Type 2 *NF1* Deletions Are Highly Unusual by Virtue of the Absence of Nonallelic Homologous Recombination Hotspots and an Apparent Preference for Female Mitotic Recombination

Katharina Steinmann, David N. Cooper, Lan Kluwe, Nadia A. Chuzhanova, Cornelia Senger, Eduard Serra, Conxi Lazaro, Montserrat Gilaberte, Katharina Wimmer, Viktor-Felix Mautner, and Hildegard Kehrer-Sawatzki

Approximately 5% of patients with neurofibromatosis type 1 (NF1) exhibit gross deletions that encompass the *NF1* gene and its flanking regions. The breakpoints of the common 1.4-Mb (type 1) deletions are located within low-copy repeats (NF1-REPs) and cluster within a 3.4-kb hotspot of nonallelic homologous recombination (NAHR). Here, we present the first comprehensive breakpoint analysis of type 2 deletions, which are a second type of recurring *NF1* gene deletion. Type 2 deletions span 1.2 Mb and are characterized by breakpoints located within the *SUZ12* gene and its pseudogene, which closely flank the NF1-REPs. Breakpoint analysis of 13 independent type 2 deletions did not reveal any obvious hotspots of NAHR. However, an overrepresentation of polypyrimidine/polypurine tracts and triplex-forming sequences was noted in the breakpoint regions that could have facilitated NAHR. Intriguingly, all 13 type 2 deletions identified so far are characterized by somatic mosaicism, which indicates a positional preference for mitotic NAHR within the *NF1* gene region. Indeed, whereas interchromosomal meiotic NAHR occurs between the NF1-REPs giving rise to type 1 deletions, NAHR during mitosis appears to occur intrachromosomally between the *SUZ12* gene and its pseudogene, thereby generating type 2 deletions. Such a clear distinction between the preferred sites of mitotic versus meiotic NAHR is unprecedented in any other genomic disorder induced by the local genomic architecture. Additionally, 12 of the 13 mosaic type 2 deletions were found in females. The marked female preponderance among mosaic type 2 deletions contrasts with the equal sex distribution noted for type 1 and/or atypical *NF1* deletions. Although an influence of chromatin structure was strongly suspected, no sex-specific differences in the methylation pattern exhibited by the *SUZ12* gene were apparent that could explain the higher rate of mitotic recombination in females.

Deletions in 17q11.2 that encompass the *NF1* gene and its flanking regions constitute the most frequently recurring mutations that cause neurofibromatosis type 1 (NF1 [MIM +162200]). Indeed, ~5% of all patients with NF1 exhibit deletions of this chromosomal region.[1–3] Three subtypes of these gross *NF1* gene deletions have been noted that differ in terms of deletion size and the positions of their respective breakpoints: type 1, type 2, and atypical *NF1* deletions. The most common of these are type 1 deletions, which encompass 1.4 Mb and lead to the loss of 14 genes, including *NF1*. Type 1 deletions are mediated by nonallelic homologous recombination (NAHR) between low-copy repeats flanking the *NF1* gene region, termed NF1-REPs A and C.[4–9] Two preferred regions of NAHR have been noted within the NF1-REPs: the paralogous recombination sites PRS1 and PRS2.[7,9] Of 60 type 1 deletions so far investigated, 40 have had breakpoints within a 3.4-kb region spanning PRS2.[9] Thus, PRS2 clearly constitutes a hotspot for NAHR within NF1-REPs A and C. Of the 60 type 1 deletions, 13 have had breakpoints within PRS1, a 1.8-kb region that represents a second preferred site for NAHR within the NF1-REPs. Importantly, NAHR underlying type 1 deletions occurs preferentially during maternal meiosis.[10–12] Atypical *NF1* deletions are rather less common than type 1 deletions, and most have nonrecurrent breakpoints.[1,4,11,13–20]

Type 2 deletions constitute the second-most-common type of gross *NF1* gene deletion. They span 1.2 Mb and are characterized by breakpoints within the *SUZ12* gene (GenBank accession number NM_015355) and its pseudogene (*SUZ12P* [GenBank accession number BC047718]). *SUZ12* is separated from NF1-REP C by 30 kb, whereas an overlap of 4 kb exists between *SUZ12P* and NF1-REP A (fig. 1). Type 2 deletions lead to the loss of 13 genes; in contrast to type 1 deletions, the functional *LRRC37B* gene within the distal NF1-REP C is retained in type 2 deletions. A total of 13 patients with type 2 deletions have been reported in the literature to date.[8,21,22] Since 6 of these 13 patients

From the Institute of Human Genetics, University of Ulm, Ulm, Germany (K.S.; C.S.; H.K.-S.); Institute of Medical Genetics, School of Medicine, Cardiff University, Cardiff (D.N.C.); Department of Maxillofacial Surgery, University Hospital Eppendorf, Hamburg (L.K.; V.-F.M.); Department of Biological Sciences, University of Central Lancashire, Preston, United Kingdom (N.A.C.); Centre de Genètica Mèdica i Molecular–Institut de Recerca Oncològica–Institut d'Investigació Biomèdica de Bellvitge (IDIBELL) (E.S.), and Translational Research Laboratory–IDIBELL, Institut Català d'Oncologia (C.L.), Hospital Duran i Reynals, and Department of Dermatology, Hospital del Mar-Institut Municipal d'Assistència Sanitària (M.G.), Barcelona; and Department of Medical Genetics, Medical University of Vienna, Vienna (K.W.)

Received May 31, 2007; accepted for publication August 3, 2007; electronically published October 31, 2007.

Address for correspondence and reprints: Dr. Hildegard Kehrer-Sawatzki, Institute of Human Genetics, University of Ulm, Albert-Einstein-Allee 11, 89081 Ulm, Germany. E-mail: hildegard.kehrer-sawatzki@uni-ulm.de

Am. J. Hum. Genet. 2007;81:1201–1220. 0002-9297/2007/8106-0008$15.00
DOI: 10.1086/522089

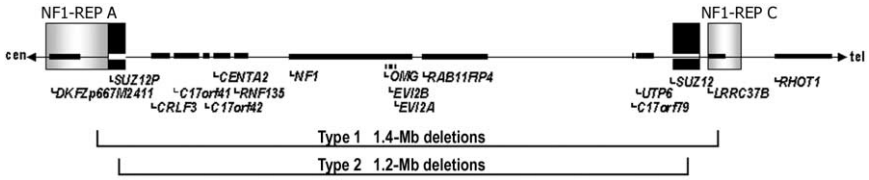

Figure 1. Map of the *NF1* gene region indicating the relative sizes and breakpoint positions of type 1 and type 2 *NF1* deletions. The gene loci are indicated as dark bars, and the positions of the NF1-REPs A and C are denoted by gray rectangles. The *SUZ12* and *SUZ12P* regions are highlighted in black. cen = Centromere; tel = telomere.

constitute mother-child pairs, they actually represent only 10 independent type 2 deletions. Remarkably, 8 of these 10 patients with type 2 deletions were somatic mosaics.[8,21] Because of this somatic mosaicism, patients with type 2 deletions frequently exhibit milder clinical manifestations of NF1. This might result in an ascertainment bias and could therefore explain the relatively small number of patients who have received a diagnosis of type 2 deletions, as compared with type 1 deletions, which are constitutional and hence are associated with a more severe clinical phenotype. So far, no somatic (mosaic) type 1 deletions have been identified unequivocally. The relatively small number of reported subjects with type 2 deletions might also be because somatic mosaicism is more difficult to detect, in practical terms.

In addition to a possible ascertainment bias due to the milder clinical phenotype, the structure, relative target size, and degree of homology between the two sets of repeats (NF1-REPs and *SUZ12/SUZ12P*) could also be responsible for the higher frequency of type 1 deletions as compared with type 2 deletions. The *SUZ12* gene and its pseudogene *SUZ12P* constitute direct continuous repeats that display 96.2% sequence identity over a length of 45 kb. By contrast, the NF1-REPs are longer, and their structure is more complicated, in that they are composed of homology blocks or subunits separated by nonhomologous regions.[5–7,16,23] NF1-REP A spans 131 kb, whereas NF1-REP C encompasses 75 kb. The largest homology block shared by NF1-REP A and NF1-REP C spans 51 kb and harbors both PRS1 and PRS2. Within this homology block, the sequence identity is 97.5%,[7] slightly higher than the 96.2% identity observed between *SUZ12* and *SUZ12P*.

Until now, the breakpoint regions of only three type 2 deletions have been precisely determined at the DNA sequence level.[8,21] It has therefore been unclear whether hotspots of recombination occur within the *SUZ12* sequences as they do in type 1 deletions, whose breakpoints cluster tightly within the NF1-REPs.

In this study, we have identified the sites of recombination underlying the remaining seven uncharacterized type 2 deletions identified in earlier studies.[8,22] Further, we characterized the breakpoints of three novel, hitherto-unreported type 2 deletions. The comparative analysis of the deletion breakpoint regions has revealed profound differences between type 1 and type 2 deletions, particularly with respect to their underlying mutational mechanisms.

Table 1. Clinical Phenotype of Three Newly Reported Female Patients with NF1 and Mosaic Type 2 Deletions

Patient	Age (years)	Clinical Investigation and Phenotype
1104	36	The patient consulted the dermatologist with respect to hyperpigmented macules, present since birth, and papulonodular lesions on macules and normal skin, which had appeared during her adolescence. Physical examination revealed hyperpigmented brownish areas with a chessboard pattern on the skin of the upper back, left shoulder, left arm, left breast, right shoulder, lumbar region, and right side of the abdomen. Axillary and inguinal freckling was absent. Four to six soft, skin-colored papulonodular lesions, 1–3 cm in diameter, were noted on the abdomen, thighs, and forearms. The histopathological investigation of one of these lesions confirmed the clinical diagnosis of a neurofibroma. Results of ophthalmological examination for Lisch nodules were negative. Abnormalities of her bones were excluded by x-ray analysis. Neurological and audiological examination also failed to reveal any anomalies. Magnetic resonance imaging (MRI) of the CNS detected multiple bihemispheric white-substance subcortical lesions, probably of vascular etiology. The angiography MRI results were, however, normal. The patient had never experienced any symptoms of stroke or neurological disease, nor was there any family history of neurofibromatosis. Her two children, aged 9 and 5 years, exhibited only light hyperpigmentation on the abdomen, as well as two and one café-au-lait spots, respectively. Results of ophthalmological examination of the children for Lisch nodules were negative, and neurological examination revealed no abnormalities.
1502	26	The patient was 168 cm in height and weighed 58 kg. Axillary and inguinal freckling and multiple café-au-lait spots but <10 subcutaneous neurofibromas were noted. She had attended secondary school but had an IQ of 90. Cerebral and whole-body MRI scans did not reveal any tumors or other anomalies. Dysmorphic facial features were not noted.
1630	15	The patient had more than six café-au-lait spots and axillary freckling but no confirmed neurofibromas, no large hands and feet, and none of the other features frequently seen in patients with gross *NF1* deletions. The patient was 176 cm in height, weighed 54 kg, and attended secondary school.

Table 2. Mosaicism in the 16 Identified Patients with Type 2 Deletions, as Determined by FISH and PCR Analysis

Patient	Mosaic	Sex	Percentage of Cells with Deletion in					
			Blood Cells		Buccal Smear	Fibroblasts	Neurofibroma	Urine
			Cultured	Uncultured				
811 (daughter)	No	F	100	…	…	…	…	…
811-M[a,b] (mother)	Yes	F	93	…	…	…	…	…
KCD-3[c]	Yes	F	92	…	…	51	…	…
697[c]	Yes	F	97	…	59	…	…	…
736[c]	Yes	F	94	…	59	…	…	…
1630[b]	Yes	F	92	…	…	…	…	…
IL39-III2 (son)	No	M	100	…	…	…	…	…
IL39[a,d] (mother)	Yes	F	70	…	…	15	…	…
1104[b]	Yes	F	84	94	8	…	…	15
SB (daughter)	No	F	100	…	…	…	…	…
WB[a,c] (mother)	Yes	F	94	…	…	…	…	…
938[c]	Yes	F	91	…	80	…	…	…
1502[b]	Yes	F	97	…	70	…	…	…
488[c]	Yes	F	98	…	56	…	…	…
928[c]	Yes	F	100	…	55	…	80	…
HC[c,e]	Yes	M	100	…	…	…	…	…

NOTE.—FISH was performed with BAC RP11-142O6 (GenBank accession number AC079915), which spans the proximal portion of the *NF1* gene, and BAC RP11-55A13 from 17q24. An ellipsis (…) = not determined.

