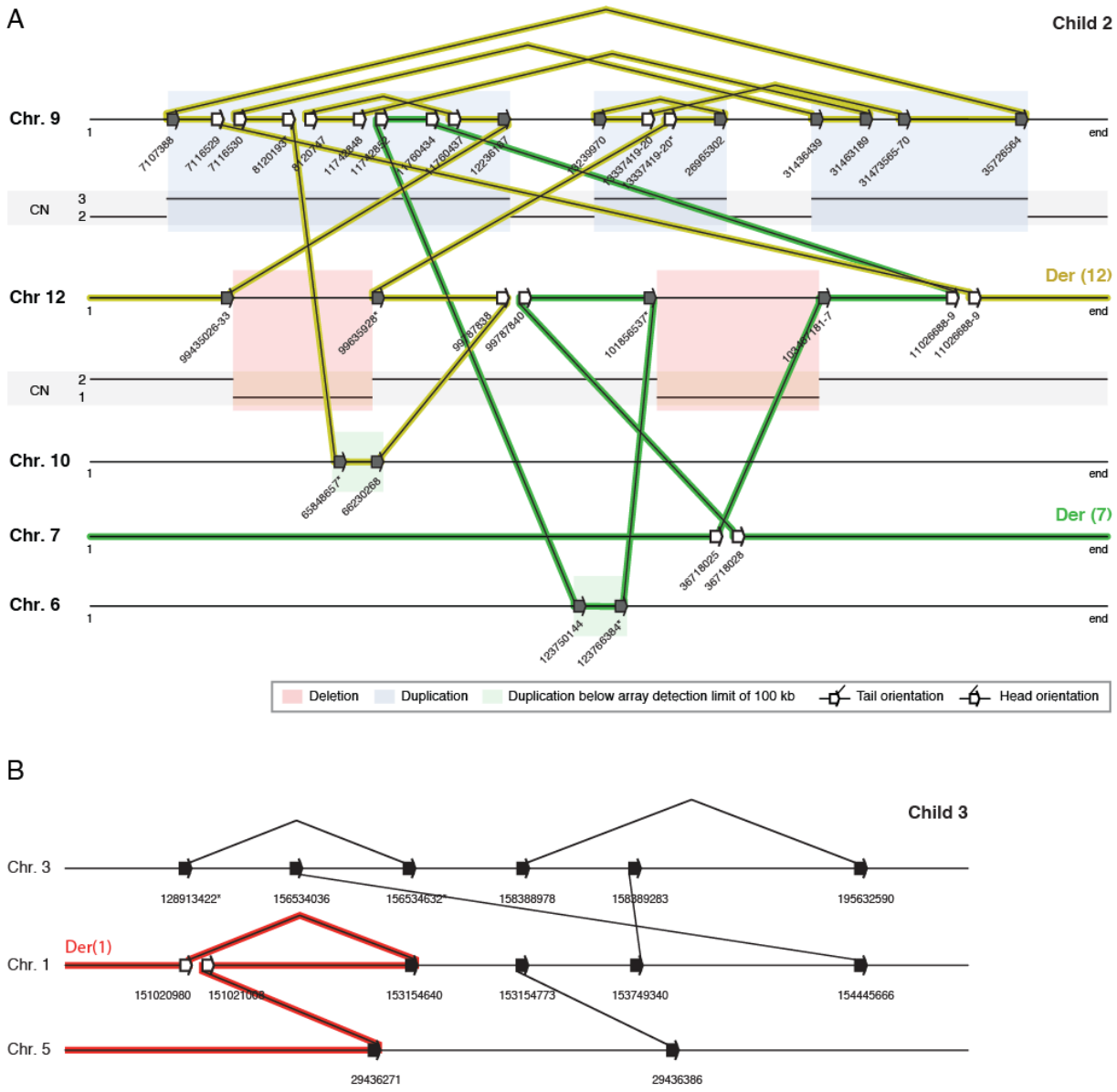

Figure S5. Rearrangements detected in child 2 and 3. Sets of adjacent white arrows indicate a double-strand break (DSB), grey arrows indicate a single end of a break that was found to be a DSB in the mother. Connecting lines between two arrows indicate breakpoint junctions. (A) Mate-pair sequencing revealed 16 breakpoint junctions in child 2. The chromosome 9 duplications can be explained by the presence of these fragments on der(12), which was inherited by the child. The deleted fragments of chromosome 12 are located on der(9), which was not inherited. Previously published data, deposited in the European Nucleotide Archive under accession number ERP001035 [8]. (B) 7 Breakpoint junctions were detected in child 3. The presence of the chromothripsis rearrangements in the mother gave rise to unstable transfer of the chromothripsis chromosomes to her child. All but one breakpoint junctions detected in the mother were also detected in her child. Due to the repetitive nature of the 3q29 region involved in the rearrangements we could not fully reconstruct the derivative chromosomes in this child.

