9 Supplemental material

9.1 Experimental protocols

9.1.1 Array CGH

For array CGH a 36k BAC sub-megabase resolution array was used, comprising the 1Mb Sanger set (clones kindly provided by Nigel Carter, Wellcome Trust Sanger Centre)(Fiegler et al., 2003) a set of 390 subtelomeric clones (assembled by members of the COST B10 initiative: Molecular Cytogenetics of solid tumours) and the human 32k Re-Array set, http://bacpac.chori.org/pHumanMinSet.htm; DNA kindly provided by Pieter de Jong) (Ishkanian et al., 2004; Krzywinski et al., 2004; Osoegawa et al., 2001).

The general overview of the process of array CGH experiment is shown in Figure 25. The following sections describe each step in detail.

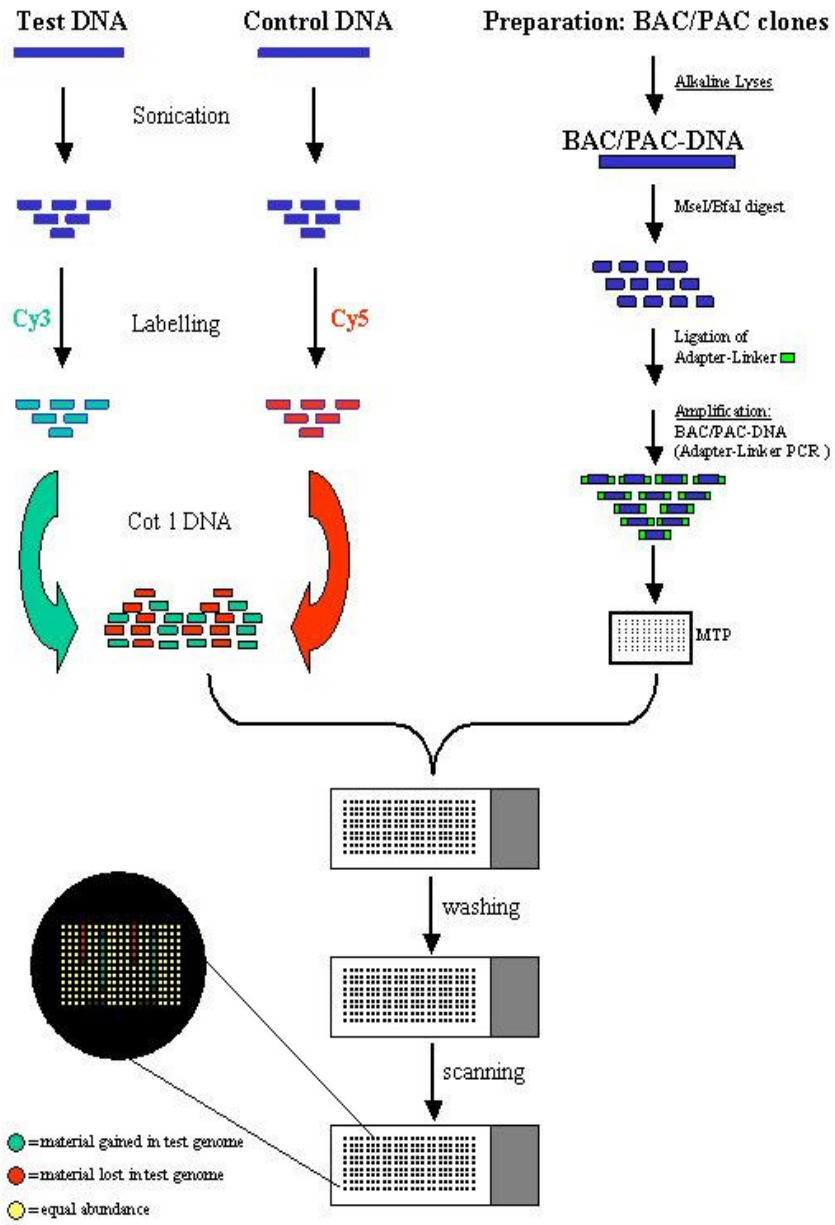

Figure 25: Procedure of array CGH (Courtesy of Dr. Erdogan). Up right part: the DNA from BAC clone inserts is isolated and amplified. The amplified products are spotted on the glass slides. Up left part: the DNA from test sample and reference sample are differentially labelled. The 2 differentially labelled DNA samples and excess unlabeled Cot-1 DNA, which can suppress the repetitive sequence, are then hybridised on the glass slides. Lower part: After hybridisation, the slides are washed and scanned.

9.1.1.1 Array production

9.1.1.1.1 BAC Insert isolation

BAC inserts were isolated in a 96 well format. First, BACs were treated by alkaline lyses according to the standard protocol. Then the DNA was treated with the exonuclease treatments to remove the remaining Ecoli DNA.

9.1.1.1.2 Amplification of BAC DNA by Linker adapter PCR

The isolated BAC DNA was amplified by linker adapter PCR. Linker adapter PCR consists of three steps: the target DNA is first digested with an appropriate restriction enzyme. And then each end is ligated to an adaptor. Finally, the known adaptor sequences are used to uniformly amplify each of the many DNA fragments representing the original samples. The following sections describe the three steps in detail.

9.1.1.1.2.1 Restriction Enzyme Digest of BAC DNA

Restriction enzyme digestion was carried out in a 7.5 µl reaction volume containing: 0.75µl of 10 x NEB1-buffer (New England Biolabs), 0,075µl of 100xBSA, 0,012µl of MseI (50U/µl), 0,15µl of BfaI (5U/µl) and 1,513 ml of H_2O, 5µl of exonuclease digested DNA. The reaction was placed in a PCR machine for 3h at 37°C. After incubation, the reaction was inactivated at 80°C for 20min. The digests were then run on a conventional 1% agarose gel to check fragment length. Restriction sizes should range from 100 bp to 1500 bp.

9.1.1.1.2.2 Ligation of Specific Primers to BAC DNA

The ligation reaction was first set up in a 8 µl reaction volume containing: 0,5µl 100µM primer-21 (5`-AGTGGGATTCCGCATGCTAGT-3´) and 0,5µl 100µM primer-12 (5`-TAACATGCATGC-3`), 0,8 10x ligase buffer (Roche), and 5,2µl of H_20 and 1µl digested BAC DNA (see above). In a thermocycler with heated lid an, the reaction was carried out at 65°C for the first 2 min to make the two oligos single stranded, and then the temperature was shifted down to 15°C, with a ramp of 1.0°C/min, to allow annealing of the two oligos. At 15°C, 0,2µl ligase buffer, 0,2µl T4-DNA-Ligase (5 U/µl; Roche), and 1,6µl of H_2O were added and the

reaction was placed in the thermocycler for an overnight incubation at 15°C (18-20h).