[a] The mosaic mother passed the deletion to her offspring.

[b] Investigated in this study.

[c] As described elsewhere.[8]

[d] As described elsewhere.[21]

[e] PCR analysis of polymorphic markers, performed using DNA extracted from buccal smears, indicated that patient HC also exhibits somatic mosaicism.

Table 3. Methods Employed to Identify the Breakpoint Regions in 13 Patients with Type 2 Deletions

Patient	Breakpoint Region Identified	Breakpoint Regions Identified by		
		PCR Analysis of Somatic-Cell Hybrids[a]	Deletion-Junction PCR[b]	Array CGH
811	This study	−	+	−
KCD-3	Kehrer-Sawatzki et al.[8]	+	+	−
697	This study	+	+	−
736	This study	−	+	+
1630	This study	−	+	+
IL39	Petek et al.[21]	+	+	–
1104	This study	+	+	−
WB	Kehrer-Sawatzki et al.[8]	+	+	−
938	This study	+	+	−
1502	This study	+	+	−
488	This study	−	+	−
928	This study	+	+	−
HC	This study	+	+	−

NOTE.—A plus sign (+) = method identified a breakpoint region; a minus sign (−) = method not applied.

[a] Somatic-cell hybrids containing only the deletion-bearing chromosome 17 were analyzed by sequence analysis of the PCR products. Evaluation of PSVs that give rise to differences between the *SUZ12* pseudogene and the functional *SUZ12* gene potentiated the approximate demarcation of the region of recombination in each case.

[b] Primers were designed to allow specific amplification across the breakpoint regions, with the forward primer located in the *SUZ12* pseudogene and the reverse primer in the functional *SUZ12* gene. Sequence analysis confirmed that the proximal sequences of the breakpoint-spanning fragments are homologous to *SUZ12P* (represented by GenBank accession number AC127024), whereas the sequences at the distal end are homologous to the *SUZ12* gene (GenBank accession number AC090616.12).

Table 4. Primers Designed to Identify the Breakpoint Regions of Type 2 *NF1* Deletions with Use of Somatic-Cell Hybrid DNA Containing Only the Deleted Chromosome 17 as PCR Template

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

Material and Methods

Identification of the Type 2 Deletion–Breakpoint Regions

Elsewhere, 13 patients with type 2 deletions were identified by use of FISH analysis.[8,21,22] Here, we report three further patients in whom type 2 deletions have been identified by FISH (patients 1104, 1630, and 1502). Patients 1630 and 1502 were referred to the NF1 Clinic in Hamburg, whereas patient 1104 presented at the Department of Dermatology (Hospital del Mar-Institut Municipal d'Assistència Sanitària, Barcelona). DNA was extracted from patient material after receipt of informed consent. The clinical phenotypes of these three patients are given in table 1. All 16 known patients with type 2 deletions are listed in table 2. Of these 16 patients, 6 constitute mother-child pairs (three sets). To identify the breakpoint regions of the 10 as-yet-uncharacterized type 2 deletions, several different methods were employed (summarized in table 3). In nine cases, the breakpoints were narrowed down by sequence analysis of PCR fragments amplified from somatic-cell hybrids harboring only the deletion-containing chromosome 17. The primer sequences used to refine the respective deletion breakpoints are listed in table 4. PCR products were sequenced by means of an ABI Prism 3100 Genetic Analyzer (Applied Biosystems). To determine whether the amplified PCR fragments were derived from the *SUZ12* gene or the *SUZ12* pseudogene, paralogous sequence variants (PSVs) were analyzed as described elsewhere.[8]

To characterize the deletions present in patients 811 and 488, deletion-junction PCRs were performed using lymphocyte genomic DNA from the two patients, since no somatic-cell hybrids were available. These deletion-junction PCRs were designed to amplify across the deletion breakpoints; the forward primer was located within *SUZ12P,* and the reverse primer within the functional gene (primer sequences are listed in table 5). PCRs were performed using the Expand 20Kb Plus PCR kit (Roche Applied Science). The respective PCR products were cloned using the TOPO TA Cloning Kit for Sequencing (Invitrogen) and were completely sequenced to identify the precise regions of strand exchange. Deletion-junction PCRs were also established to confirm the breakpoint positions of the deletions first determined by the analysis of somatic-cell hybrids (table 5).

For two deletions (in patients 736 and 1630), the breakpoints were identified by comparative genomic hybridization (CGH) with HG18 CHR17 FT arrays (NimbleGen Systems). These oligonucleotide arrays are human chromosome 17–specific arrays, with a median probe spacing of 160 bp. The *SUZ12* gene and its pseudogene were represented by 153 and 58 oligonucleotide probes, respectively. Sample labeling, array manufacture, hybridization, scanning data extraction, and primary data analysis were performed by NimbleGen as described elsewhere.[24] Copy-number changes were determined by automated segmentation analysis; this involved the normalization of signal intensities of the test versus reference, with use of qspline normalization.[25] On the basis of positions of the deletion boundaries and with use of a 4-kb window size for the mean $\log_2$ values of each segment, primers were designed to amplify across the deletion breakpoints. These deletion-junction PCRs were performed using primers listed in table 5. The respective PCR products were cloned and sequenced from both ends. To complete the sequence of these deletion-junction (breakpoint-spanning) fragments of 5.5 kb (in patient 736) and 4 kb (in patient 1630), PCRs with primers listed in tables 6 and 7 were performed using cloned fragment DNA as a template. Sequencing of the respective PCR products then served to identify the regions of strand exchange during recombination.

Table 5. Primers Designed to Amplify across the Deletion Breakpoints (Deletion-Junction PCRs)

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

Table 6. Primers Designed to Complete the Cloned 4-kb Breakpoint-Spanning PCR Fragment Sequence Amplified from Genomic DNA of Patient 1630

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

Sequence Analysis of the Breakpoint Regions

DNA sequences surrounding 30 deletion-junction sites (230–540 bp in length) flanking the 13 type 2 *NF1* deletions (the "case data set" constituting the known regions of recombination between the *SUZ12* gene and pseudogene) were scanned for the presence of 37 DNA sequence motifs known to be associated with genomic instability,[26] as well as various "super-hotspot motifs" found in the vicinity of microdeletions, microinsertions, and indels.[27] The 30 analyzed breakpoint-flanking regions were derived from 22 regions flanking the 11 deletions caused by homologous recombination and 8 regions flanking the breakpoints of the 2 deletions that arose via recombination at nonhomologous sites between *SUZ12* and *SUZ12P.* In addition, complexity analysis[28] was used to generate decompositions of concatenated flanking sequences, with respect to direct and inverted repeats, as described elsewhere.[29] The origins of all fragments $\geq$8 bp in length derived from these decompositions were tracked, and the number and position of their occurrences in different flanking sequences were determined. A fragment that was found in six or more sequences, either as a direct or an inverted copy, was regarded as a potential recurring sequence motif.

Complexity analysis was also used to search for direct and inverted repeats and symmetric elements capable of forming non-B DNA structures that are known to induce DNA breakage.[30] Thus, the following sequences were sought: direct repeats $\geq$8 bp in length that were <20 bp apart and could form slipped structures; tetraplexes formed by four GGG, GGGG, or GGGGG repeats and

Table 7. Primers Designed to Complete the Cloned 5.5-kb Breakpoint-Spanning Fragment Sequence PCR Amplified from Genomic DNA of Patient 736

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

A

	Patient 697 Blood		697 Hybrid	Father		Mother		Sister	
D17S1873	142	130	130	142	126	130	134	142	130
D17S2093	159	159	159	159	159	159	159	159	159
D17S1841	270	270	270	270	272	270	278	270	270
D17S1532	169	175	175	169	169	175	169	169	175
D17S975	258	258	258	258	254	258	258	258	258
D17S1307	206	del	del	206	210	206	210	206	206
D17S1849	229	del	del	229	207	229	207	229	229
NF1PCR3	266	del	del	266	266	270	266	266	270
UT 80	202	del	del	202	198	198	194	202	198
IVs27TG24.8	272	del	del	272	280	272	280	272	272
IVs27ac28.4	207	del	del	207	205	207	213	207	207
D17S2237	400	del	del	400	404	400	404	400	400
D17S1800	278	del	del	278	264	276	276	278	276
D17S1927	150	150	150	150	150	150	150	150	150
D17S1782	93	93	93	93	93	93	93	93	93
D17S1369E	81	81	81	81	81	81	81	81	81
D17S1880	188	188	188	188	194	188	192	188	188
D17S2139	130	130	130	130	130	130	130	130	130
D17S1850	270	266	266	270	264	266	270	270	266
D17S1293	280	270	270	280	264	270	260	280	270
D17S907	286	286	286	286	326	286	322	286	286
D17S1833	156	156	156	156	154	156	154	156	156
D17S1660	220	220	220	220	220	220	220	220	220
D17S1788	156	156	156	156	156	156	156	156	156

B

	Patient HC blood		HC hybrid cell line	HC buccal smear		Father		Mother	
D17S1873	136	144	144	—		136	130	144	130
D17S1841	270	272	272	—		270	272	272	264
D17S975	254	258	258	—		254	258	258	262
D17S1307	210	del	del	—		210	210	214	210
D17S1849	207	del	del	—		207	233	207	207
NF1PCR3	266	del	del	—		266	262	262	274
IVs27TG24.8	272	del	del	272	272	272	280	272	272
IVs27ac28.4	213	del	del	213	207	213	209	207	207
D17S2237	404	del	del	—		404	400	404	404
D17S1800	268	del	del	268	276	268	280	276	280
D17S1880	188	174	174	—		188	188	174	174
D17S907	314	286	286	314	286	314	324	286	324
D17S1833	162	156	156	—		162	154	156	154
D17S1788	154	164	164	—		154	154	164	154

Figure 2. Analysis of polymorphic chromosome 17 markers to determine the mechanism underlying the type 2 deletion of patient 697 (*A*) and to investigate the somatic mosaicism exhibited by patient HC (*B*). The rectangles highlight markers located within the region that is deleted on the maternal chromosome in patients 697 and HC.

separated from each other by up to 5 bp; cruciforms formed by two inverted repeats with the minimum stem size and maximum loop length set to 8 bp and 20 bp, respectively; triplexes formed by two symmetric ⩾8-bp elements comprising at least 75% R•Y bases and separated by up to 20 bp (symmetric elements with at least 7 bp overlap were also considered to have the potential to form triplexes); left-handed Z DNA formed by GY•RC repeated at least six times; and triplet repeat sequences of the form GAA•TTC, CGG•CCG, and CTG•CAG, with a total length ⩾9 bp, which are known to induce genetic instability.

To assess the significance of our findings, 100 data sets comprising 15 random sequences matching the corresponding original breakpoint regions, in terms of their length and nucleotide composition, were simulated by reshuffling the original sequences. A *Z* score was calculated[31] for each motif and for the non-B DNA–forming sequences described above. Any *Z* score in

Table 8. Primers Used to PCR Amplify Fragments of *SUZ12* and *SUZ12P*, to Investigate the Methylation Status of CpG Dinucleotides in Males and Females

The table is available in its entirety in the online edition of *The American Journal of Human Genetics*.

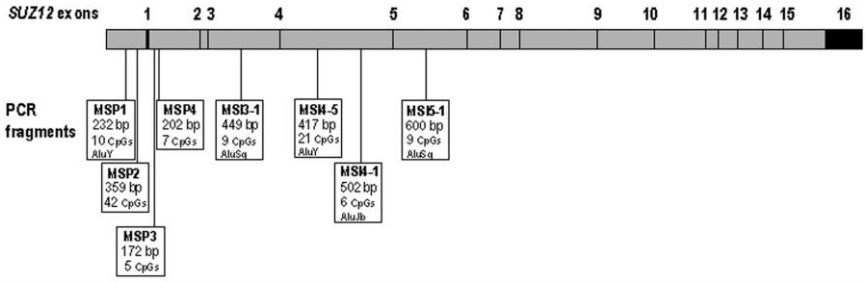

Figure 3. Locations of the eight PCR fragments, MSP1–MSI5-1, within the *SUZ12* gene. These fragments were analyzed to assess the methylation status of the *SUZ12* gene/pseudogene sequences in males and females. Numbers in bold indicate the positions of the *SUZ12* gene exons. In addition to the designation of the respective PCR fragments, their sizes (in bp) and the numbers of analyzed CpGs within each fragment are provided. Several fragment-spanning *Alu* repetitive sequences are also indicated.

the case data set that exceeded the 99th percentile of the maximum *Z* score for 100 random data sets was deemed to be statistically significant at the 1% level. Order statistics, *R-scans*,[32] were used to search for significant clustering of the breakpoint regions observed along the sequences of the functional *SUZ12* gene and *SUZ12P* by comparing their distributions with the uniform Poisson distribution. Each breakpoint region was represented by its (1) left boundary, (2) right boundary, or (3) middle position.