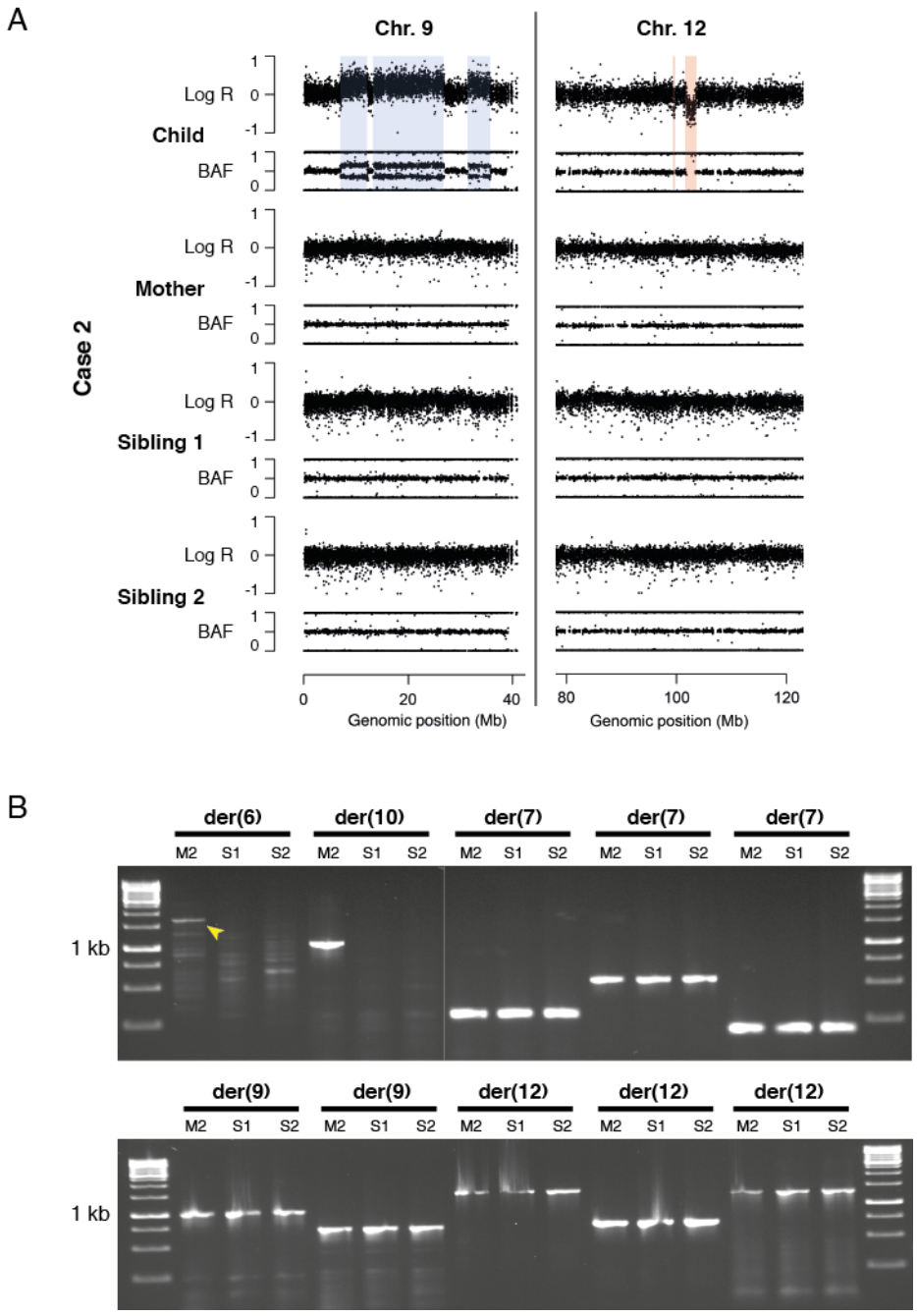

Figure S6. PCR and SNP array results for the two siblings of child 2. Partial inheritance of a different subset of chromothripsis chromosomes from the mother leads to a more copy number balanced state in two siblings of child 2. (A) SNP array results for the child, mother and two siblings in case 2. Child 2 carries three duplications on chromosome 9 and two deletions on chromosome 12 which are unique to this child. (B) PCR and Sanger sequencing revealed both siblings, like child 2, also inherited der(7) and der(12) and not der(6) and der(10) from their mother; however, unlike child 2, they additionally inherited der(9), leading to the copy neutral state of chromosome 9 and 12. Abbreviations are as follows: M2, mother 2; S1, sibling 1; S2, sibling 2.

Supplemental Tables

Table S1. Phenotypic and genomic characteristics of the children and mothers described in this study. Coordinates are in hg19. Abbreviations are as follows: C, child; M, mother; MP-seq, mate-pair sequencing.

C/M	Clinical phenotype	Birth/pregnancy history	Cytogenetic result	# of breakpoint junctions detected (MP-seq)	Chromosomes involved
C1	Ambiguous genitals; prominent forehead; hyperteloric; downslant; periorbital oedemia; beaked nose; thin lips; low-set ears; pointy chin; clinidactyly 5th digits; right foot: syndactyly (2nd-3rd digits), 2 rudimentary digits (4th and 5th). Li- g.b; asymmetrically enlarged ventricles, gallbladder agenesis; multiple VSDs and open ductus, hypotonic	1st pregnancy, caesarean section at 34.7 weeks of gestation Birth weight: 1900g (10th centile) Birth length: unknown Head circumference: 31 cm (10th centile)	46,XY,der(9)t(9;14;10;16)(p24.3;q32.1;q23.1;q21)mat.arr 9p24.3(193,993-615,714)x1,9p23(9,043,125-12,048,982)x1,10q23.1(84,551,375-85,687,314)x3,16q21q24.3(61,374,670-88,690,771)x3	5	9, 10, 14, 16
M1	No clinical phenotype	1.5 years until first pregnancy, no miscarriages/spontaneous abortions, PCOS; 6 ovolutations a year	46,XX,t(9;14;10;16)(p24.3;q32.1;q23.1;q21)	13	9, 10, 14, 16