9.1.1.1.2.3 Ligation Mediated PCR

The ligation mediated PCR was carried out in a reaction volume of 50 µl containing: 1µl ligation product, 5µl 10x PE buffer, 10µl dNTPs (1mM each), 0,5µl primer-21 (100µM), 32,5µ H_2O. Overlaying the sample with 30µl mineral oil to avoid evaporation at high temperature, the PCR program started at 68°C for 4 min to remove the MseI/BfaI-Lig12-Primer and 1 µl (10 units) of DNA polymerase was added and a 4 min. incubation for the fill-in reaction. After 3 min. at 95°C denaturation step, the PCR cycled at 95°C for 40 sec, 59°C for 30 sec, 90 sec (+2 sec/cycle) for 35 cycles. A 7 min extension at 72°C completed the protocol. Some of PCR products were run on a conventional 1% agarose gel to check fragment length. Size of the PCR product should range from 70 to 1500 bp, with the highest concentration of product around 200 to 800 bp.

9.1.1.1.2.4 Re-PCR of Ligation Mediated PCR

The ligation mediated PCR is used as a template in a Re-PCR reaction to generate DNA for spotting. 1µl of the primary PCR product was amplified under the following condition. After 3 min. at 95°C denaturation step, the PCR cycled at 95°C for 40 sec, 59°C for 30 sec, 90 sec (+2 sec/cycle) for 35 cycles. A 7 min extension at 72°C completed the protocol. Again some of the PCR products were run on a conventional 1% agarose gel to check fragment length. Size of the PCR product should range from 200 to 1500bp.

9.1.1.1.3 Preparation of Spotting Solutions from Re-PCR used for array CGH

The Re-PCR products were precipitated by adding 150 µl pre-chilled 100% ethanol and sodium acetate (pH 5.2). Then the dried DNA pellet was dissolved in 3xSSC/1,5 M Betaine.

9.1.1.1.4 Production of array

The products were robotically spotted onto epoxy coated glass slides (Nunc, Wiesbaden, Germany) using an in-house modified Qarray (originally from

Genetix, new Milton, U.K.) and Pointech (Gibbon, MN) Tungston PTL 2500 slit pins. Here, the microspotting technique was applied, where a spotting roboter disposed the PCR products directly on the slides. Epoxy slides were chosen due to several reasons. First, epoxy slides are especially suitable for covalent immobilization of oligonucleotides (10 to 80 bases), PCR products as well as cDNA molecules. Second, additional amino-modifications of the nucleic acids are not required. Third, their hydrophobic surface allows small spot diameters (100 to 130 µm, depending on the type of pins and spotting buffer) to create high-density arrays. Finally their surface chemistry is very stable and remains active even during very long spotting runs.

9.1.1.2 DNA labelling and hybridisation

9.1.1.2.1 Random Primed Labelling of genomic DNA for array CGH analysis

Genomic DNA samples were sonicated to generate fragments 200–2,000 bp in size. Test and reference DNA was labelled by random prime labelling (BioPrime DNA Labeling System, Invitrogen, Carlsbad, California) with fluorolink Cy3-dUTP and Cy5-dUTP (Amersham Biosciences, Piscataway, NJ). Briefly 1µg of DNA were mixed with 2.5 x random primer solution, incubated at 95°C for 10 min. and then immediately cooled on ice for 5 min. Consequently, 5µl dNTP mix (2mM dATP, 2mM dCTP, 2mM dGTP, and 1mM dTTP, in TE buffer), 3µl 1mM Cy3-dUTP or Cy5-dUTP and 1µl Klenow fragment was added to the reaction mix and incubated overnight at 37°C. The labelling reaction was stopped by adding 5µl stop solution. Probes were then purified by Qiaquick purification kit to remove the unincorporated nucleotides according to manufacture instructions. Two reactions as described above were pooled for each channel.

9.1.1.2.2 Slide Processing

Slides was prehybridised at 42°C for 1h in the blocking solution (200µl heringsperm DNA, 0,1% SDS, 4xSSC, 0,5% BSA). Afterwards, the Slides were immediately rinsed 5 times with Millipore water and air-dried by centrifugation for 5 min. at 150g.

9.1.1.2.3 Hybridisation of labeled genomic DNA

The labelled test genomic DNA and the labelled reference genomic DNA from two purified random priming reactions were pooled with 500 µg of human Cot-1 DNA (Invitogen, Roche). Pooled DNA was then precipitated by adding 2.5 volumes of ice-cold 100% ethanol and 0.1 volume of 3 M sodium acetate (pH 5.2). The precipitated DNA was dissolved in 6,8µl 10%SDS, 3,4µl yeast tRNA (100µg/µl, Invitrogen), and 24µl master hybridisation mix (70% formamide, 2,8 x SSC, 8% dextran sulphate), and denatured at 70°C for 15 min. After denaturation, the hybridisation mix was incubated at 42°C for 2h to allow the Cot1 DNA to anneal to repetitive sequences on both the sample and reference DNA. The labelled probes were then placed on the slide under a coverslip. The arrays were incubated for 24 hours under humidified conditions using a slide booster from Implen (Munich, Germany). After hybridisation the slides were washed with 50% formamide 2xSSC, 0,1% SDS for 15min. at 42°C, followed by a 10 minute wash in PN buffer (0,2 M sodium phosphate with 0,001% NP40) at room temperature. The slides were then incubated for a 30 seconds in 1xPBS and 2-3sec in millipore water. Finally the slides were dried by centrifugation at 150g for 5min.

9.1.1.3 Scanning

Following hybridisation, slides were scanned at 532 nm (Cy3) and 635 nm (Cy5) using a GenePix 4000B laser scanner (Axon Instruments, Union City, CA) in order to read out the fluorescence signal intensities in each channel. The resulting 16 bit TIFF images were analysed employing Genepix Pro 5.0 software (Axon Instruments).

9.1.2 Mapping balanced translocation breakpoint by chromosome sorting and array painting

9.1.2.1 Cell culture and preparation

A t(1;13)(p31.2;q22.1)-containing cell line was established by Epstein-Barr virus transformation of peripheral blood lymphocytes and cultured in RPMI 1640 medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, and antibiotics at 37°C in a humidified atmosphere containing 5% CO2. Cells in log

phase were treated for 16h with colcemid (0.05 mg/mL final concentration) to arrest cells in metaphase.