Polymorphic-Marker Genotyping

To determine whether the deletion found in patient 697 occurred by inter- or intrachromosomal recombination and whether patient HC exhibited somatic mosaicism, analysis of polymorphic markers on chromosome 17 (fig. 2) was performed by capillary electrophoresis on an ABI Prism 3100 (Applied Biosystems).

Determination of SUZ12 *Gene/Pseudogene Methylation Status*

DNA samples were isolated from the peripheral blood of five healthy control subjects (two males and three females). Aliquots (2 μg) of each sample were treated with the Epitect Bisulfite kit (Qiagen). For the subsequent PCR amplifications, specific primer sets (listed in table 8) were designed to amplify the sense strand of the bisulphite-converted *SUZ12* genomic DNA. In total, eight regions of the *SUZ12/SUZ12P* sequences were analyzed (fig. 3). These eight fragments contain a total of 109 CpG sites. PCR fragments were amplified from bisulphite-treated DNA from peripheral-blood samples of four healthy donors (two males and two females). For PCR fragments MSI3-1 and MSI5-1, DNA samples from three female donors were used as a template. To assess the methylation status of the CpG sites, all PCR fragments were cloned into the TOPO TA cloning vector (Invitrogen). In total, 310 positive plasmid clones were sequenced using the M13 sequencing system, and the methylation status of ~3,271 CpG sites was evaluated (fig. 4).

Analysis of SUZ12 *and* SUZ12P *in Chimpanzee*

A 15-kb deletion in the genomic sequence of chimpanzee *SUZ12P* was identified by BLAT comparison of the human *SUZ12* cDNA and genomic (GenBank accession number AC090616.12) sequences with the chimpanzee genome (Ensembl PanTro 2.1). This deletion was confirmed in both chimpanzee and bonobo by PCR analysis, with primers located in *SUZ12* intron 4 (5′-TTTAGACGCTTTCCCCAGAA-3′) and intron 5 (5′-TGATGGATTGAACTCGTGGA-3′) of the chimpanzee genomic sequence.

RT-PCR Experiments to Analyze Hypothetical SUZ12P/SUZ12 *Fusion Transcripts*

RNA was isolated using the RNeasy Mini-Kit (Qiagen) from (1) somatic-cell hybrids that contained only the deletion-bearing chromosome 17 of patients 697, 928, HC, WB, and 1502 and (2) fibroblast cultures of patient KCD. As a control, RNA was also isolated from somatic-cell hybrids that contained only the normal chromosomes of patient 697 and from blood cells of a healthy male donor. Reverse transcription was performed with the SuperScript First Strand Synthesis System (Invitrogen). RT-PCR primers used in the search for hypothetical fusion transcripts are listed in table 9.

The positions of the primers were chosen according to the analysis of ESTs derived from *SUZ12P,* together with the predicted structure of fusion transcripts (summarized in fig. 5).

Results

In addition to the 13 patients with type 2 deletions reported elsewhere,[8,21,22] we identified, by FISH, 3 further patients with type 2 deletions (patients 1104, 1630, and 1502) (tables 1 and 2). Thus, in total, 16 patients with type 2 deletions have been reported to date. Three of these patients have constitutional deletions that they inherited from their mothers, who displayed somatic mosaicism for these lesions. Thus, with the exclusion of these three inherited deletions, 13 independent de novo type 2 deletions have been identified so far, all of whom are somatic mosaics composed of cells with and without the deletion.

The figure is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Figure 4. Methylation pattern of fragments MSP1–MSI5-1, as determined in two male and two female control DNA samples. The legend is available in its entirety in the online edition of *The American Journal of Human Genetics*.

Mosaicism in 8 of these 12 patients has been demonstrated elsewhere,[8,21] whereas 5 patients were investigated for the first time, with respect to mosaicism, in this study. The relative proportions of deletion-bearing to wild-type cells in each patient, as determined by FISH and marker analysis in different tissues, are summarized in table 2. It appears that patient HC has a constitutional type 2 deletion, as determined by the FISH analysis of peripheral-blood cells.[8] In this study, buccal smear DNA from patient HC was analyzed by PCR genotyping of polymorphic markers, to demonstrate that this patient was a somatic mosaic (fig. 2).

Identification of the Breakpoint Regions in Type 2 Deletions

The breakpoint region was identified elsewhere in only 3 of the 13 known independent type 2 deletions.[8,21] In this study, we determined the precise locations of the deletion breakpoints of the remaining 10 type 2 deletions, by various methods (summarized in table 3). In each case, cloning and sequencing of the deletion junctions revealed the regions of recombination. There is invariably some ambiguity with respect to the actual breakpoint sites, because of the identity between the *SUZ12* and *SUZ12P* sequences within this region (see the "Sequence Identity in Recombination Regions" section). Despite this, flanking PSVs clearly indicate that the breakpoints must have occurred within these regions. Although the deletion breakpoints were found to occur preferentially within introns 4 and 5 of *SUZ12* (table 10), these breakpoints do not cluster in discrete hotspot regions spanning a few kilobases, as observed for type 1 deletions. Most (11 of 13) of the type 2 deletions nevertheless possess breakpoints that lie within sequences that are homologous between *SUZ12* and *SUZ12P,* which implies that these events originated as a consequence of NAHR between *SUZ12* and its pseudogene. However, two deletions (those of patients 928 and HC) have breakpoints in apparently nonhomologous regions between *SUZ12* and *SUZ12P.*

Elsewhere, we reported that three type 2 deletions (those of patients KCD-3, IL39, and WB)[8,21] were generated by intrachromosomal mitotic recombination. In the present study, SNP genotyping of patient 697 and her family indicated that this type 2 deletion was also mediated by an intrachromosomal mechanism (fig. 2).

Table 9. Primers Used for RT-PCR Analysis of Chimeric Transcripts Generated by Fusion of *SUZ12P* with *SUZ12* in Patients with Type 2 Deletions

The table is available in its entirety in the online edition of *The American Journal of Human Genetics.*

Sequence Identity in Recombination Regions

Of the 13 type 2 deletions, 11 were caused by homologous recombination between *SUZ12* and *SUZ12P.* In these cases, the recombination events occurred within regions that are 100% identical between *SUZ12* and its pseudogene. Thus,

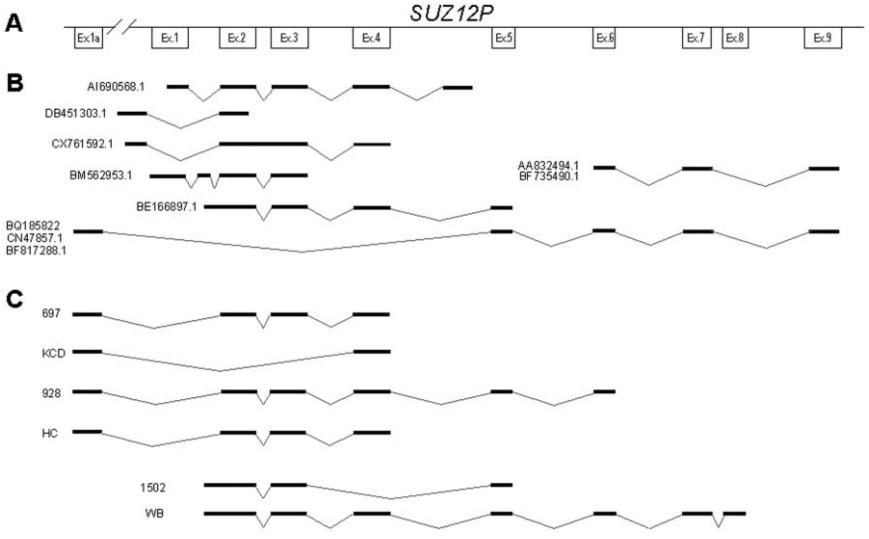

Figure 5. Alternative splice variants of the *SUZ12* pseudogene. *A, SUZ12P* exons 1–9, assigned according to BLAST sequence alignments of the functional *SUZ12* gene (GenBank accession number NM_015355 vs. accession number AC127024). Exon 1a is unique to the *SUZ12P* and has no homologue in the functional *SUZ12* gene. In panel B, the GenBank accession numbers of the *SUZ12P* ESTs are indicated on the left of each splice variant. Horizontal bars indicate exons, whereas angled lines represent splicing. The EST with the GenBank accession number BQ185822 contains an upstream exon 1a encompassing the genomic region 26.060.815–26.060.955 of chromosome 17 (Ensembl 44, Human Genome Assembly 18 [hg18], National Center for Biotechnology Information [NCBI] build 36). *C,* RT-PCR products, detected in patients 697, KCD, 928, and HC, amplified with primers KSFP1f and KSFP1r, and RT-PCR products detected in patients 1502 and WB, amplified using primers KSFP2f and KSFP1r.

Table 10. Properties of the Breakpoint Regions of 13 Type 2 *NF1* Deletions

	Breakpoint			Recombination		
Patient	Intron Position(s) in *SUZ12P/SUZ12*	Regions in *SUZ12P*[a]	Regions in *SUZ12*[a]	Type	Mechanism[b]	Sequences in Breakpoint Region(s)
811-M[c,d]	4	26093151–26093196	27297456–27297501	Homologous	ND	*Alu*Sq
KCD-3[d]	4	26095311–26095419	27299613–27299721	Homologous	Intrachromosomal	Unique
697[c,d]	4	26100313–26100381	27304631–27304699	Homologous	Intrachromosomal	*Alu*Jo
736[c,d]	4	26104591–26104678	27308894–27308981	Homologous	ND	*Alu*Sx
1630[d]	4	26108823–26108934	27313160–27313271	Homologous	ND	Unique
488[c,d]	5	26111549–26111605	27318189–27318245	Homologous	ND	*Alu*Sg
1502[d]	5	26111715–26111905	27318355–27318545	Homologous	ND	*Alu*Sc
IL39[d]	5	26115506–26115552	27322062–27322108	Homologous	Intrachromosomal	Unique
1104[c]	5	26115608–26115673	27322160–27322225	Homologous	ND	Unique
WB[d]	8	26125231–26125482	27331895–27332146	Homologous	Intrachromosomal	L1
938[c,d]	9	26127646–26127756	27334336–27334446	Homologous	ND	*Alu*Sg
928[c,d]	6	26117984	27327893	Nonhomologous	ND	Unique
HC[c,d]	4 and 10	2610411	27339771	Nonhomologous[e]	ND	*Alu*Sx and *Alu*Y

[a] According to the sequence of chromosome 17 in hg18 (March 2006), NCBI build 36.1, database version 44.36f.

[b] ND = not determined.

[c] These breakpoints were identified in this study.

[d] These patients were reported elsewhere[8] to have type 2 deletions. However, the breakpoint regions were identified[8,21] for only three of them (patients KCD-3, IL39, and WB).

[e] Although the breakpoints occurred in nonhomologous sites in *SUZ12P* and *SUZ12,* high sequence similarity between 57-bp segments of *Alu* elements was observed at the breakpoints.

the exact breakpoints cannot be precisely defined, and we refer instead to "regions of recombination" that are flanked by DNA sequences containing PSVs that can be used to distinguish between *SUZ12* and *SUZ12P.* These recombination regions are 47–253 bp in size (tables 10 and 11). When the sequence identity between *SUZ12* and *SUZ12P* is determined in windows of 500 bp, including the recombination region and surrounding sequences, the sequence identity varies between 93.6% and 97.8%, with a mean value of 95.9% (table 11). Thus, the sequences surrounding the recombination regions exhibit much lower levels of homology than do the recombination regions themselves.

Sequence Analysis of Breakpoint Regions

Two of 37 motifs known to be associated with site-specific cleavage/recombination, high-frequency mutation, and gene rearrangement[26]—namely, short (5-bp) polypurine and polypyrimidine tracts—were found to be overrepresented at the 1% level within the breakpoint regions of the type 2 deletions investigated in this study. None of the so-called super-hotspot motifs found in the vicinity of microdeletions, microinsertions, and indels, as reported elsewhere,[27] were significantly overrepresented within the breakpoint regions. However, two novel motifs (T_{10} and A_{10}) were found to be overrepresented ($P < .01$) in the vicinity of the breakpoints. In addition, symmetric T•C–rich repeats capable of non-B DNA triplex structure formation were also found to be significantly overrepresented ($P < .01$).