C2	Severe mental and growth retardation, microcephaly, epicanthus, low-set ears, micrognatia, clinodactyly and hypoplastic phalanges of the fifth fingers, hypoplasia or absence of toenails, and extremely small genitals (For detailed description see: de Pater et al[10])	2nd pregnancy, uneventful and ended at 38.2 weeks of gestation. Birth weight: 2570g(<3rd centile) Birth length: 49 cm (25 centile) Head circumference: 32 cm (<3rd centile) [10]	46,XY,der(12)-ins(12;9)(q24.1;p22p24)mat.arr 9p23-p24.1(7,143,431-12,234,642)x3,9p21.2-p23(13,300,367-26,978,170)x3,9p31.3-p21.1(31,591,132-35,648,008)x3, 12q23.1(99,435,026-99,635,928)x1, 12q23.2(101,916,750-103,544,418)x1[10]	16	6, 7, 9, 10, 12
M2	Delayed psychomotor development and major learning difficulties, no dysmorphic features[10]	No history of previous miscarriages/ spontaneous abortions, 1st and 3rd pregnancy ended in the birth of male infants sharing only the mothers' phenotype[10]	46,XX,ins(12;9)(q24.1;p22p24) [10]	23	6, 7, 9, 10, 12
C3	Hypotonia, mild facial dysmorphisms, severe psychomotor development, non-progressie white matter abnormalities, small hands with single transverse creases, feeding difficulties, recurrent upper airway infections, and severe intellectual disability (For detailed description see: van Binsbergen et al[11])	3rd pregnancy, ended at 38 weeks of gestation Birth weight: 2,425 g (5th centile) Birth length: 50 cm (50th centile) Head circumference: 36 cm (80th centile) [11]	46,XY, der(1)(1pter->1q21.3::5p13.3->5pter),der(3)(3pter -> 3q22::3q29-> 3qter::3q29-> 3qter),der(5)(1qter -> 1q21.3::3q22 -> 3q29::5p13.3 -> 5qter).arr1q21.3(153751264-154439066) x 1,3q29(195420586-197837049) x 3 dn[11]	7	1, 3, 5
M3	No clinical phenotype[11]	One previous elective abortion and one early miscarriage[11]	46,XX, der(1)(1pter->1q21.3::5p13.3->5pter),der(3)(3pter -> 3q22::3q29 -> 3qter::1q21.3 -> 1q21.3:),der(5)(1qter -> 1q21.3::3q22 -> 3q29::5p13.3 -> 5qter). arr(1-22,X) x 2[11]	8	1, 3, 5