9.1.2.2 Flow karyotyping and sorting

Chromosomes were stained with chromomycin-A3 (CA3) and Hoechst 33258 (Ho) and analysed on a dual-laser beam flow cytometer (FACSVantage™ SE; Becton Dickinson, Franklin Lakes, NJ, USA). CA3 was excited with an argon ion laser tuned at 458 nm at 100 mW, and CA3 fluorescence was measured through a 550 nm longpass filter. Ho was excited with an argon ion laser tuned into the UV range (351 and 364 nm) at 125 mW laser power, and Ho fluorescence was measured through two KV 408 filters. The system was triggered on the CA3 fluorescence signal. In total, 6000 to 12000 chromosomes were sorted from each cluster. DNA obtained in this way was amplified by GenomePlex.

9.1.2.3 DNA amplification using GenomePlex

For the amplification of flow-sorted chromosomes we used the GenomePlex Whole Genome Amplification (WGA) Kit (Rubicon Genomics). GenomePlex WGA is based upon random chemical fragmentation and conversion of genomic DNA into a library of DNA molecules flanked by universal priming sites. DNA Fragments are amplified by standard PCR using universal oligonucleotide primers.

9.1.2.4 Array painting using 36K BAC array

Labeling and hybridization of DNA amplified by GenomePlex on 36K BAC array were performed as described in the Section 8.1.1.2.

9.1.2.5 Fluorescence in situ hybridisation

To confirm the breakpoint regions determined by array-CGH, we employed fluorescence in situ hybridization (FISH) experiments. A permanent lymphoblastoid cell line of patient 2 was established by EBV transformation according to standard protocols after informed consent. FISH was performed using three BAC clones at each breakpoint region. For the breakpoint on

chromosome 1, BAC clones RP11-55O04, RP11-764P11 and RP11-746B05 were employed. BAC clones RP11-490G20, RP11-339I10 and RP11-702F16 were used for the breakpoint on chromosome 13. DNA samples were prepared according to standard protocols and were labeled by nick translation with either biotin-16-dUTP or digoxigenin-11-dUTP. Immunocytochemical detection of probes was performed as described elsewhere (Wirth et al., 1999). Chromosomes were counterstained with 4'-6-diamino-2-phenyl-indole (DAPI). Metaphases were analysed with a Zeiss epifluorescence microscope.

9.1.2.6 PCR fragment subarray

The genomic sequence of specific breakpoint spanning BAC clones for chromosome 1 (RP11-764P11) and chromosome 13 (RP11-339I10) were chosen as targets for the design of a PCR amplicon subarray.

A customized Perl script incorporated within CGHPRO was used to design the primers for the amplicons. The script first divided the BAC clones into evenly distributed intervals of 2 kb. Then, by using Primer3 (Rozen and Skaletsky, 2000), within each interval, it designed primers for generating PCR fragments ranging from 500 to 800 bp in size. To facilitate the subsequent amplifications using the same condition, primers were selected to have the same annealing temperature. Finally, to confirm that the amplicons are specific for the target region, the script searched the whole human genome for the presence of the amplicon sequences by using BLAST with the default parameter. Amplicons with more than one matches in the genome were excluded from PCR amplification and spotting. The primers designed for RP11-764P11 and RP11-339I10 are listed in the Table 6 and Table 7, respectively.

Table 6: List of primer pairs designed for RP11-764P11 (chromosome 1)

NO.	FORWARD PRIMER	REVERSE PRIMER
1	TCCAGCTTCATTCATAGGGC	CTCAAAGCGCTCTTACCCAC
2	ATTCTGGCTAGGTGTGGTGG	TTAAGCACCTGTGACCTCCC
3	TTGAGGAACTGGGGACATTC	CAGTCTCTGCTTTTGAGGGG
4	GCTGAGCAGAGAGGGATTTG	TGGCCTTAAAACTGGACCAC
5	CTCAGCTACAGGAACCCCAG	TGGGTAAAATGTCCCTCCTG
6	TAACTGGATCTTCCGCATCC	CCCACCTGACCAATATGGAG
7	CAGCACACTGATGGGTCTTG	ACGTGAAGAATGCAGAAGCC
8	GGAGCTGGTTTTTCAAAAGG	CACGTGCCTGTAAGCCTAGC
9	CTTCAGGCAAAAGGTTGAGC	TGAACCCAGGAGAAGTCCAG
10	GCCACAGAGTCTAACGAGGC	TTCTCTGCATTCCTCACACG
11	TGGCCTCGTGTCTGTAGTTG	CTGGATTCAGGCCCTAAGTG
12	TCTGTGGTGTTTGCTGCTTC	TACCTCTGATGATGGGGAGC
13	GAGCCAGGCGTTCTGTTTAG	CTGCAATTGACCCACAAATG
14	TAGCAGGTCACCCAGAGTCC	CCCTCGGAGCCTCTATTTTC
15	ATCACCAGTGAAAAGGACGG	GAAGAGTCTGGCCTCCAGTG
16	ATCTGGGACAACAGAGCTGG	CTGACAGAAGGCTCCAGACC
17	CTTTGGAAGACTGAGGCAGG	TAGTTTGGCTGTGTTTCCCC
18	GGAAGAGCTTTTCATGCCAG	TGCATAAGCTTTTGTGCCAG
19	CCAGTCCTCTCTTGCCTGAC	ACTCATGGCCTATGACCCAG
20	TCAGAAGAATGGCCCCATAG	TTATGTCCAGCCCCCAGTAG
21	CTCTAGCCTCATCACCCAGC	TTTTGCACTCTGTCACCCAG
22	GATGCCTGCTTTCTTCCAAC	TCACCTCACAGCGAAGTCAC
23	AAATCACATCAAGGAACCGC	TGCCAAGTGTAGTGTCTGCC
24	TGTGTCAGAGCCACAGAAGG	AACATCGTGCGTTTACCTCC
25	ATTCCCTTGGCTGTCAAATG	CTCAGCCCTTGGAGAAACAG
26	CACCTCCATGATCCCAATTC	ACTCATGGGAACAGGAAACG
27	AATTCCAGCACTCCGTGTTC	ACAGTGGACAGGTTTGAGGC
28	ATGGGCATGAAGATAGGCAC	TTCTACAGAGGGCACATCCC
29	TACCAGATGTGCAGAGCCAC	GGGCACAGTGGTATTATGGG
30	CTCAGCACACAGTAGGCCAG	GAGGCAGCCATCATTCTCTC
31	TAAATTCCCTGCCATTCTGG	GGTTGCTTGCTTGTAATCCC
32	ACAGCCATCTTTCAACCCTG	CACTAACAGGCCCTCTCTGC
33	CATACCTGGGTTGCTTCCAC	CTCATGGGCTTTAGCAGCTC
34	TTCTGCCCCTGTTTTCATTC	TTTTGGCTCTTTTTGGTTCC
35	TTGGAGACTCAGAAATGGGG	TTCCACATTTTCTTCCAGCC
36	CAGTGTCCCCAAAGAGGAAG	TTTAATCAGGGCTGGAGTGG
37	GTGAACTGGGACTAGCCAGC	ATGAGGATAAGGACCCCCAC
38	GAGCTCTGACTTCTGGGTGG	TCACCTCTTTCCTGGGATTG
39	GCCTGAAATGCTCTCTACCG	AATGCTCATCCAGCCAAATC
40	GATCTTCAGGGAAGGAAGGG	CTCCTCTGGATAAGGGGCTC
41	CTGCCTTGAGTGAAAGGAGG	GTTTCCTCTTCATGCCTTGC
42	GAAGGGTTGCTTCAGACTGG	ATGCAAGAAAGCACATGCAG
43	AGGCAGGTGGATCATTTGAG	AAAATGGCATTACTGGGCTG
44	TGAGATTTGGGGTGTGATTG	ATCCACCCTCTTTGGTCTCC
45	ATTAGGAGTGCAGTGGCACC	GCTAAGGCTGATGAAGTGGC