NAHR between the *SUZ12* gene and its pseudogene appears to be the mutational mechanism underlying 11 of the 13 type 2 deletions listed in table 10. Two deletions, those of patients 928 and HC, manifested breakpoints in apparently nonhomologous regions of the *SUZ12* sequences. Thus, for patient HC, the proximal deletion breakpoint is located within intron 4 of the *SUZ12P* pseudogene, whereas the distal deletion breakpoint is located in intron 10 of the functional *SUZ12* gene. Short interrupted homologies of <57 bp were evident at each breakpoint because of the presence of partial *AluSx* and *AluY* elements (fig. 6). It is therefore likely that the deletion in patient HC was mediated by homologous recombination between the *Alu* repeats.

The breakpoints of the deletion in patient 928 are located within different regions of intron 6 of *SUZ12P* and *SUZ12* (if the *SUZ12* and *SUZ12P* sequences are aligned, the breakpoints are separated by 3.2 kb). The sequence similarity at the breakpoints is limited to a hexanucleotide motif (TTTTGT). We propose that this deletion could have been mediated by (1) the presence of a short non-B DNA triplex structure[30] whose formation was potentiated by a TTTTGTCTGTTTT symmetric element in the vicinity of the second breakpoint, followed by (2) the binding of the disconnected strand to a hexanucleotide motif in the 3′→5′ strand at the first breakpoint (fig. 6) through the formation of another non-B DNA triplex structure. This mechanism indicates how the symmetric T•C–rich repeats found to be overrepresented in the case data set could have mediated the reported deletions. Cluster analysis failed, however, to reveal any clustering of the NAHR breakpoints within the *SUZ12* gene ($P < .05$).

Analysis of Hypothetical SUZ12P/SUZ12 *Fusion Transcripts*

The functional full-length SUZ12 protein comprises 739 aa encoded by exons 1–16. NAHR-mediated type 2 dele-

Table 11. Sequence Identity between *SUZ12* and *SUZ12P* in 500-bp Segments, Including the Recombination Regions of 13 Type 2 *NF1* Deletions

Patient	Breakpoint: Intron Position(s) in *SUZ12/SUZ12P*	Breakpoint: Region in *SUZ12P*[a]	Breakpoint: Region in *SUZ12*[a]	Length of Recombination Regions (bp)[b]	Sequence Identity[c] (%)
811-M	4	26093151–26093196	27297456–27297501	47	95.1
KCD-3	4	26095311–26095419	27299613–27299721	110	97.4
697	4	26100313–26100381	27304631–27304699	70	95.6
736	4	26104591–26104678	27308894–27308981	89	93.6
1630	4	26108823–26108914	27313160–27313251	93	96.4
488	5	26111549–26111605	27318189–27318245	58	94.0
1502	5	26111715–26111905	27318355–27318545	192	96.7
IL39	5	26115506–26115552	27322062–27322108	48	97.8
1104	5	26115608–26115673	27322160–27322225	67	94.9
WB	8	26125231–26125482	27331895–27332146	253	95.6
938	9	26127646–26127756	27334336–27334446	112	95.6
928	6	26117984	27327893	…	96.8
HC	4 and 10[d]	26104116	27339771	…	96.4[e]

[a] According to the sequence of chromosome 17 in the hg18 (March 2006), NCBI build 36.1, database version 44.36f.

[b] In the recombination regions indicated, *SUZ12* and *SUZ12P* exhibit 100% sequence identity.

[c] In 500-bp windows surrounding and including recombination regions. By way of example, in the case of the deletion of patient 811-M, we compared the sequence identity between *SUZ12* and *SUZ12P* in a 500-bp–spanning fragment containing 226 bp proximal to the recombination region, the 47 bp of the recombination region itself, and 227 bp distal to the recombination region.

[d] No analogous sequence to *SUZ12* exon 10 is present in the *SUZ12* pseudogene.

[e] For intron 4.

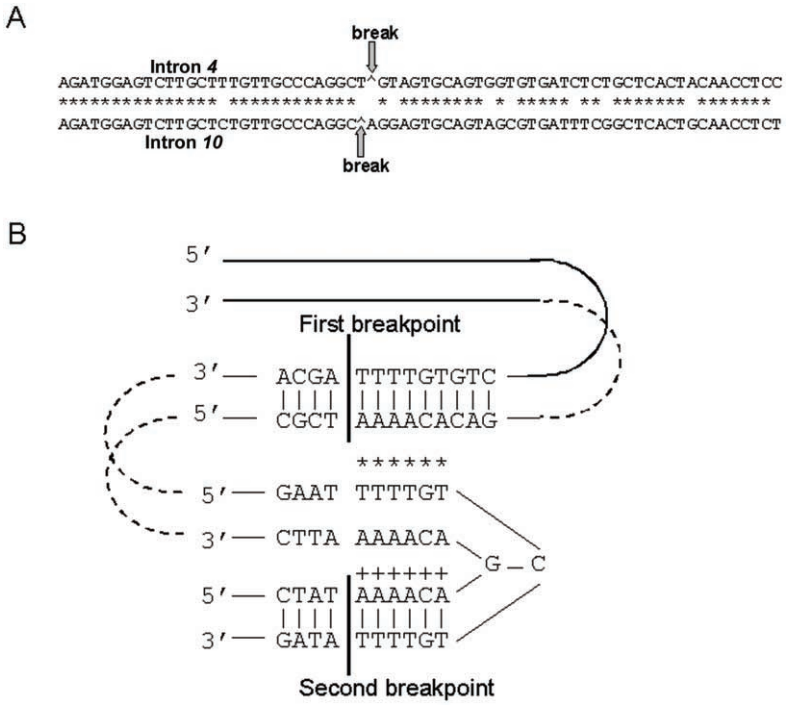

Figure 6. Sequence analysis of the deletion breakpoint regions in patients HC (*A*) and 928 (*B*). *A,* Complexity analysis revealed 57-bp stretches with 87% identity within the breakpoint regions of patient HC in intron 4 of the *SUZ12* pseudogene (*upper line*) and intron 10 (*lower line*) of the *SUZ12* gene. *B,* Hexanucleotide motif TTTTGT, found at the boundaries of the deletion in patient 928. This motif forms part of a symmetric element (TTTTGTCTGTTTT) that may have the propensity to form a non-B DNA triplex structure.

tions result in the fusion of the genomic sequences of *SUZ12P* and *SUZ12*. The resulting chimeric *SUZ12P/SUZ12* genes potentially encode SUZ12-like proteins with altered N-termini. EST analysis indicates that the *SUZ12P* gene is transcribed as shown schematically in figure 5. Taking the different splice variants of *SUZ12P* into account, we assembled hypothetical *SUZ12P/SUZ12*–fusion transcripts *in silico,* according to the relative positions of the 13 type 2 deletion breakpoints. Investigation of all three ORFs indicated that the majority of these potential fusion transcripts contain premature stop codons that would be predicted to encode truncated proteins, with little similarity to the amino acid sequence of SUZ12 (table 12).

The expression of several chimeric *SUZ12P/SUZ12* transcripts was confirmed by RT-PCR analysis of cDNAs derived from primary fibroblasts of patient KCD and from somatic-cell hybrids containing the deletion-bearing chromosome 17 of patients 697, 928, HC, 1502, and WB (fig. 5). By contrast, *SUZ12/SUZ12P*–fusion transcripts could not be amplified from cDNA derived either from somatic-cell hybrids containing only the normal chromosome 17 of patient 697 or from blood cells from a healthy donor. However, none of the *SUZ12P/SUZ12*–fusion transcripts detected in patients with type 2 deletions contain an ATG initiation codon within an optimal sequence context for translational initiation (namely, a Kozak[79] recognition sequence). This would appear to render the productive translation of these putative *SUZ12P/SUZ12*–fusion transcripts rather unlikely.

Preponderance of Females among Patients with Type 2 Deletions

Intriguingly, 12 of the 13 type 2 deletions so far identified have been found in females. To determine whether this female preponderance was statistically significant, it

Table 12. Predicted Lengths of Proteins Encoded by Hypothetical *SUZ12P/SUZ12* Fusion Transcripts Assembled *in Silico*

GenBank Accession numbers of ESTs Derived from *SUZ12P* (Tissue Origin) and ORF	Predicted Length[a] of Putative Proteins in Patients with Type 2 Deletion					
	811-M, KCD, 697, 736, and 1630[b]	488, 1502, IL39, 6, and 1104[c]	WB[d]	928[e]	938[f]	HC[g]
BQ185822 (fetal eye):						
ORF 1	63[h]	76[i]	113[i]	613[i,j,k]	113[i]	31[h]
ORF 2	63[h,l]	83[h,l]	83[h,l]	83[h,l]	83[h,l]	48[h,l]
ORF 3	634[i,l,m]	62[h,l]	62[h,l]	62[h,l]	62[h,l]	384[i,l,n]
DB451303.1 (testis):						
ORF 1	68[i]	68[i]	68[i]	68[i]	68[i]	68[i]
ORF 2	147[h,l]	147[h,l]	147[h,l]	147[h,l]	147[h,l]	147[h,l]
ORF 3	5[h,l]	5[h,l]	5[h,l]	5[h,l]	5[h,l]	5[h,l]
CX761592.1 (blastocyst):						
ORF 1	119[h,l]	119[h,l]	119[h,l]	119[h,l]	119[h,l]	119[h,l]
ORF 2	49[h,l]	49[h,l]	49[h,l]	49[h,l]	49[h,l]	49[h,l]
ORF 3	88[i]	88[i]	88[i]	88[i]	88[i]	88[i]
BM562953.1 (adenocarcinoma cell line):						
ORF 1	86[i]	86[i]	86[i]	86[i]	86[i]	86[i]
ORF 2	118[h,l]	118[h,l]	118[h,l]	118[h,l]	118[h,l]	118[h,l]
ORF 3	49[h,l]	49[h,l]	49[h,l]	49[h,l]	49[h,l]	49[h,l]
BE166897.1 (head and neck):						
ORF 1	6[h,l]	6[h,l]	433[i,k,o]	14[h]	83[h]	6[h,l]
ORF 2	11[h,l]	573[i,k,p]	11[h,l]	11[h,l]	11[h,l]	11[h,l]
ORF 3	9[h,l]	9[h,l]	9[h,l]	9[h,l]	9[h,l]	9[h,c]

NOTE.—Predictions were made in all three possible translation frames and with consideration of the different human EST-derived splice variants of *SUZ12P*.

[a] In amino acids, encoded by the various possible fusion transcripts of *SUZ12P* and *SUZ12*.
[b] Breakpoint located in intron 4.
[c] Breakpoint located in intron 5.
[d] In intron 8.
[e] In intron 6.
[f] In intron 9.
[g] In introns 4 and 10.
[h] The ORF of this transcript does not correspond to the amino acid sequence of *SUZ12*.
[i] The ORF of this transcript corresponds to the amino acid sequence of *SUZ12*.
[j] The hypothetical protein encoded comprises exons 1a, 5, and 6 of *SUZ12P* and exons 7–16 of *SUZ12*.
[k] The start codon is not in an optimal Kozak[79] sequence context.
[l] This transcript does not include a start codon.
[m] The hypothetical encoded protein comprises exon 1a of *SUZ12P* and exons 5–16 of *SUZ12*.
[n] The hypothetical encoded protein comprises exon 1a of *SUZ12P* and exons 11–16 of *SUZ12*.
[o] The hypothetical encoded protein comprises exon 5 of *SUZ12P* and exons 9–16 of *SUZ12*.
[p] The hypothetical encoded protein comprises exon 5 of *SUZ12P* and exons 6–16 of *SUZ12*.

was necessary to consider the entire cohort of 934 patients with NF1 referred to the NF1 Clinic at the Department of Maxillofacial Surgery, University of Hamburg, from which 10 of the 13 type 2 deletion patients were drawn. All these patients were investigated by analyzing polymorphic markers located within the *NF1* gene region as described elsewhere,[3] and those exhibiting homozygosity/hemizygosity for all markers were further analyzed by FISH. In total, 43 patients with gross *NF1* gene deletions were identified. Of these, 33 (77%) had type 1 or atypical *NF1* deletions, whereas 10 (23%) had type 2 deletions. All 10 patients with a type 2 deletion were female, compared with 19 females and 14 males with type 1 or atypical deletions. By comparison with type 1 and atypical deletions, the female preponderance among patients with type 2 deletions was highly significant ($P = .001$, by two-tailed Fisher's exact test). Thus, a marked difference in the sex ratios between type 1 and type 2 deletions is apparent.