Table S2. Breakpoint junctions detected in all three cases (mothers and children) by mate-pair sequencing and PCR and Sanger sequencing. Additional information per column:

[a] For every breakpoint junction our analysis pipeline specifies the coordinates of the boundaries of the two genomic fragments that are connected together. Coordinates represent the outer boundaries derived from all mate-pair clones supporting the breakpoint junction. Chr1, s1 and e1 specify the chromosome and left and right coordinate of the first fragment respectively. Similarly chr2, s2 and e2 specify the chromosome and coordinates of the second fragment.

[b] Individuals the specific breakpoint junction was detected in: M, mother; F, father; C, child. Count represents the number of reads covering the breakpoint junction.

[c] Orientation of the breakpoint (BP) junction between the two chromosomal fragments specified in each row. A fragment can be detected at its head (H) or tail (T) side to another fragment. The first letter (H or T) indicates the orientation of fusion of the first fragment (chr1, s1, e1) and the second letter indicates the fusion orientation of the second fragment (chr2, s2, e2).

[d] Total number of unique mate-pair reads supporting the breakpoint junction. All reads were uniquely identified in the indicated mother and or child and not in any of 150 control mate-pair datasets.

[e] For every breakpoint junction, chr1 and bp1 indicate the exact breakpoint location of the first, and chr2 and bp2 that of the second of the two fragments that are connected together at the specific breakpoint junction. Abbreviations are as follows: nd, not detected by PCR; nt, not tested by PCR.

[f] Individuals the breakpoint junction was confirmed in by PCR and Sanger sequencing; M, mother; C, child.

[g] Characteristics of the breakpoint junction. Chromosomal fragments can be either fused blunt, the connection can be guided by microhomologous sequences at either end of the junction or a few (non-templated) nucleotides can be inserted between the two connected ends. Bp change represents microhomology or inserted sequence found at the breakpoint junction.

[h] PCR products for these breakpoint junctions were >1.5 kb, which exceeds the maximum read length of Sanger sequencing (~650 bp). Sequencing using the forward primer identified the first fragment, the reverse primer identified the second fragment of the breakpoint junction.