46	CACAATCTGCATGCTGTTCC	TTTCAAGCATAGGTCCTGGG
47	CTGCTCAGTCCTTCAGGGTC	GGCAAGATGAAGAGTCCTGC
48	TGACTTCCCCTCATGACTCC	CCCTTCCATCTTCTTCCCTC
49	TGACAGGGAAGGAAATGAGG	GCAGGAAATCTCTGAGGCTG
50	TGTGCAGTCCACTCAAAAGC	AAGATCCACACATTCCTGCC
51	ACGTCCACTCACCCTTTGTC	TCGTCTGGGTCTCAAATTCC
52	TTGGTGAGGAAGACCGAATC	GAGAGCGGAAATGGAAGTTG
53	TGCAGTTCAGGCAAATGAAG	TAAAAGAAAATTGGTGCGGG
54	ACATGGCAAGTCTCCCTCAC	TGGTCCTCATGTTCAAGCTG
55	ATGTGACCAAAGGATCTCCG	AGCCCAGTACCTGGATGTTG
56	CCAGGCAGTTCCAAGAAGAG	AGTCAATGGGGTGACTTTGC
57	TTGTCAAGGAGGGGAGAATG	AACACAAGAGTGGGCAAACC
58	GAAGGCAGGGAGATCACTTG	TCCCCAAACAGAGGACAATC
59	GAGCCTGTCTCTCAAGCACC	GAGCAAGTGAAGGGATCAGC
60	TCAAGGTTCTTACGGGCATC	AATGTTGATGAGGAGCACCC
61	CTTTAAATGTGCTCCTCGGC	TGGAAGGAGAGAACCACCAG
62	CTTTGTGGTTTTGGGTCCAC	TGATGTTTGAGGCCTTTTCC
63	TTTGATCTTAGACAGCCGCC	TGATCCAGCACATGCTTCTC
64	AGCATATCCTGGGCATGAAG	TGTGCTGGGTGCTCTGTTAG
65	CCCACAAATGTGACTGCAAG	CCCAAAGTGAGGTTGTTTCC
66	ACCATGGGAGCATGTTAAGC	GGGTCTGTAATGGCTTCCTG
67	AAGTGGTATTTGCTGGGTGC	TGGGACAAAGTCCAGGCTAC
68	AATTGGTTGGTTGGCTTCTG	TAGTTGCTTTTCCCCACCAC
69	ATGCTCCGTGCTAAGCTCTC	TGTCCATCCTTCCTTCCTTG
70	TGCCACATACTGAAAGCACC	CTGACTGGACTTCCTGGTCC
71	AAAATAAGGTCTGGTGGGGG	TGCATATTGGTCACAATGGG
72	CCTAAAATAGGGAGGACGGC	TTGCGGAAGCAACATACAAG
73	CAAAAGCAGAGGGAGACCAG	TCTCTGTCGGTGCGTATCAG
74	TACTTCCAGTCATGAGGCCC	AGCCTGAGTCATCATTTGGC
75	CATCCCTGAATACAAACGGG	CACCTTTCGGTCTTCTGCTC
76	GCACTCTCGGGACTTCTCAC	AGATCCACAATCAGGCAACC
77	ACAGAAACTGATGCCAAGGG	GCAGCTGCTAGCAATATCCC
78	TGGAACGAGCACAGTAGCAG	GTGGAAAGGAGGAAGCAATG
79	TGAAGGGCATTTACTCCAGG	CTAAGCCCCAGTCTGAGTCG
80	CTAGCTGGGAGGCTTGTTTG	CCCAATGAGCTCCTAAATGC
81	TGTACAATGACTTGGGGTGG	ACCCAGGCAGATTGTACAGC
82	CAATGGGAGAACACATGCAC	GGCAAGCCTAAAGCAGTCAC
83	AACTCACCCCTACCCAATCC	CCCTTTCCCTTCCTTCACTC
84	AGTCATGAAATCATGCCGTG	GCCCTTAGCCTTGTGTCATC
85	ACTCAGCATCATTGAAGCCC	AGGAACATGCCAGGAGTGAC
86	CTAGCACAGCCCAGATAGCC	TGCTTCCAGGATTTTTCCAG
87	AGGTGGACAGTAGGTGTGCC	GTAAGTGCTGGTCTGGGAGC
88	TAGGCAACCTCACCCTGAAC	TGAGCTGGCAAGTGTGAATC

Table 7: List of primer pairs designed for RP11-339I10 (chromosome 13)