Comparative Analysis of Methylation Status at the SUZ12 *Locus in Males and Females*

The marked predominance of mitotic NAHR between *SUZ12* and *SUZ12P* in females, compared with males, could be indicative of a sex-specific difference in chromatin conformation in this region of chromosome 17. Since an inverse relationship has been reported between DNA methylation and mitotic recombination,[33,34] we investigated whether there might be sex-specific differences in the degree of DNA methylation of the *SUZ12* gene sequences.

The methylation patterns exhibited by eight *SUZ12* regions were examined; four of these regions are located in the *SUZ12* gene promoter (fragments MSP1–4), with four additional regions located in introns 3, 4, and 5 (fragments MSI3-1–MSI5-1) (fig. 3). Five of these fragments encompass *Alu* sequences. PCR fragments were amplified from bisulphite-treated DNA isolated from peripheral-blood samples taken from four healthy donors (two males and two females). In the case of fragments MSI3-1 and MSI5-1, DNA samples from three female donors were analyzed. To assess the methylation status of the CpG sites in each fragment, all PCR fragments were cloned, and at least 10 clones per fragment per individual were sequenced. In total, 109 CpG sites were investigated in this way (figs. 3 and 4). The analysis of PSVs, which may be used to discriminate between *SUZ12* and *SUZ12P,* indicated that fragments MSP1, MSP2, and MSI4-1 were amplified exclusively from the *SUZ12* gene. For fragments MSP3 and MSP4, no PSVs were present that could have been used to discriminate between the *SUZ12* and *SUZ12P* sequences. Since the sequence identity between *SUZ12* and *SUZ12P* within these regions is 100%, we may infer that PCR products MSP3 and MSP4 are probably derived from both the gene and pseudogene in unknown proportions. Indeed, PSV analysis of cloned PCR fragments MSI3-1, MSI4-1, and MSI5-1 indicated that sequences derived from both *SUZ12* and *SUZ12P* were coamplified to some extent (table 13). Altogether, the methylation status of a total of 109 CpG sites in 310 different clones was investigated (fig. 4). Modeling CpG site methylation as a function of sex in a generalized model did not, however, reveal any significant difference in the overall degree of cytosine methylation between females and males ($P = .67$).

Analysis of SUZ12 *and* SUZ12 *in Chimpanzee and Macaque*

BLAST alignments of exons 1–16 of the functional *SUZ12* gene with the sequence of BAC CTD-2349P21 (GenBank accession number AC127024) spanning the genomic sequence of human *SUZ12P* indicated that the latter is a nonprocessed pseudogene comprising exons 1–9 of *SUZ12* (the *SUZ12P* pseudogene contains a premature stop codon in exon 2). Within the coding region of exons 1–9, the average sequence identity between human *SUZ12* and *SUZ12P* is 97.8%, with no evident gaps (table 14). Sequence alignments of the human *SUZ12* gene against the macaque (*Macaca mulatta*) genome sequence (MMUL 1.0, February 2006)[35] indicated that only the functional *SUZ12* gene is present in the macaque. A comparison of the coding region of human *SUZ12* with its chimpanzee (*Pan troglodytes*) orthologue revealed 99.9% sequence identity (table 15). When the *SUZ12P* coding sequences were compared between human and chimpanzee, 98% sequence identity was apparent (table 16). However, the chimpanzee *SUZ12P* pseudogene was found to lack exon 5, as well as flanking portions of introns 4 and 5. We confirmed this deletion by PCR analysis and sequencing of an 840-bp fragment spanning the deletion breakpoints. This ~15-kb deletion was also detected by PCR in the bonobo (*P. paniscus*), the sister species of *P. troglodytes,* indicating that

Table 13. Origin of the Sequenced PCR Fragments

PCR Fragment and Investigated Individual	No. of Cloned Fragments Derived from	
	SUZ12	*SUZ12P*
MSI3-1:		
Male 1	3	6
Male 2	4	5
Female 1	5	4
Female 2	5	5
MSI4-5:		
Male 1	5	1
Male 2	9	1
Female 1	7	0
Female 2	5	1
MSI5-1:		
Male 1	10	3
Male 2	9	4
Female 1	11	3
Female 2	10	3
Female 3	8	4

NOTE.—According to the analysis of PSVs between *SUZ12* and *SUZ12P*.

Table 14. Sequence Identity between Human *SUZ12* and *SUZ12P*

Exon	Length of (bp)		Sequence Identity (%)
	Exon in *SUZ12P* and *SUZ12*	Aligned Sequence	
1[a]	468	468	93.8
2	47	47	97.9
3	65	65	100
4	69	69	100
5	50	50	93.2
6	86	86	100
7	232	232	97.0
8	94	94	97.9
9	106	106	100

NOTE.—The exons of the functional *SUZ12* gene included in the Ensembl transcript ENSG00000178691 were aligned against the human genome (hg18 [March 2006]), with use of the BLAT search program at the UCSC Genome Bioinformatics Web site.

[a] Without the 5′ UTR, starting from the ATG codon.

this deletion must have predated the separation of both chimpanzee lineages 0.86–2 million years ago.[36,37]

The proximal breakpoint of the deletion observed in chimpanzee *SUZ12P* occurs within a unique sequence in intron 4, whereas the distal breakpoint is located in an *Alu*Sg sequence within *SUZ12P* intron 5. Residual homologies were, however, noted in the regions flanking the proximal and distal breakpoints in the extant human sequence (figs. 7*A* and 7*B*). In the chimpanzee ancestral sequence (after divergence from the human lineage), these homologies could have been responsible for generating double-strand breaks through the formation of non-B DNA slipped structures.[30] Although the homology in the vicinity of the proximal breakpoint appears weak, one may postulate that, over evolutionary time, the human sequence acquired mutations that have served to decrease the homology between the breakpoints to the rather limited level now evident in this region. Single-stranded DNA ends flanking the breakpoints were compared as described elsewhere.[38] We postulate that DNA ends B and C could have formed a hairpin-loop structure mediated by the 21-bp exact inverted repeats still present in the human sequence (fig. 7*C*). We further speculate that excision of this hairpin gave rise to the deletion present in the extant chimpanzee genome. A schematic representation of the structural elements that could have mediated the deletion in the chimpanzee is shown in fig. 7*D*.

When compared with the analogous sequence on human chromosome 17, this 15-kb region, which is deleted in chimpanzee *SUZ12P,* extends from position 26100965 to 26115578. This region also harbors the proximal breakpoints of 5 of the 13 type 2 deletions studied here. The breakpoints of two further type 2 deletions map close to the deletion boundaries, separated from them by 30 bp and by 584 bp (fig. 8).

Discussion

NAHR between low-copy repeats constitutes the mutational mechanism underlying a number of different genomic disorders.[39] Gross NF1 deletions also belong to this group of conditions characterized by local genomic architecture that has the propensity to undergo chromosomal rearrangement. Although the breakpoints of type 1 *NF1* deletions have been extensively characterized,[5–7,9] only three breakpoints of type 2 *NF1* deletions have been investigated in any detail.[8,21] In this study, we have characterized the breakpoints of 10 type 2 deletions and have compared the breakpoint positions with those identified elsewhere.

One of the aims of this study was to investigate whether the type 2 deletion breakpoints mapped to specific locations or whether they were widely dispersed within the *SUZ12* gene. By means of cluster analysis, no evidence of a significant NAHR hotspot for type 2 deletions was found. Instead, the breakpoint regions of the 13 characterized type 2 deletions are dispersed over ~32 kb of the 64-kb–spanning *SUZ12* gene (table 10). Although 6 of the 13 type 2 deletion breakpoints were located within intron 4, the overrepresentation of breakpoints in intron 4 is not altogether surprising, because intron 4 is the largest intron of *SUZ12,* spanning ~20.3 kb of the 64-kb *SUZ12* gene. Furthermore, the intron 4 breakpoints were not restricted to a single location; rather, they spanned a region of ~16.4 kb. Thus, in contrast to other genomic disorders that are characterized by spatially very confined NAHR hotspots (summarized in table 17), no NAHR hotspot was identified in this study within either the *SUZ12* gene or its pseudogene, even though these sequences are homologous along 45 kb of their shared length. This finding clearly serves to distinguish type 2 *NF1* deletions not merely from other known genomic disorders but also from type 1 *NF1* deletions. It is, nevertheless, worthy of note that 5 of the 13 type 2 deletions have proximal breakpoints within a region of the *SUZ12* pseudogene encompassing exon 5, the exon that has been deleted from the chimpanzee genome. Two other type 2 deletions, those of patients 1104 and 697, have proximal breakpoints that map close to the boundaries of the chimpanzee-specific *SUZ12P* deletion, at distances of 30 bp and 584 bp, respectively (fig. 8). That this region of *SUZ12P* is involved both in human mitotic recombination yielding type 2 *NF1* deletions and in mediating an evolutionary rearrangement in the chimpanzee is indicative of its general genetic instability.

Since no NAHR hotspot that could account for the majority of type 2 deletions appears to be present within the *SUZ12* gene, we sought to ascertain whether certain sequence motifs might be present in the breakpoint regions that could have contributed to double-strand breakage in these regions. The sequence analysis of the breakpoints underlying the type 2 deletions investigated here revealed the significant overrepresentation of short (5-bp) and long (≥10-bp) polypurine/polypyrimidine tracts and triplex-

Table 15. Sequence Identity of the *SUZ12* Coding Regions between Human and Chimpanzee

SUZ12 Exon	Length of (bp)		Sequence Identity (%)	Position on Chimpanzee Chromosome 17[b]
	Exon in Chimpanzee[a] and Human	Aligned Sequence		
1[c]	274	274	100	25304630–25304903
2	47	47	100	25301808–25301854
3	65	65	100	25301654–25301718
4	69	69	100	25294073–25294141
5	50	50	100	25275015–25275064
6	86	86	98.9	25267589–25267674
7	232	232	100	25265114–25265345
8	94	94	100	25264212–25264305
9	106	106	100	25257223–25257328
10	178	178	100	25251822–25251999
11	92	92	100	25247009–25247100
12	144	144	100	25246333–25246476
13	158	158	100	25245619–25245776
14	199	199	100	25244579–25244777
15	80	80	100	25243461–25243540
16[d]	346	346	99	25241342–25241687

[a] The exons of the human *SUZ12* gene included in the Ensembl transcript ENSG00000178691 were aligned against the chimpanzee genome (Ensembl PanTro 2.1 assembly [July 2006]), with use of the BLAT search program at the UCSC Genome Bioinformatics Web site.

[b] Numbering is according to the sequence of the chimpanzee genome (Ensembl PanTro 2.1 assembly [July 2006]).

[c] Without the 5′ UTR, starting from the ATG codon.

[d] Up to the TGA codon.

forming sequences. These motifs have the propensity to form non-B DNA structures that are known to be involved in chromosome breakage, mutation, recombination, and repair.[26,30,46–49] Of the 13 known NF1-associated type 2 deletions characterized here, 11 result from homologous recombination (table 10). In these cases, the breakpoints occurred within homologous regions between the *SUZ12* gene and its pseudogene. However, two deletions (those of patients HC and 928) are characterized by breakpoints that are located in apparently nonhomologous regions of *SUZ12* and *SUZ12P*. In patient HC, *Alu* sequences located in intron 4 of the *SUZ12* pseudogene and intron 10 of the functional *SUZ12* gene appear to have been responsible for mediating the deletion. *Alu-Alu* recombination has also been observed in several other deletions associated with genomic disorders, particularly those with nonrecurring breakpoints residing within the respective low-copy repeats; see, for example, the works of Inoue et al.,[50] Shaw and Lupski,[51] and Uddin et al.[52] This suggests that, whereas the nonallelic pairing of genomic regions is facilitated by higher-order genomic structures such as segmental duplications, the recombination events themselves may be triggered by shorter repetitive elements such as *Alu* sequences.[51] In patient 928, a hexanucleotide motif (TTTTGT) was noted at the boundaries of the deletion. This homology is almost certainly too short for it to have been solely responsible for mediating this deletion by homologous recombination; therefore, a nonhomologous mechanism such as nonhomologous end-joining could have been involved. However, this motif is part of a symmetric element (TTTTGTCTGTTTT) that occurs in the vicinity of the deletion breakpoints and that may have the potential to form a non-B DNA triplex structure (fig. 6). This structure could, in turn, have induced double-strand breaks, thereby initiating the process of aberrant recombination that led to the deletion in this patient. Such destabilizing DNA structures have been identified elsewhere for non-hotspot-related breakpoints of deletions in Sotos syndrome.[53]

NAHR leading to rearrangements in other genomic disorders appears to occur preferentially in 300–500-bp regions of extremely high (>99%) sequence identity.[6,40–43] Regions of uninterrupted identity within the size range of 134–232 bp have been reported to represent "minimum efficient processing segments" required for strand exchange during somatic recombination in mammalian cells.[54] However, the somatic, non-disease-related NAHR that is evident between human α-globin genes displays little requirement for the long, identical homology blocks shared by these paralogues, since exchanges occur within intervals as short as 34 bp.[55] This is in line with what we have observed for type 2 *NF1* deletions: the recombination regions characterizing 9 of the 11 deletions that have arisen by homologous recombination between *SUZ12* and *SUZ12P* are shorter than 114 bp (table 11).