	Mate-Pair sequencing results									PCR and Sanger sequencing results						
Case	chr1[a]	s1[a]	e1[a]	chr2[a]	s2[a]	e2[a]	Individual (counts)[b]	BP junction orientation[c]	Total count[d]	chr1[e]	bp1[e]	chr2[e]	bp2[e]	Bp conf.[f]	Breakpoint signature[g]	Bp change[g]
case 1	9	10640282	10641388	9	12067163	12069048	M(4)F(0)C(0)	TT	4	9	10642756	9	12069307	M	Microhomology	A
case 1	10	85708440	85709723	16	62811907	62813013	M(4)F(0)C(0)	HT	4	10	85707849	16	62814401	M	Blunt	-
case 1	9	10643066	10645130	14	87959160	87960589	M(7)F(0)C(0)	HH	7	9	10642768	14	87958400	M	Insertion	T
case 1	9	9874447	9875887	9	10397725	10400216	M(8)F(0)C(0)	TT	8	9	9876859	9	10400356	M	Microhomology	TGTCAT
case 1	9	9877636	9880121	9	10184962	10187413	M(11)F(0)C(0)	HH	11	9	9876864	9	10184654	M	Microhomology	GA
case 1	16	62814654	62819507	16	62837225	62840110	M(2)F(0)C(15)	HH	17	16	62814403	16	62837185	C,M	Microhomology	T
case 1	9	9039041	9043602	9	12069501	12073773	M(5)F(0)C(16)	TH	21	9	9043619	9	12069309	C,M	Blunt	-
case 1	9	633474	636448	9	9043966	9046015	M(2)F(0)C(19)	HT	21	9	633072	9	9046145	C,M	Blunt	-
case 1	9	9043986	9045832	10	84550840	84554624	M(2)F(0)C(24)	HH	26	9	9043621	10	84550262	C,M	Microhomology	A
case 1	10	85704055	85707766	16	62832514	62837165	M(8)F(0)C(29)	TT	37	10	85707839	16	62837183	C,M	Insertion	T
case 1	14	87956251	87957402	9	630972	632755	M(3)F(0)C(0)	TT	3	14	87958400	9	633072	M	Blunt	-
case1	10	84549834	84549834	9	9048033	9048033	M(1)F(0)C(0)	TH	1	10	84550260	9	9046159	M	Blunt	-
case1	9	10182312	10184604	9	10400974	10404035	M(7)F(0)C(0)	TH	7	9	10184643	9	10400358	M	Blunt	-