NO.	FORWARD PRIMER	REVERSE PRIMER
1	GAATGTATGGGCTGTTTGCC	TTTTTAATGCCTTTGCCCTG
2	GGAAGCAGCAGGAGACAAAC	GCTCCCTCTCATCTCCACAC
3	ACCAGATTTCTCCCATGTGC	CTTGTGTGGGTGGTTCAGTG
4	GAGAAACGGTCATTTTTGCC	CTTTTCCTTCAAGGTGCTGC
5	GAGTCCACTCTTGCCTCCAG	TTTTGGGACTTGCATACCTTG
6	GCCGTCTGCTAGCTTTTG	AAGACCCATCAGTGTGCTG
7	TCATCGCGTAGTTCTTGTGC	CACCATTGCAGAGGATTGAG
8	TGTCCGTGTTTTTCAAATGG	CCGTGTTAACCAGGATGGTC
9	ATGGATCATGAAGGACCCTG	ACGTCTCACCCTAATGTGGC
10	TGCTGAATGCAGTTTTCCTG	CCTAACCCCCTTGTCTCTCC
11	AGCAGAGTTGTATGGGTGCC	TTGGCTTAGGTTTACTGGGG
12	TAATTCCCTTGTGGAGCTGG	AAATTGGGATCCGAGAAACC
13	GGGATTGGCTATTCCTCCTC	TGTGGTCCATTGTTTGTTGG
14	TCCAAAATGAGAAAATGCCC	CGCTGAGCCTTGGTTTCTAC
15	GGTGCCTCTCAAGGTACAGG	TCTCTGCCCAAGCTGACTTAG
16	TCAAAGCACAATTGAGGGTG	TCAACTTGTCCCTCAGAGCC
17	AATTCAACCTTTTGCCTCCC	CCCTAATGAATGGGATGTGG
18	CCTTCCTTGAGGGAGGAAAC	AAATCCAGTGATGGTAGGCG
19	ACAAGGCAAAGGGTCACTTG	GTGAGAGAGCCTGACATCCC
20	CTTGAGGTCAAGCCAGAAGG	TTCCCAAAGGCTTATGATGG
21	GGGCTAACCAATCAAAATGC	AATGCCCTTCTATGTGTGTGG
22	AACAACCGTGGATTCTCAGG	CTTGCAAACTCTCCTTTGCC
23	ATCACCTTTTCTGGCCACTG	GGGAACCAGAAGATGCAAAG
24	CTCCATTGTTCCATGGCTTC	TGGAAGGAAATTCCAAGCAC
25	AATTTTGGATCTTCCCCCAG	AGTTCACCTTGGACCCACTG
26	ATGGCCACAGTATGTCCTCC	GAGTTGTTTCTGGCCTCACC
27	TGAGTGAGCACGGACAAGAC	GGCTGGTGTCCTGAGACTTC
28	AGGCACTATGCTATGGCCTG	AGCCACCAGCTTTGTCTCTC
29	ACCTGCAAAGCAATTCCAAC	CTTTCACTCAGGCCAGGAAC
30	CACAATCATAGGCATGTCGC	CAGAGACAGAGCAATGCGAG
31	TGGCAATTGTCTCCACTTTG	GCAAAGCAGGAGTAAGGCAC
32	GCATAGGTAGGTGCCTCTCG	ACTGGGCAAGTCATCGAATC
33	AATTTTAAAAGTTGCCCCCTC	TGCCAGATGGAGAGATTGC
34	CCCCAGAAGTCCTCTGTTTC	AAATGTTGATATCCTGGCCG
35	CATGGGATTTTGTACAGGGC	AGTTACCAACCCCTGATCCC
36	CTCGGTTGGAACAAAAGAGG	AAACCATTGCCAAATCCAAG
37	CGCAAAAGCAGGAAGTATGG	GATATCGCTCCCATTTCTGG
38	ATGCCTTGAAAAAGAGGCAC	AGCAGGAAATGGTGATGAGG
39	TGTGTGAGTGCTTGGTAGCC	ACAAGGGTGTTCTGACCTGC
40	ACACAGGAAAAGCCCTTGTG	TCTTTCCCCATTCCTAACCC
41	AAAGACCGCATATGCCAAAC	CAAAGATGCTGTCCTTTCCC
42	AGGACTGCTCTGCCTACGAG	TAAGGTGGGAGGATTGCTTG
43	TGAGATAAAGCCAGGGATGG	GAAGGAACAAAGCAAGCAGG
44	AAATTCCTGGAGCATTGTGC	CCAGTGTTCCTCCTCTCTGC
45	AGGGGATGCCCAAAAATATC	GACTTTTGGCTGTTTCTCCG

46	TTGGATGGGGGAGTGTTATC	TACCCTCTAATGCGGCAATC
47	CTCAGAGGCAGATAGCCCAC	AGATGGGCAATGAGATCCAG
48	ATTTCCCTCCTTTGTGGGAC	TATTGCAGGGAGATTTTGGC
49	TGCTGAGGATTCATCCCTTC	GCTGTTTGCAACATTCATCC
50	TGGCAGATTTATGAGCTTTGC	TTCTTGGCTCCACCCTTATG
51	TTCCTCAGGCCTTCTTTCAG	CACCCCGTGTCCAAGTATTC
52	AGATGTTTGGAAGGCAATGG	TGAGATGAGGACTGGCTGTG
53	TGGCAATTTCCTGTGAATCC	GCCAGTGAGGAAGAGTCAGC
54	AATTTCTCCATCATCTGGGG	TCTCGTTTGGGAATCTCTCG
55	CGAATTTTCACCAGTCCCTG	TGGGAATGAGAAAGACAGGC
56	GACTTTGCCAACTTTGGAGC	CTTTGCTCTGTCTTTTCGCC
57	TTGAAGAGACTGCCCTTTCC	AAAAATGCCTTATGAAAGCCC
58	AACGCAGTTCAAGTCCAGTG	CTTCTTTCTGTGAGCTGCCC
59	ATATGGGCAGGATAAAGGGC	TCACACTTGGAAATGCCTTG
60	CTTGAGCCCAGAGACCACTC	AGGGCCCATTTTCATTTTTC
61	AAAGTGGGTGGAACACAAGG	CACATGACCACAAGGTGAGG
62	GGTTGAGCCAGCTCTCTTTG	AATTTCTGTGGTGTGCTCCC
63	GCAAGGGGTATAGCAGATGG	AACAACCACAAGAAGGCTGG
64	ACTGCTTCAAGCTCAGGCTC	GAATGATACGGTGGGGAATG
65	TAGACCCCAGATTGCTGTCC	TGCAGTTTAAACCTGGAGGC
66	ATCTTGGACTGTGTGGGCTC	TGCGTTTCTCCTGTGTATGC
67	ATGAAAAGTCCCTGTGGCTG	GCATTGTGACCAAGCATGAG
68	TTTGAGGAGTTGGGATCCAG	TGCAATGAAAGCCACAGAAG
69	TACCCCTGAAAATGACACGG	GGCAGATTGCTCTCGAACTC
70	GAGCTGGTGGACGAGTTAGG	ACCAAAACCAGCAAAGGTTG
71	GTCATGCCTTTGGAGGTTTC	TCCTGAGGAGCTGGGAGTAG
72	ATCCTCATTCATCCACTGCC	GGGTTGCAATGATTTTGAGC
73	AGGAAATGAAATCCCCCATC	GCTTTTCCTTGAATTGCAGC
74	TCTTACATTCACCCGCTTCC	TTGCAATTGCTTCCTGTGTC
75	GCCAGTAAGTGGGAAAGCTG	CTGGAGGAAAAAGCACAAGG
76	CAGCTCTGCCTTTGGAGAAC	AAAATTCCTTCTGTGCCTGC
77	AGAACTGGGAGTGTCCATGC	TCTTCAGCAACTCTGCCAAG
78	TTTGAGGCCCTGAGTTATGG	AGTAACTGTGCCTTGTGGGG
79	TGCCGTTTTAATCTGGTTCC	TCTTCTTTTGGACACCTGGC
80	AAACCTGGCAAGCACAGTTC	TTGCTTATCTAGGCATGGGC
81	TGGCTGACTGAATAAAGGGC	ATTGGACTGCTGGCCACTAC
82	GAGGGAGGGAGGGAAAGAC	CTCCCTCAGTTACCCTGCTG
83	TTGCAGTGCAGAGCAAAATC	CTCCGGTTACAGGTTTGAGC
84	TTTAGTGGTGGTTGCAGTGG	AGTGGTAATGTTCATGGGGC
85	AGGGCACGTGAGATACAGAC	GGCTTTTAAACACCCCTTGG
86	GGTGAAATTTCCACAATGGG	AGGGCCCTTATCTCTCTTGC
87	TAGGGACTGCCAAAAACTCC	TTTAGTTCCAGCATTTGGGC