Detailed analysis of the NAHR hotspot within the CMT1A repeats has indicated that this hotspot coincides spatially with a hotspot for allelic homologous recombination (AHR).[56] Intriguingly, a similar situation may be the case for NF1. Thus, whereas an NAHR hotspot has been identified at the breakpoints of type 1 *NF1* deletions,[5–7,9]

Table 16. Sequence Identity of the *SUZ12* Pseudogene between Human and Chimpanzee in Regions Homologous to the Coding Exons of the Functional *SUZ12* Gene

SUZ12P Exon[a]	Exon Length (bp)		Aligned Sequence Length (bp)	Sequence Identity (%)	Position[b]
	Chimpanzee	Human			
1	238	238	238	97.1	26582143–26582478
2	47	47	47	97.9	26579285–26579331
3	65	65	65	100	26579130–26579195
4	69	69	69	100	26571918–26571988
5	Deleted	50	…	…	…
6	86	86	86	98.9	26563142–26563216
7	231	231	231	98.3	26560657–26560889
8	94	94	94	97.9	26554152–26554262
9	106	106	106	100	26559755–26559857

[a] *SUZ12P* exons were assigned according to their homology with the coding exons of the functional *SUZ12* gene extracted from the Ensembl PanTro 2.1 chimpanzee genome assembly (July 2006).

[b] According to Ensembl PanTro 2.1, release 44.

comparative analysis of the NF1-REPs on chromosome 17 and a third repeat copy on chromosome 19 has indicated that REP19 contains an AHR hotspot in the same region as the NAHR hotspots within the NF1-REPs on chromosome 17.[9] Taken together, these findings imply that meiotic NAHR and AHR are closely related processes that can display spatial coincidence.

According to the phase I HapMap[57] and Perlegen genotype data, *SUZ12* and *SUZ12P* represent regions of low meiotic recombination frequency. This contrasts sharply with the meiotic recombination hotspot observed in the closely flanking NF1-REPs A and C. The absence of AHR hotspots could, therefore, simply be a reflection of the lack of NAHR hotspots in the *SUZ12* sequences. However, the identification of AHR hotspots might also have been hampered by the lack of high-quality genotyping data from duplicated sequences.[56,57] Thus, resequencing of the *SUZ12* gene and its pseudogene in multiple individuals would be necessary to exclude unequivocally the presence of AHR hotspots within these regions.

A quite unprecedented finding from this study was that 12 of the 13 mosaic patients with type 2 deletions were females (table 2). The overrepresentation of females among patients with type 2 deletion, compared with those with type 1 and atypical deletions, was found to be statistically significant. The underlying reason for the apparently higher rate of mitotic NAHR within the *SUZ12* sequences in females compared with males is, however, unknown. Significantly higher frequencies of female mitotic recombination at the HLA-A locus, resulting in loss of heterozygosity (LOH) and loss of expression of one of the codominant HLA-A alleles, have been reported elsewhere.[58] Further, in the asymptomatic parents of patients with facioscapulohumeral muscular dystrophy, somatic mosaicism occurs more frequently in the maternal samples.[59] Meiotic (allelic) recombination events are generally more common in females than in males.[60–63] However, highly localized sex-specific differences in recombination frequency have been reported in the context of extensive interindividual differences in both meiotic[63,64] and mitotic recombination frequencies.[58] LOH caused by mitotic recombination is a frequent mechanism that leads to somatic *NF1* inactivation and tumorigenesis in patients with NF1.[65,66] However, as yet, neither sex-specific differences in LOH frequency nor hotspots of recombination have been identified.

It has been suggested that DNA methylation serves to suppress mitotic recombination.[33,34] Indeed, hypomethylation of the immunoglobulin genes is an important requirement for allele-specific recombination during B-cell maturation,[67,68] whereas the degree of DNA methylation correlates inversely with the frequency of V(H) gene rearrangement.[69] Given the possible influence of DNA methylation on the mitotic recombination rate, we sought to compare the methylation status of the *SUZ12* gene in healthy males and females. These analyses did not, however, reveal any obvious sex-specific difference in the methylation pattern that would account for the inferred disproportionately higher frequency of mitotic NAHR leading to type 2 *NF1* deletions in females. Other epigenetic modifications such as histone methylation and deacetylation could, however, conceivably induce alterations in chromatin structure that might account for the sex-specific difference in the rate of mitotic NAHR involving the *SUZ12* gene.

SUZ12 is known to be essential for the activity of the polycomb repressive complex 2, which represses transcription through histone H3 methylation[70,71] and is required for specific gene-expression regulation during ES cell differentiation.[72] Whereas homozygous Suz12 inactivation is lethal in mouse embryos at early postimplantation stages, Suz12 heterozygous mice exhibit no obvious differences in fertility and/or morphology compared with wild-type mice.[73] The consequences of *SUZ12* hemizygosity for pa-

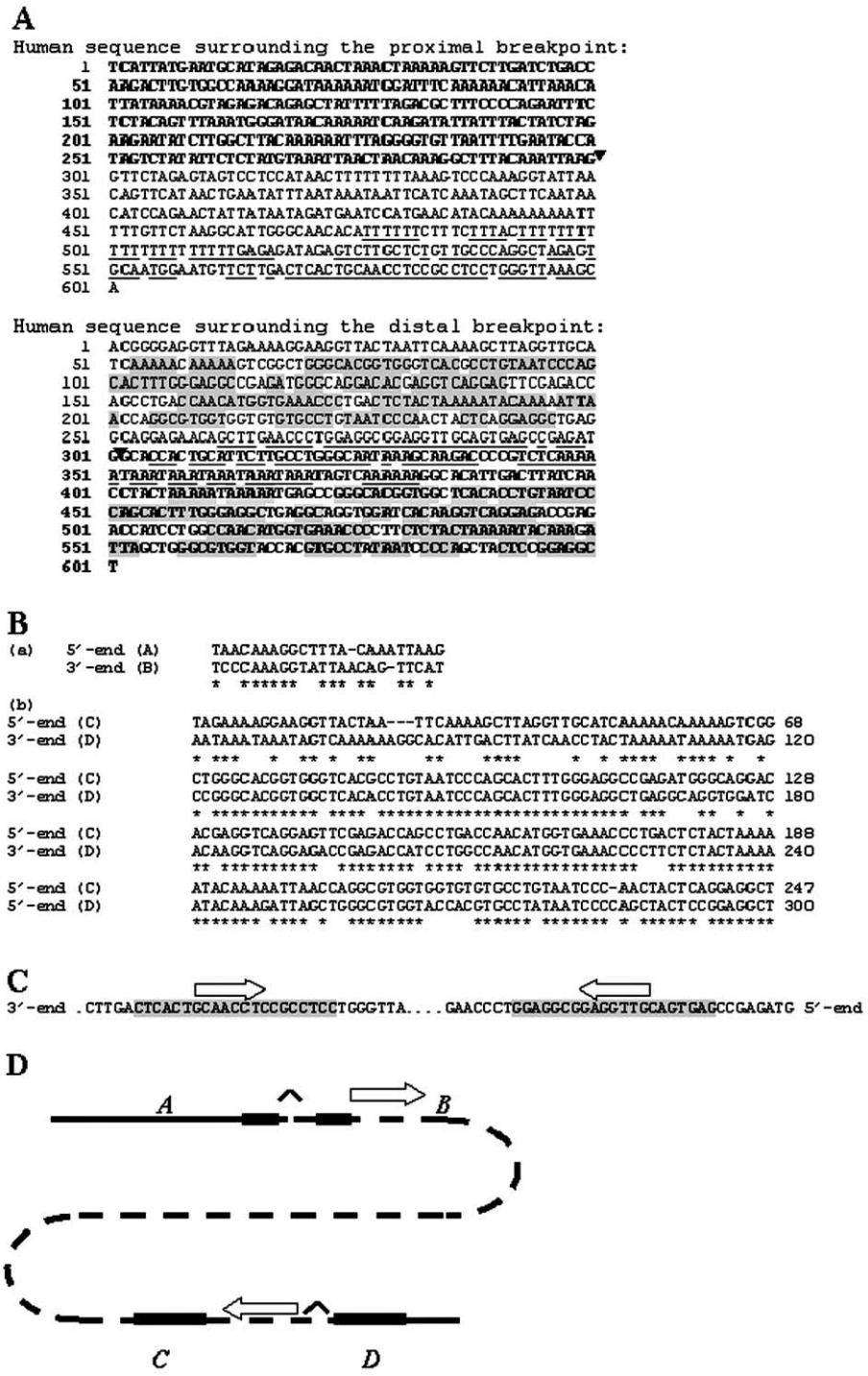

Figure 7. *A,* Extant human DNA sequence surrounding the breakpoints of the deletion detected in chimpanzee *SUZ12P*. The DNA sequence shown in bold letters is homologous to that still present in the chimpanzee genome flanking the evolutionary breakpoints, which are indicated by arrowheads. The human DNA sequence in plain font has been deleted from the chimpanzee genome. With the assumption that the human sequence is similar to the ancestral sequence, the deletion in the chimpanzee lineage could have been facilitated by direct repeats (*gray*) or inverted repeats (*underlined*). *B,* Homology evident between the regions flanking the proximal (*a*) and the distal (*b*) breakpoints in the human sequence. The uppercase letters assigned to the DNA ends are given in parentheses. *C,* Inverted repeats that could have mediated the deletion. *D,* Schematic representation of the structural elements that could have mediated the deletion. Sequences flanking the proximal and distal breakpoints are designated (1) *A* and *B* and (2) *C* and *D,* respectively. Breakpoint junctions are marked by a caret (∧). Homologous regions depicted in panel B are denoted by horizontal bars. Inverted repeats depicted in panel C and their orientations are indicated by horizontal arrows. The DNA sequence deleted as a result of the resolution of the secondary structure by excision is illustrated by broken lines.

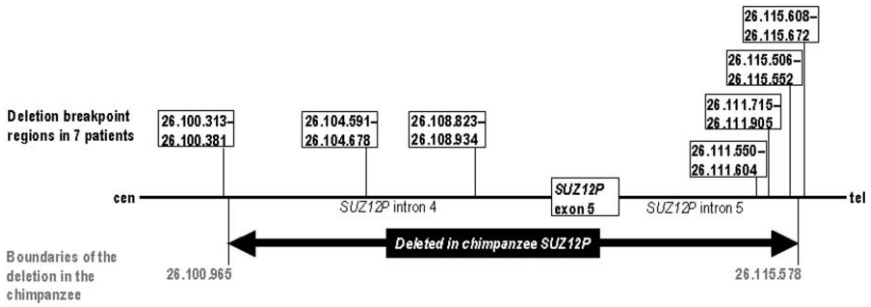

Figure 8. Scheme of the region encompassing portions of introns 4 and 5 and exon 5 of *SUZ12P* that is deleted from the chimpanzee genome (*black arrow*). The numbers indicate nucleotide positions on human chromosome 17, according to the hg18, NCBI build 36. Numbers in gray refer to the boundaries of the deletion in chimpanzee. Numbers in black indicate the recombination regions identified in seven patients with type 2 deletions.

tients with type 1 or type 2 *NF1* deletions are, however, unknown.