case 2	7	36718108	36721671	12	99787831	99790893	M(11)F(0)C(10)	HH	21	7	36718028	12	99787840	C,M	Microhomology	G
case 2	9	11739802	11742710	9	31460247	31463007	M(4)F(0)C(3)	TT	7	9	31463189	9	11742848	C,M	Blunt	-
case 2	9	11758195	11759972	12	110263742	110266538	M(3)F(0)C(2)	TT	5	9	11760434	12	110266889	C,M	Blunt	-
case 2	9	7104763	7106960	12	103464346	103466774	M(11)F(0)C(0)	TT	11	9	7107289	12	103467187	M	Flanks identified[h]	-
case 2	9	12236188	12238989	9	26965673	26968013	M(10)F(0)C(0)	HH	10	9	12236168	9	26965304	M	Microhomology	TTG
case 2	9	13236665	13239425	12	101859464	101861141	M(7)F(0)C(0)	TH	7	9	13239726	12	101859126	M	Flanks identified[h]	-
case 2	6	123750357	123752276	9	11743072	11746336	M(9)F(0)C(3)	HH	12	6	123750144	9	11742852	C,M	Microhomology	A
case 2	6	123765872	123766384	12	101856510	101856537	M(2)F(0)C(0)	TT	2	nd	nd	nd	nd	nd	-	-
case 2	9	31433337	31436333	9	31463517	31465231	M(4)F(0)C(0)	TH	4	9	31436410	9	31463194	M	Microhomology	GA
case 2	9	31471300	31473229	12	99435218	99436825	M(6)F(0)C(0)	TH	6	9	31473570	12	99435026	M	Blunt	-
case 2	6	123748945	123750012	6	123768304	123769906	M(6)F(0)C(0)	TH	6	6	123749994	6	123768338	M	Flanks identified[h]	-
case 2	10	66227201	66230197	12	99785257	99787355	M(3)F(0)C(2)	TT	5	10	66230268	12	99787838	C,M	Complex	-
case 2	9	8118030	8119888	10	65848657	65850658	M(2)F(0)C(2)	TH	4	9	8120193	10	65848657	C,M	Only one flank identified	-
case 2	7	36714976	36717979	12	103467238	103470137	M(5)F(0)C(7)	TH	12	7	36718025	12	103467181	C,M	Blunt	-
case 2	9	7107641	7109780	9	35723601	35726283	M(8)F(0)C(6)	HT	14	9	7107388	9	35726564	C,M	Microhomology	A
case 2	9	13334672	13337076	9	31473792	31476201	M(6)F(0)C(5)	TH	11	9	13337420	9	31473565	C,M	Microhomology	TT
case 2	9	13240004	13241557	9	26962041	26964677	M(2)F(0)C(4)	HT	6	9	13239970	9	26965302	C,M	Blunt	-
case 2	9	7116561	7120299	9	31436447	31440029	M(7)F(0)C(5)	HH	12	9	7116530	9	31436439	C,M	Blunt	-
case 2	9	8121313	8122780	9	11760865	11763707	M(5)F(0)C(4)	HH	9	9	8120747	9	11760437	C,M	Microhomology	TGAG
case 2	9	12233990	12236078	12	99431406	99434124	M(11)F(0)C(6)	TT	17	9	12236167	12	99435033	C,M	Microhomology	G
case 2	9	7113602	7116395	12	110267135	110269429	M(10)F(0)C(6)	TH	16	9	7116529	12	110266889	C,M	Blunt	-
case 2	9	13337440	13340068	12	99636114	99638904	M(4)F(0)C6)	HH	10	9	13337419	12	99635928	C,M	Only one flank identified	-
case 2	10	65844061	65847557	10	66230913	66234003	M(7)F(0)C(0)	TH	7	nt	nt	nt	nt	nt	-	-
case 2	12	99634461	99634461	9	35727377	35727377	M(1)F(0)C(0)	TH	1	nt	nt	nt	nt	nt	-	-
case 3	3	158384939	158388876	3	195632595	195635632	M(8)F(0)C(6)	TH	14	3	158388978	3	195632590	C,M	Microhomology	GT
case 3	1	154441965	154444308	3	195630328	195631725	M(6)F(0)C(0)	TT	6	1	154444963	3	195632204	M	Flanks identified[h]	-
case 3	1	154445895	154448216	3	156530510	156533437	M(3)F(0)C(6)	HT	9	1	154445666	3	156534036	C,M	Microhomology	CG
case 3	1	153155549	153158347	5	29436496	29438223	M(10)F(0)C(3)	HH	13	1	153154773	5	29436386	C,M	Microhomology	AT
case 3	1	151021102	151023743	5	29433067	29436167	M(5)F(0)C(4)	HT	9	1	151021008	5	29436271	C,M	Microhomology	GC
case 3	3	128911732	128913422	3	156534632	156536932	M(8)F(1)C(2)	TH	11	nt	nt	nt	nt	nt	-	-
case 3	1	153746565	153748730	3	158389428	158392531	M(7)F(0)C(1)	TH	8	1	153749340	3	158389283	C,M	Microhomology	ACT
case 3	1	151018679	151020295	1	153151785	153154013	M(6)F(0)C(4)	TT	10	1	151020980	1	153154640	C,M	Insertion	TGGGAAGCTGCAC

Table S3. Comparison of the rearrangement characteristics for each of the three mothers to the criteria for inference of chromothripsis as defined by Korbel and Campbell[14]

Criteria for inference of chromothripsis	Mother 1	Mother 2	Mother 3
1. Clustering of breakpoints	+	+	+
2. Regularity of oscillating copy-number states	-	-	-
3. Prevalence of regions with interspersed loss and retention of heterozygosity	-	-	-
4. Prevalence of rearrangements affecting a single haplotype	+	+	+
5. Randomness of DNA segment order and fragment joins	+	+	+
6. Ability to 'walk' the derivative chromosome if all rearrangements in a region with chromothripsis are detectable	+	+	+[a]

[a] Due to the repetitive nature of one of the regions involved in the rearrangements we were unable to detect all breakpoints in this individual and could therefore not reconstruct all derivative chromosomes

Table S4. Effects of chromothripsis breakpoints on protein-coding genes.