HPLC-purified oligonucleotides were obtained from MWG Biotech AG (Ebersberg, Germany). Lyophilized primer were dissolved in H_2O to a final concentration of 100 µM and stored frozen at -20°C. All PCR products were amplified in a thermal cycler (TC9700, Perkin Elmer) under the conditions described below. Reaction mixtures of 50 µl contained 200 µM of each dNTP (Roche Molecular Biochemicals, Mannheim, Germany), 20 pmol of each primer, 100 ng of a genomic DNA preparation, 1× PCR buffer and 2U AmpliTaq (Perkin Elmer). Since the primers were selected to have the same annealing temperature, PCR products were amplified under the same conditions as following: initial denaturation (5 min at 94°C) followed by 35 cycles of denaturation (1 min at 94°C), annealing (40 sec at 59°C) and extension (2 min at 72°C), a final extension step was carried out for 5 min at 72°C. PCR products were ethanol precipitated, dissolved in 3xSSC/1.5M betaine and spotted on epoxy-coated slides (NUNC). Labeling and hybridisation of DNA from sorted chromosomes on the PCR amplicon subarray were performed as described in the Section 8.1.1.2.

9.1.2.7 Long-range PCR and sequencing

Long-range PCR amplification of breakpoint-spanning fragments was performed using specific primer pairs with one primer mapping to chromosome 1, and a corresponding primer mapping to chromosome 13 (Table 8). The ExpandTM Long Template PCR system (Roche Applied Science) was employed.

Table 8: Primer pairs designed to amplify the breakpoint-spanning fragments

NO.	FORWARD PRIMER	REVERSE PRIMER
1	ACAGCCATCTTTCAACCCTG	AGATGGGCAATGAGATCCAG
2	CACTAACAGGCCCTCTCTGC	CTCAGAGGCAGATAGCCCAC
3	ACAGCCATCTTTCAACCCTG	CTCAGAGGCAGATAGCCCAC
4	ACAGCCATCTTTCAACCCTG	AGATGGGCAATGAGATCCAG
5	CATACCTGGGTTGCTTCCAC	TATTGCAGGGAGATTTTGGC
6	CTCATGGGCTTTAGCAGCTC	ATTTCCCTCCTTTGTGGGAC
7	CATACCTGGGTTGCTTCCAC	ATTTCCCTCCTTTGTGGGAC
8	CTCATGGGCTTTAGCAGCTC	TATTGCAGGGAGATTTTGGC

Amplifications were performed in a thermal cycler (TC9700, Perkin Elmer) in 20µl reactions containing 50ng of DNA, 30 pmol of each primer, 10µl of FailSafeMix (0.4mM of each dNTP, 1×PCR buffer) and 3.5U of polymerase mix, with the following cycling parameters: after initial denaturation at 92°C for 2 min, 10 cycles of denaturation 30 sec at 92°C, annealing 30 sec at 58°C and extension 5 min at 68°C, followed by 20 cycles [30 sec at 92°C, 30 sec at 58°C and 5 min at 68°C (+20sec/cycle)] and a final extension step at 68°C for 7 min.

The three PCR products (No. 3, 4, 8) were used as templates for sequencing in both directions by use of BigDye Terminator chemistry (PE Biosystems). Separation and visualisation was performed on an Applied Biosystem 3730xl DNA Analyzer.

9.1.2.8 Agarose gel electrophoresis

DNA fragments were separated and visualized by agarose gel electrophoresis. Gels of 1% agarose (Invitrogen) in TBE buffer (0.1 M Tris, 0.1 M boric acid, 2 mM EDTA) were supplemented with 0.5 µg/ml ethidium bromide. At least 0.2 volumes of gel loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, and 30% glycerol) was added to the nucleic acid solutions before loading into the wells. DNA size markers HyperLadder I 100 Lanes (Bioline) were also loaded. Gels were run at 100 V for 30-45 min. Nucleic acids were visualized and pictures were taken using the E.A.S.Y Win32 gel documentation system (Herolab, Wiesloch, Germany).