It may well be pertinent to note that the *SUZ12* (previously termed "*JJAZ1*") gene has been frequently found to be fused to the *JAZF1* gene on chromosome 7 in endometrial stromal tumors by the recurrent translocation t(7;17).[74–77] These somatic translocations lead to the fusion of two zinc-finger genes; the fusion product is an mRNA containing 5′-*SUZ12* and 3′-*JAZF1* sequences. It is currently unknown whether this fusion inactivates both genes or whether the fusion leads to a gain of function. Importantly, *SUZ12* is not the only fusion partner of *JAZF1,* which has also been found to be fused to the PHD finger protein 1 (*PHF1*) gene at chromosome 6p21.[78] However, the *JAZF1/PHF1*–fusion transcript is expressed under the control of the *PHF1* promoter, in contrast to the *JAZF1/SUZ12* fusion.[78] Although the molecular and cellular sequelae of the *JAZF1*-fusion transcripts remain to be identified, the recurrence of the tumor-specific t(7;17) involving the *SUZ12* gene provides further evidence of the somatic instability of the *SUZ12* gene.

The breakpoint analyses performed in this study have shown that the *SUZ12* gene and its pseudogene constitute preferred regions for mitotic NAHR in the *NF1* gene region that lead to type 2 deletions. Remarkably, RT-PCR experiments indicated that *SUZ12P/SUZ12*–fusion mRNAs are transcribed in patients with type 2 deletions. However, none of these transcripts contains an AUG translational initiation codon within an optimal context (as defined by Kozak[79]). Further detailed analyses at the protein level are necessary to assess whether these chimeric transcripts could be effectively translated into SUZ12-like proteins.

Type 1 and type 2 *NF1* deletions appear to differ not only in terms of their size and breakpoint locations but also with regard to the underlying mechanism of mutagenesis. Whereas type 1 deletions are almost exclusively caused by interchromosomal NAHR during maternal meiosis,[12] type 2 deletions appear to arise by an intrachromosomal recombinational mechanism during mitosis, as evidenced by three deletions investigated elsewhere[8,21] and the deletion in patient 697 from this study. Intrachromosomal recombination has therefore been observed in a total of four type 2 deletions investigated to date, whereas interchromosomal recombination has been demonstrated in the case of six type 1 deletions.[12] In a similar vein, whereas the germline duplication of the CMT1A region is known to arise almost exclusively via exchanges between homologous chromosomes,[40] somatic deletions involving the duplicated α-globin genes in normal human blood cells invariably occur by intrachromosomal recombination.[55] Taken together, it would appear that there are basic differences between somatic and germline NAHR, with significant roles played by homologue pairing during meiotic NAHR and intrachromosomal exchange during mitotic NAHR.

It has recently been reported that NAHR-mediated inversions between duplicated sequences occur at a much higher frequency in human somatic cells than was previously suspected.[80] These rearrangement-prone duplicate sequences are quite widespread in the human genome, although the described inversions appear not to be directly disease associated. Similarly, deletions mediated by the duplicated α-globin genes have been found in sperm cells of healthy human males.[55] Thus, duplicated sequences in the human genome possess considerable potential to mediate recurring rearrangements in somatic cells, thereby potentially contributing to genetic disease, particularly if dosage-sensitive genes are involved, as in the case of the type 2 *NF1* deletions. Remarkably, all 13 known type 2 *NF1* deletions investigated here were associated with somatic mosaicism and hence must have been mediated by mitotic NAHR. By contrast, type 1 deletions appear to be invariably constitutional and mediated by NAHR that occurs mostly during maternal meiosis.[12] Thus, a positional preference for NAHR must exist in the *NF1* gene region: during meiosis, NAHR occurs between the NF1-REPs, giving rise to type 1 deletions, whereas, during mitosis, NAHR appears to occur almost exclusively between the *SUZ12* gene and its pseudogene, thereby generating type 2 deletions. To our knowledge, such a clear distinction be-

Table 17. Summary of Human Genomic Disorders in Which NAHR Occurs within Hotspots or Preferred Regions of Recombination

			Repeat		Hotspot/Preferred Recombination Region	
Genomic Disorder	Rearrangement	Chromosomal Region	Block(s) of Shared Homology	Length (kb)	Length[a]	Frequency of NAHR (%)
Charcot-Marie-Tooth disease type 1A	1.5-Mb duplication	17p12	CMT1A-REPs	24	557 bp	90[40]
Hereditary neuropathy with liability to pressure palsies	1.5-Mb deletion					91[41]
Azoospermia (male infertility)	780-kb deletion	Yq11	HERV15 provirus sequences	~10	1.3	100[42]
NF1	1.4-Mb type 1 deletion	17q11.2	NF1-REPs A and C	51	3.4[b]	67[9]
Sotos syndrome	2-Mb deletions	5q35	Sos-PREP subunit C and Sos-DREP subunit C′	62.8 and 50.1	2.5	67[43]
Smith-Magenis syndrome	3.7-Mb deletion	17p11.2	SMS-REPs	170	12	50[44]
Reciprocal duplication	Dup(17)(p11.2p11.2)				12	23[44]
Williams-Beuren syndrome	1.55-Mb deletion	7q11.23	WBS duplicons, homology block B	105	12	27[45]

[a] In kilobases, unless indicated otherwise.
[b] Termed "PRS2."

tween the preferred sites of mitotic versus meiotic recombination has not been described in any other human gene or genomic disorder.

Acknowledgments

We thank Dr. Josef Högel, for the statistical analysis of our data, and Helene Spöri and Antje Kollak, for technical assistance. This work was funded by Deutsche Krebshilfe grant 106982.

Web Resources

Accession numbers and URLs for data presented herein are as follows:

Ensembl, http://www.ensembl.org/index.html (for *SUZ12* transcript [accession number ENSG00000178691] and the PanTro 2.1 chimpanzee genome assembly)
GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for *SUZ12* [accession number NM_015355], *SUZ12P* [accession number BC047718], cDNA genomic sequence [accession number AC090616.12], BAC CTD-2349P21 [accession number AC127024], *Homo sapiens* cDNA clone ESTs [accession numbers BQ185822, AI690568.1, DB451303.1, CX761592.1, BM562953.1, BE166897.1, CN47857.1, and BF817288.1], and BAC RP11-142O6 [accession number AC079915])
HapMap, http://www.hapmap.org/
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for NF1)
Perlegen, http://genome.perlegen.com/
UCSC Genome Bioinformatics, http://www.genome.ucsc.edu/

References

1. Cnossen MH, van der Est MN, Breuning MH, van Asperen CJ, Breslau-Siderius EJ, van der Ploeg AT, de Goede-Bolder A, van den Ouweland AM, Halley DJ, Niermeijer MF (1997) Deletions spanning the neurofibromatosis type 1 gene: implications for genotype-phenotype correlations in neurofibromatosis type 1? Hum Mutat 9:458–464
2. Rasmussen SA, Colman SD, Ho VT, Abernathy CR, Arn PH, Weiss L, Schwartz C, Saul RA, Wallace MR (1998) Constitutional and mosaic large *NF1* gene deletions in neurofibromatosis type 1. J Med Genet 35:468–471
3. Kluwe L, Siebert R, Gesk S, Friedrich RE, Tinschert S, Kehrer-Sawatzki H, Mautner V-F (2004) Screening of 500 unselected neurofibromatosis 1 patients for deletions of the *NF1* gene. Hum Mutat 23:111–116
4. Dorschner MO, Sybert VP, Weaver M, Pletcher BA, Stephens K (2000) NF1 microdeletion breakpoints are clustered at flanking repetitive sequences. Hum Mol Genet 9:35–46
5. Jenne DE, Tinschert S, Reimann H, Lasinger W, Thiel G, Hameister H, Kehrer-Sawatzki H (2001) Molecular characterization and gene content of breakpoint boundaries in patients with neurofibromatosis type 1 with 17q11.2 microdeletions. Am J Hum Genet 69:516–527
6. López-Correa C, Dorschner M, Brems H, Lazaro C, Clementi M, Upadhyaya M, Dooijes D, Moog U, Kehrer-Sawatzki H, Rutkowski JL, et al (2001) Recombination hotspot in NF1 microdeletion patients. Hum Mol Genet 10:1387–1392
7. Forbes SH, Dorschner MO, Le R, Stephens K (2004) Genomic context of paralogous recombination hotspots mediating recurrent *NF1* region microdeletion. Genes Chrom Cancer 41: 12–25
8. Kehrer-Sawatzki H, Kluwe L, Sandig C, Kohn M, Wimmer K, Krammer U, Peyrl A, Jenne DE, Hansmann I, Mautner VF (2004) High frequency of mosaicism among patients with neurofibromatosis type 1 (NF1) with microdeletions caused by somatic recombination of the *JJAZ1* gene. Am J Hum Genet 75:410–423
9. De Raedt T, Stephens M, Heyns I, Brems H, Thijs D, Messiaen L, Stephens K, Lazaro C, Wimmer K, Kehrer-Sawatzki H, et al (2006) Conservation of hotspots for recombination in low-copy repeats associated with the NF1 microdeletion. Nat Genet 38:1419–1423
10. Lazaro C, Gaona A, Ainsworth P, Tenconi R, Vidaud D, Kruyer H, Ars E, Volpini V, Estivill X (1996) Sex differences in mutational rate and mutational mechanism in the *NF1* gene in neurofibromatosis type 1 patients. Hum Genet 98:696–699
11. Upadhyaya M, Ruggieri M, Maynard J, Osborn M, Hartog C, Mudd S, Penttinen M, Cordeiro I, Ponder M, Ponder BA, et al (1998) Gross deletions of the neurofibromatosis type 1 (*NF1*) gene are predominantly of maternal origin and commonly associated with a learning disability, dysmorphic features and developmental delay. Hum Genet 102:591–597
12. López-Correa C, Brems H, Lazaro C, Marynen P, Legius E