Mother (case)	Gene symbol	Gene name	Breakpoint location[a]	Phenotype MIM number	Probability of haploinsufficiency[b]
1	*PTPRD*	PROTEIN-TYROSINE PHOSPHATASE, RECEPTOR-TYPE, DELTA	2 Breaks in intron; 2 breaks in 5' region of gene		0.855
1	*KANK1*	KN MOTIF- AND ANKYRIN REPEAT DOMAIN-CONTAINING PROTEIN 1	Break in intron	612900	0.653
1	*NRG3*	NEUREGULIN 3	Break in intron		0.251
2	*AOAH*	ACYLOXYACYL HYDROLASE	Break in intron		0.160
2	*ANKS1B*	ANKYRIN REPEAT AND STERILE ALPHA MOTIF DOMAINS-CONTAINING PROTEIN	3 Breaks in different introns		0.313
2	*TRPV4*	TRANSIENT RECEPTOR POTENTIAL CATION CHANNEL, SUBFAMILY V, MEMBER 4	Break in 5' region of gene	184095;181405;168400;156530;113500;184252;600175;613508;606835;606071	0.228
2	*KDM4C*	LYSINE-SPECIFIC DEMETHYLASE 4C	2 Breaks in intron		Not defined
2	*MPDZ*	MULTIPLE PDZ DOMAIN PROTEIN	Break in intron	615219	0.450
2	*IFT74*	INTRAFLAGELLAR TRANSPORT 74, CHLAMYDOMONAS, HOMOLOG OF	Break in intron		0.357
2	*TRDN*	TRIADIN	2 Breaks in intron	615441	0.288
2	*TLN1*	TALIN 1	Break in intron		0.417
2	*CREB3*	cAMP RESPONSE ELEMENT-BINDING PROTEIN 3	Break in 5' region of gene		0.298
3	*GFM1*		Break in intron		0.188
3	*LXN*	LATEXIN	Break in intron		0.303
3	*TNK2*	TYROSINE KINASE, NONRECEPTOR, 2	Break in intron		0.514
3	*SHE*	SH2 DOMAIN-CONTAINING PROTEIN E	2 Breaks in 3' regions of gene		0.143
3	*IL6R*	INTERLEUKIN 6 RECEPTOR	Break in 3' region of gene	614689;614752	0.209
3	*C1orf56*		Break in exon		0.143
3	*BNIPL*	BCL2/ADENOVIRUS E1B 19-KD PROTEIN-INTERACTING PROTEIN 2-LIKE	Break in 3' region of gene		Not defined
3	*CDC42SE1*		Break in 3' region of gene		0.316
3	*MLLT11*	MYELOID/LYMPHOID OR MIXED-LINEAGE LEUKEMIA, TRANSLOCATED TO, 11	Break in 5' region of gene		0.220
3	*LEKR1*	LEUCINE-, GLUTAMATE-, AND LYSINE-RICH PROTEIN 1	Break in 5' region of gene		0.232
3	*PA2G4P4*		Break in 5' region of gene		Not defined
3	*SLC27A3*	SOLUTE CARRIER FAMILY 27 (FATTY ACID TRANSPORTER), MEMBER 3	Break in intron		0.126
3	*INTS3*	INTEGRATOR COMPLEX SUBUNIT 3	Break in 3' region of gene		0.564

[a] We arbitrarily defined a 20kb region on the 5' and 3' part of the gene as the 5' and 3' region of the gene, respectively.

[b] Based on the method described by Huang et al.[17]