9.2 Supplementary tables

Table S1: Segmental duplication content and DNA copy number polymorphisms in 25 patients with unbalanced aberrations (200kb)*.

case No.	LCR content** upper breakpoint	CNPs*** upper breakpoint	LCR content** lower breakpoint	CNPs*** lower breakpoint	size of homologous sequence(kb)	sequence identity
1	2.464	-	5.796	-	146	0.995
2	7.625	-	8.512	-	265	0.983
3	1.156	+	3.031	+	0	0
4	0	-	0	+	0	0
5	2.591	+	4.161	+	43	0.941
6	1.808	+	2.787	-	38	0.987
7	**4.992**	**+**	**5.082**	**+**	**190**	**0.982**
8	3.67	-	0	+	0	0
9	0.009	-	0.079	-	0	0
10	0	-	0	-	0	0
11	0	-	0	+	0	0
12	0.01	-	0.016	-	0	0
13	0	-	0	-	0	0
14	0	-	0	-	0	0
15	**2.755**	**+**	**2.182**	**+**	**6**	**0.906**
16	**1.14**	**+**	**2.973**	**-**	**0**	**0**
17	0	-	0	-	0	0
18	0.144	-	0.052	+	0	0
19	1.533	+	0	-	0	0
20	2.532	-	1.987	+	123	0.929

21	0	-	0	-	0	0
22	0	-	0	+	0	0
23	0	-	0	+	0	0
24	0	-	0	-	0	0
25	0	-	0	-	0	0

*Three patients with previously known genomic disorders are shown in bold. The 25 cases are sorted by aberration size. Calculation is based on a 200 kb interval centered around the breakpoint.
**LCR (Low Copy Repeats, same as segmental duplications) content is calculated using the formula (∑Length of Duplication * Copy Number)/ Length of Clone
*** CNP: DNA Copy Number Polymorphism

Table S2: Segmental duplication content and DNA copy number polymorphisms in 41 mentally retarded patients with balanced translocation (200kb)*.

case No.	LCR content** breakpoint 1	CNPs*** breakpoint 1	LCR content** breakpoint 2	CNPs*** breakpoint 2	size of homologous sequence(kb)	sequence identity
1	0.006	-	0	-	0	0
2	0	-	0.021	-	0	0
3	0	-	0	-	0	0
4	0	-	0	+	0	0
5	0	-	0	-	0	0
6	0	-	0	-	0	0
7	0.639	-	0	-	0	0
8	0	-	0	-	0	0
9	0	-	0	-	0	0
10	0	-	0	-	0	0
11	0	-	0	-	0	0

12	0	+	0	-	0	0
13	0	-	0.497	-	0	0
14	0	-	0	-	0	0
15	0	-	0.02	-	0	0
16	0	-	0.272	-	0	0
17	0	-	0	-	0	0
18	0	-	0	-	0	0
19	0	-	0	-	0	0
20	0	-	0	-	0	0
21	3.007	-	2.676	+	0	0
22	0.057	-	1.307	+	0	0
23	0	-	0	-	0	0
24	0	-	0	-	0	0
25	0	-	0.008	-	0	0
26	0.016	-	0	-	0	0
27	0	+	0.023	+	0	0
28	0	-	0	-	0	0
29	0.032	+	0	-	0	0
30	0	-	0.175	-	0	0
31	11.858	-	0	-	0	0
32	0	-	0	-	0	0
33	0	-	0	-	0	0
34	0	-	0.007	-	0	0
35	0	+	0	-	0	0
36	0	-	0	+	0	0
37	0	-	0.008	-	0	0

38	0	-	0	-	0	0
39	0	-	0	-	0	0
40	0	-	0.517	-	0	0
41	0.208	+	0	-	0	0

*Calculation is based on a 200 kb interval centered around the breakpoint.
**LCR (Low Copy Repeats, same as segmental duplications) content is calculated using the formula (∑Length of Duplication * Copy Number)/ Length of Clone
*** CNP: DNA Copy Number Polymorphism

Table S3: Segmental duplication content and DNA copy number polymorphisms in patients with unbalanced aberrations (deVries et al., 2005) (200kb interval)*.

patient	LCR content** upper breakpoint	CNPs*** upper breakpoint	LCR content** lower breakpoint	CNPs*** lower breakpoint	Size of homologous sequence(kb)	Sequence identity
1	0.016	-	0.01	-	0	0
2	0	-	0.02	-	0	0
3	0.027	-	2.032	+	0	0
4	0.075	-	0	-	0	0
5	0.023	-	0	-	0	0
6	0	-	0	-	0	0
7	0	+	0.008	-	0	0
8	0	-	0	-	0	0
9	0	-	0	-	0	0
9	0	+	0	+	0	0
9	0	+	1.14	+	0	0
9	1.795	-	0.116	-	0	0
10	5.32	+	3.321	+	250	0.989

11	0.834	-	6.779	-	0	0
12	0.021	-	0	+	0	0
13	0.108	-	5.826	+	0	0
14	2.083	+	0	-	0	0
15	1.8	+	2.962	-	52	0.934

*Calculation is based on a 200 kb interval centered around the breakpoint.
**LCR (Low Copy Repeats, same as segmental duplications) content is calculated using the formula (∑Length of Duplication * Copy Number)/ Length of Clone
*** CNP: DNA Copy Number Polymorphism

Table S4: Segmental duplications at breakpoints of balanced and unbalanced aberrations in patients with mental retardation (200kb interval)*

	Unbalanced aberrations**	**Balanced translocations**
Total No. of aberrations	22	41
Aberrations with no LCRs***flanking the breakpoints	11 (50%)	23(56%)
Aberrations with LCRs*** flanking one breakpoint	2 (9%)	16 (39%)
Aberrations with LCRs*** flanking both breakpoints, but without homology	4 (18%)	2 (5%)
Aberrations with homologous LCRs*** flanking both breakpoints	5 (23%)	0 (0%)

*Data are based on a 200 kb interval centered around the breakpoint.
** Three cases with previous known genomic disorders (7, 15 and 16) have been excluded
**LCRs: Low Copy Repeats, same as segmental duplications

Table S5: DNA copy number polymorphisms at breakpoints of balanced and unbalanced aberrations with mental retardation (200kb interval)*

	Unbalanced aberrations**	Balanced translocation
Total No. of Aberrations	22	41
No. of Aberration without CNPs*** at the breakpoints	11 (50%)	32 (78%)
No. of Aberration with CNPs*** at one breakpoint	9(41%)	8 (20%)
No. of Aberration with CNPs*** at both breakpoints	2 (9%)	1 (2%)

*Data are based on a 200 kb interval centered around the breakpoint.
** Three cases with previously known genomic disorders (7, 15 and 16) have been excluded
**CNPs: copy number polymorphisms

Table S6: Segmental duplication content and DNA copy number polymorphisms in 25 patients with unbalanced aberrations (1Mb)*

Case No.	LCR content** upper breakpoint	CNPs*** upper breakpoint	LCR content** lower breakpoint	CNPs*** lower breakpoint	Size of homologous sequence (kb)	Sequence identity
1	2.298	-	1.185	-	218	0.988
2	3.051	+	1.873	+	549	0.979
3	1.319	+	0.925	+	138	0.983
4	0	-	0	+	0	0
5	1.332	+	1.38	+	433	0.975
6	0.695	+	0.712	-	70	0.987
7	**1.318**	**+**	**1.477**	**+**	**373**	**0.979**
8	0.737	-	0	+	0	0
9	0.002	-	0.017	-	0	0
10	0	-	0.003	-	0	0
11	0.008	-	0.016	+	0	0