(2000) Unequal meiotic crossover: a frequent cause of NF1 microdeletions. Am J Hum Genet 66:1969–1974
13. Kayes LM, Riccardi VM, Burke W, Bennett RL, Stephens K (1992) Large *de novo* DNA deletion in a patient with sporadic neurofibromatosis 1, mental retardation, and dysmorphism. J Med Genet 29:686–690
14. Kayes LM, Burke W, Riccardi VM, Benett R, Ehrlich P, Rubinstein A, Stephens K (1994) Deletions spanning the neurofibromatosis I gene: identification and phenotype of five patients. Am J Hum Genet 54:424–436
15. Riva P, Corrado L, Natacci F, Castorina P, Wu BL, Schneider GH, Clementi M, Tenconi R, Korf BR, Larizza L (2000) NF1 microdeletion syndrome: refined FISH characterization of sporadic and familial deletions with locus-specific probes. Am J Hum Genet 66:100–109
16. Jenne DE, Tinschert S, Dorschner MO, Hameister H, Stephens K, Kehrer-Sawatzki H (2003) Complete physical map and gene content of the human *NF1* tumor suppressor region in human and mouse. Genes Chrom Cancer 37:111–120
17. Kehrer-Sawatzki H, Tinschert S, Jenne DE (2003) Heterogeneity of breakpoints in non-LCR-mediated large constitutional deletions of the 17q11.2 *NF1* tumour suppressor region. J Med Genet 40:E116
18. Kehrer-Sawatzki H, Kluwe L, Funsterer C, Mautner VF (2005) Extensively high load of internal tumors determined by whole body MRI scanning in a patient with neurofibromatosis type 1 and a non-LCR-mediated 2-Mb deletion in 17q11.2. Hum Genet 116:466–475
19. Venturin M, Gervasini C, Orzan F, Bentivegna A, Corrado L, Colapietro P, Friso A, Tenconi R, Upadhyaya M, Larizza L, et al (2004) Evidence for non-homologous end joining and non-allelic homologous recombination in atypical *NF1* microdeletions. Hum Genet 115:69–80
20. Mantripragada KK, Thuresson AC, Piotrowski A, Diaz de Stahl T, Menzel U, Grigelionis G, Ferner RE, Griffiths S, Bolund L, Mautner V, et al (2006) Identification of novel deletion breakpoints bordered by segmental duplications in the *NF1* locus using high resolution array-CGH. J Med Genet 43:28–38
21. Petek E, Jenne DE, Smolle J, Binder B, Lasinger W, Windpassinger C, Wagner K, Kroisel PM, Kehrer-Sawatzki H (2003) Mitotic recombination mediated by the *JJAZF1* (KIAA0160) gene causing somatic mosaicism and a new type of constitutional *NF1* microdeletion in two children of a mosaic female with only few manifestations. J Med Genet 40:520–525
22. Spiegel M, Oexle K, Horn D, Windt E, Buske A, Albrecht B, Prott EC, Seemanova E, Seidel J, Rosenbaum T, et al (2005) Childhood overgrowth in patients with common *NF1* microdeletions. Eur J Hum Genet 13:883–888
23. De Raedt T, Brems H, Lopez-Correa C, Vermeesch JR, Marynen P, Legius E (2004) Genomic organization and evolution of the NF1 microdeletion region. Genomics 84:346–360
24. Selzer RR, Richmond TA, Pofahl NJ, Green RD, Eis PS, Nair P, Brothman AR, Stallings RL (2005) Analysis of chromosome breakpoints in neuroblastoma at sub-kilobase resolution using fine-tiling oligonucleotide array CGH. Genes Chrom Cancer 44:305–319
25. Olshen AB, Venkatraman ES, Lucito R, Wigler M (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 5:557–572
26. Abeysinghe SS, Chuzhanova N, Krawczak M, Ball EV, Cooper DN (2003) Translocation and gross deletion breakpoints in human inherited disease and cancer I: nucleotide composition and recombination-associated motifs. Hum Mutat 22: 229–244
27. Ball EV, Stenson PD, Krawczak M, Cooper DN, Chuzhanova NA (2005) Micro-deletions and micro-insertions causing human genetic disease: common mechanisms of mutagenesis and the role of local DNA sequence complexity. Hum Mutat 26:205–213
28. Gusev VD, Nemytikova LA, Chuzhanova NA (1999) On the complexity measures of genetic sequences. Bioinformatics 15: 994–999.
29. Chuzhanova NA, Krawczak M, Thomas N, Nemytikova LA, Gusev VD, Cooper DN (2002) The evolution of the vertebrate β-globin gene promoter. Evolution 56:154–162
30. Wells RD (2007) Non-B DNA conformations, mutagenesis, and disease. Trends Biochem Sci 32:271–278
31. Marino-Ramirez L, Spouge JL, Kanga GC, Landsman D (2004) Statistical analysis of over-represented words in human promoter sequences. Nucleic Acids Res 32:949–958
32. Karlin S, Mackin C (1991) Some statistical problems in the assessment of inhomogeneities of DNA sequence data. J Am Statist Assoc 86:27–30
33. Chen RZ, Pettersson U, Beard C, Jackson-Grusby L, Jaenisch R (1998) DNA hypomethylation leads to elevated mutation rates. Nature 395:89–93
34. Rizwana R, Hahn PJ (1999) CpG methylation reduces genomic instability. J Cell Sci 112:4513–4519
35. Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, et al (2007) Evolutionary and biomedical insights from the rhesus macaque genome. Science 316:222–234
36. Yoder AD, Yang Z (2000) Estimation of primate speciation dates using local molecular clocks. Mol Biol Evol 17:1081–1090
37. Won YJ, Hey J (2004) Divergence population genetics of chimpanzees. Mol Biol Evol 22:297–307
38. Chuzhanova N, Abeysinghe SS, Krawczak M, Cooper DN (2003) Translocation and gross deletion breakpoints in human inherited disease and cancer. II. Potential involvement of repetitive sequence elements in secondary structure formation between DNA ends. Hum Mutat 22:245–251
39. Lupski JR, Stankiewicz P (2005) Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet 1:e49
40. Lopes J, Ravise N, Vandenberghe A, Palau F, Ionasescu V, Mayer M, Levy N, Wood N, Tachi N, Bouche P, et al (1998) Fine mapping of *de novo* CMT1A and HNPP rearrangements within CMT1A-REPs evidences two distinct sex-dependent mechanisms and candidate sequences involved in recombination. Hum Mol Genet 7:141–148
41. Reiter LT, Hastings PJ, Nelis E, De Jonghe P, van Broeckhoven C, Lupski JR (1998) Human meiotic recombination products revealed by sequencing a hotspot for homologous strand exchange in multiple HNPP deletion patients. Am J Hum Genet 62:1023–1033
42. Kamp C, Hirschmann P, Voss H, Huellen K, Vogt PH (2000) Two long homologous retroviral sequence blocks in proximal Yq11 cause AZFa microdeletions as a result of intrachromosomal recombination events. Hum Mol Genet 9:2563–2572
43. Kurotaki N, Stankiewicz P, Wakui K, Niikawa N, Lupski JR (2005) Sotos syndrome common deletion is mediated by directly oriented subunits within inverted Sos-REP low-copy repeats. Hum Mol Genet 14:535–542

44. Bi W, Park SS, Shaw CJ, Withers MA, Patel PI, Lupski JR (2003) Reciprocal crossovers and a positional preference for strand exchange in recombination events resulting in deletion or duplication of chromosome 17p11.2. Am J Hum Genet 73: 1302–1315
45. Bayes M, Magano LF, Rivera N, Flores R, Perez Jurado LA (2003) Mutational mechanisms of Williams-Beuren syndrome deletions. Am J Hum Genet 73:131–151
46. Bacolla A, Jaworski A, Larson JE, Jakupciak JP, Chuzhanova N, Abeysinghe SS, O'Connell CD, Cooper DN, Wells RD (2004) Breakpoints of gross deletions coincide with non-B DNA conformations. Proc Natl Acad Sci USA 101:14162–14167
47. Bacolla A, Wells RD (2004) Non-B DNA conformations, genomic rearrangements, and human disease. J Biol Chem 279: 47411–47114
48. Pearson CE, Nichol Edamura K, Cleary JD (2005) Repeat instability: mechanisms of dynamic mutations. Nat Rev Genet 6:729–742
49. Bacolla A, Wojciechowska M, Kosmider B, Larson JE, Wells RD (2006) The involvement of non-B DNA structures in gross chromosomal rearrangements. DNA Repair (Amst) 5:1161–1170
50. Inoue K, Osaka H, Thurston VC, Clarke JT, Yoneyama A, Rosenbarker L, Bird TD, Hodes ME, Shaffer LG, Lupski JR (2002) Genomic rearrangements resulting in *PLP1* deletion occur by nonhomologous end joining and cause different dysmyelinating phenotypes in males and females. Am J Hum Genet 71:838–853
51. Shaw CJ, Lupski JR (2005) Non-recurrent 17p11.2 deletions are generated by homologous and non-homologous mechanisms. Hum Genet 116:1–7
52. Uddin RK, Zhang Y, Siu VM, Fan YS, O'Reilly RL, Rao J, Singh SM (2006) Breakpoint associated with a novel 2.3 Mb deletion in the VCFS region of 22q11 and the role of *Alu* (SINE) in recurring microdeletions. BMC Med Genet 7:18
53. Visser R, Shimokawa O, Harada N, Niikawa N, Matsumoto N (2005) Non-hotspot-related breakpoints of common deletions in Sotos syndrome are located within destabilised DNA regions. J Med Genet 42:e66
54. Waldman AS, Liskay RM (1988) Dependence of intrachromosomal recombination in mammalian cells on uninterrupted homology. Mol Cell Biol 8:5350–5357
55. Lam KW, Jeffreys AJ (2006) Processes of copy-number change in human DNA: the dynamics of α-globin gene deletion. Proc Natl Acad Sci USA 103:8921–8927
56. Lindsay SJ, Khajavi M, Lupski JR, Hurles ME (2006) A chromosomal rearrangement hotspot can be identified from population genetic variation and is coincident with a hotspot for allelic recombination. Am J Hum Genet 79:890–902
57. International HapMap Consortium (2005) A haplotype map of the human genome. Nature 437:1299–1320
58. Holt D, Dreimanis M, Pfeiffer M, Firgaira F, Morley A, Turner D (1999) Interindividual variation in mitotic recombination. Am J Hum Genet 65:1423–1427
59. van der Maarel SM, Deidda G, Lemmers RJ, van Overveld PG, van der Wielen M, Hewitt JE, Sandkuijl L, Bakker B, van Ommen GJ, Padberg GW, et al (2000) De novo facioscapulohumeral muscular dystrophy: frequent somatic mosaicism, sex-dependent phenotype, and the role of mitotic transchromosomal repeat interaction between chromosomes 4 and 10. Am J Hum Genet 66:26–35
60. Donis-Keller H, Green P, Helms C, Cartinhour S, Weiffenbach B, Stephens K, Keith TP, Bowden DW, Smith DR, Lander ES, et al (1987) A genetic linkage map of the human genome. Cell 51:319–337
61. Li W, Fann CS, Ott J (1998) Low-order polynomial trends of female-to-male map distance ratios along human chromosomes. Hum Hered 48:266–270
62. Lynn A, Kashuk C, Petersen MB, Bailey JA, Cox DR, Antonarakis SE, Chakravarti A (2000) Patterns of meiotic recombination on the long arm of human chromosome 21. Genome Res 10:1319–1332
63. Cheung VG, Burdick JT, Hirschmann D, Morley M (2007) Polymorphic variation in human meiotic recombination. Am J Hum Genet 80:526–530
64. Calabrese P (2007) A population genetics model with recombination hotspots that are heterogeneous across the population. Proc Natl Acad Sci USA 104:4748–4752
65. Serra E, Rosenbaum T, Nadal M, Winner U, Ars E, Estivill X, Lazaro C (2001) Mitotic recombination effects homozygosity for *NF1* germline mutations in neurofibromas. Nat Genet 28: 294–296
66. De Raedt T, Maertens O, Chmara M, Brems H, Heyns I, Sciot R, Majounie E, Upadhyaya M, De Schepper S, Speleman F, et al (2006) Somatic loss of wild type *NF1* allele in neurofibromas: comparison of *NF1* microdeletion and non-microdeletion patients. Genes Chrom Cancer 45:893–904
67. Hsieh C-L, Lieber MR (1992) CpG methylated minichromosomes become inaccessible for V(D)J recombination after undergoing replication. EMBO J 11:315–325
68. Engler P, Weng A, Storb U (1993) Influence of CpG methylation and target splicing on V(D)J recombination in a transgenic substrate. Mol Cell Biol 13:571–577
69. Espinoza CR, Feeney AJ (2007) Chromatin accessibility and epigenetic modifications differ between frequently and infrequently rearranging V(H) genes. Mol Immunol 44:2675–2685
70. Squazzo SL, O'Geen H, Komashko VM, Krig SR, Jin VX, Jang SW, Margueron R, Reinberg D, Green R, Farnham PJ (2006) Suz12 binds to silenced regions of the genome in a cell-type-specific manner. Genome Res 16:890–900
71. de la Cruz CC, Kirmizis A, Simon MD, Isono K, Koseki H, Panning B (2007) The Polycomb Group Protein SUZ12 regulates histone H3 lysine 9 methylation and HP1α distribution. Chromosome Res 15:299–314.
72. Pasini D, Bracken AP, Hansen JB, Capillo M, Helin K (2007) The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol Cell Biol 27:3769–3779
73. Pasini D, Bracken AP, Jensen MR, Lazzerini Denchi E, Helin K (2004) Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. EMBO J 23:4061–4071
74. Koontz JI, Soreng AL, Nucci M, Kuo FC, Pauwels P, van Den Berghe H, Cin PD, Fletcher JA, Sklar J (2001) Frequent fusion of the *JAZF1* and *JJAZ1* genes in endometrial stromal tumors. Proc Natl Acad Sci USA 98:6348–6353
75. Hrzenjak A, Moinfar F, Tavassoli FA, Strohmeier B, Kremser ML, Zatloukal K, Denk H (2005) *JAZF1*/*JJAZ1* gene fusion in endometrial stromal sarcomas: molecular analysis by reverse transcriptase-polymerase chain reaction optimized for paraffin-embedded tissue. J Mol Diagn 7:388–395
76. Nucci MR, Harburger D, Koontz J, Cin PD, Sklar J (2007) Molecular analysis of the *JAZF1-JJAZ1* gene fusion by RT-PCR

and fluorescence *in situ* hybridization in endometrial stromal neoplasms. Am J Surg Pathol 31:65–70
77. Sato K, Ueda Y, Sugaya J, Ozaki M, Hisaoka M, Katsuda S (2007) Extrauterine endometrial stromal sarcoma with *JAZF1/JJAZ1* fusion confirmed by RT-PCR and interphase FISH presenting as an inguinal tumor. Virchows Arch 450:349–353
78. Micci F, Panagopoulos I, Bjerkehagen B, Heim S (2006) Consistent rearrangement of chromosomal band 6p21 with generation of fusion genes *JAZF1*/*PHF1* and *EPC1*/*PHF1* in endometrial stromal sarcoma. Cancer Res 66:107–112
79. Kozak M (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene 361:13–37
80. Flores M, Morales L, Gonzaga-Jauregui C, Dominguez-Vidana R, Zepeda C, Yanez O, Gutierrez M, Lemus T, Valle D, Avila MC, et al (2007) Recurrent DNA inversion rearrangements in the human genome. Proc Natl Acad Sci USA 104:6099–6106