12	0.055	-	0.002	-	0	0
13	0.001	-	0.01	+	0	0
14	0.022	-	0	-	0	0
15	**2.649**	**+**	**2.698**	**+**	**558**	**0.966**
16	**0.471**	**+**	**1.431**	**-**	**112**	**0.986**
17	0.003	-	0	-	0	0
18	0.03	-	2.291	+	0	0
19	0.981	+	0.002	-	0	0
20	2.035	-	1.856	+	362	0.951
21	0	-	0.065	+	0	0
22	0.008	-	0.715	+	0	0
23	0	-	0	+	0	0
24	0.004	+	0	-	0	0
25	0	+	0.004	-	0	0

*Three patients with previouly known genomic disorders are shown in bold. The 25 cases are sorted based on the aberration size. Calculation is based on a 1 Mb interval centered around the breakpoint.
**LCR (Low Copy Repeats, same as segmental duplications) content is calculated using the formula (∑Length of Duplication * Copy Number)/ Length of Clone
*** CNP: DNA Copy Number Polymorphism

Table S7: Segmental duplication content and DNA copy number polymorphisms in 41 mentally retarded patients with balanced translocations (1Mb interval)*.

case No.	LCR content** breakpoint 1	CNPs*** breakpoint 1	LCR content** breakpoint 2	CNPs*** breakpoint 2	size of homologous sequence(kb)	sequence identity
1	0.006	+	0.003	-	0	0
2	0	+	0	-	0	0

3	0	+	0.005	+	0	0
4	0.061	+	0.007	+	0	0
5	0.003	-	0.03	-	0	0
6	0.009	-	0.035	+	0	0
7	0.421	-	0	-	0	0
8	0	-	0	-	0	0
9	0	-	0	-	0	0
10	0	-	0.001	+	0	0
11	0.003	-	0	-	0	0
12	0.001	+	0.009	+	0	0
13	0.003	-	1.82	-	0	0
14	0.126	-	0	-	0	0
15	0	-	0.006	-	0	0
16	0.001	-	0.326	-	0	0
17	0.001	-	0	-	0	0
18	0.006	-	0	-	0	0
19	0	-	0.003	-	0	0
20	0.019	+	0	+	0	0
21	0.999	-	1.858	+	0	0
22	0.053	+	0.87	+	0	0
23	0.031	-	0	-	0	0
24	0.001	-	0.004	-	0	0
25	0.007	+	0.002	-	0	0
26	0.005	+	0.01	-	0	0
27	0	+	0.005	+	0	0
28	0.002	+	0.018	-	0	0

29	0.006	+	0.019	-	0	0
30	0.019	-	0.035	-	0	0
31	2.828	+	0.002	-	0	0
32	0.003	-	0	+	0	0
33	0.905	+	0.006	+	0	0
34	0	-	0.013	-	0	0
35	0.007	+	0.019	-	0	0
36	0	-	0.008	+	0	0
37	0	-	0.004	-	0	0
38	0	-	0.003	-	0	0
39	0.005	+	0	-	0	0
40	0.003	-	1.819	-	0	0
41	0.968	+	0.919	+	0	0

*Calculation is based on a 1 Mb interval centered around the breakpoint.
**LCR (Low Copy Repeats, same as segmental duplications) content is calculated using the formula (∑Length of Duplication * Copy Number)/ Length of Clone
*** CNP: DNA Copy Number Polymorphism

Table S8: Segmental duplication content and DNA copy number polymorphisms in pateints with unbalanced aberrations (deVries et al., 2005) (1Mb interval)*

patient	LCR content** upper breakpoint	CNPs*** upper breakpoint	LCR content** lower breakpoint	CNPs*** lower breakpoint	size of homologous sequence (kb)	sequence identity
1	0.033	-	0.013	+	0	0
2	0.019	-	0.008	+	0	0
3	0.008	+	0.418	+	0	0
4	0.03	-	0.004	-	0	0
5	0.005	-	0.002	-	0	0

6	0	-	0	-	0	0
7	0.002	+	0.002	+	0	0
8	0	-	0.002	-	0	0
9	0.019	+	0.0090	+	4	0.934
9	0.004	+	0	+	0	0
9	0	+	0.471	+	0	0
9	1.145	+	0.972	-	264	0.992
10	1.321	+	1.617	+	385	0.979
11	2.395	+	4.556	+	353	0.916
12	0.004	-	0	+	0	0
13	2.857	+	2.518	+	282	0.967
14	0.751	+	0.751	+	0	0
15	0.362	+	0.919	-	113	0.934

*Calculation is based on a 1 Mb interval centered around the breakpoint.
**LCR (Low Copy Repeats, same as segmental duplications) content is calculated using the formula (∑Length of Duplication * Copy Number)/ Length of Clone
*** CNP: DNA Copy Number Polymorphism

Table S9: Segmental duplications at breakpoints of balanced and unbalanced aberrations in patients with mental retardation (1Mb interval)*

	Unbalanced aberrations**	Balanced translocations
Total No. of Aberrations	22	41
No. of Aberration with no LCRs*** flanking the breakpoints	2 (9%)	3 (7%)
No. of Aberration with LCRs*** flanking one breakpoint	7 (32%)	18 (44%)
No. of Aberration with LCRs*** flanking both breakpoints, but without homology	7 (32%)	20 (49%)
No. of Aberration with homologous LCRs*** flanking both breakpoints	6 (27%)	0 (0%)

*Data are based on a 1 Mb interval centered around the breakpoint.
** Three cases with known genomic disorders (7, 15 and 16) have been excluded
**LCRs: Low Copy Repeats, same as segmental duplications

Table S10: DNA copy number polymorphisms at breakpoints of balanced and unbalanced aberrations in patients with mental retardation (1Mb interval)

	Unbalanced aberrations**	Balanced translocation
Total No. of Aberrations	22	41
No. of Aberration without CNPs*** at the breakpoints	6 (27%)	19 (46%)
No. of Aberration with CNPs*** at one breakpoint	13 (59%)	14 (34%)
No. of Aberration with CNPs*** at both breakpoints	3 (14%)	8 (20%)

*Data are based on a 1 Mb interval centered around the breakpoint.
** Three cases with known genomic disorders (7, 15 and 16) have been excluded
**CNPs: copy number polymorphisms