Am. J. Hum. Genet. 73:1302–1315, 2003

Reciprocal Crossovers and a Positional Preference for Strand Exchange in Recombination Events Resulting in Deletion or Duplication of Chromosome 17p11.2

Weimin Bi,[1] Sung-Sup Park,[1,*] Christine J. Shaw,[1] Marjorie A. Withers,[1] Pragna I. Patel,[1,2] and James R. Lupski[1,3,4]

Departments of [1]Molecular and Human Genetics, [2]Neurology, and [3]Pediatrics, Baylor College of Medicine, and [4]Texas Children's Hospital, Houston

Smith-Magenis syndrome (SMS) is caused by an ~4-Mb heterozygous interstitial deletion on chromosome 17p11.2 in ~80%–90% of affected patients. Three large (~200 kb), complex, and highly homologous (~98%) low-copy repeats (LCRs) are located inside or flanking the SMS common deletion. These repeats, also known as "SMS-REPs," are termed "distal," "middle," and "proximal." The directly oriented distal and proximal copies act as substrates for nonallelic homologous recombination resulting in both the deletion associated with SMS and the reciprocal duplication: dup(17)(p11.2p11.2). Using restriction enzyme *cis*-morphism analyses and direct sequencing, we mapped the regions of strand exchange in 16 somatic-cell hybrids that harbor only the recombinant SMS-REP. Our studies showed that the sites of crossovers were distributed throughout the region of homology between the distal and proximal SMS-REPs. However, despite ~170 kb of high homology, 50% of the recombinant junctions occurred in a 12.0-kb region within the *KER* gene clusters. DNA sequencing of this hotspot (positional preference for strand exchange) in seven recombinant SMS-REPs narrowed the crossovers to an ~8-kb interval. Four of them occurred in a 1,655-bp region rich in polymorphic nucleotides that could potentially reflect frequent gene conversion. For further evaluation of the strand exchange frequency in patients with SMS, novel junction fragments from the recombinant SMS-REPs were identified. As predicted by the reciprocal-recombination model, junction fragments were also identified from this hotspot region in patients with dup(17)(p11.2p11.2), documenting reciprocity of the positional preference for strand exchange. Several potential *cis*-acting recombination-promoting sequences were identified within the hotspot. It is interesting that we found 2.1-kb AT-rich inverted repeats flanking the proximal and middle *KER* gene clusters but not the distal one. The role of any or all of these in stimulating double-strand breaks around this positional recombination hotspot remains to be explored.

Introduction

The term "genomic disorders" refers to a group of human genetic diseases caused by chromosome rearrangements, including deletions, duplications, and inversions (Lupski 1998, 2003). Susceptibility to these DNA rearrangements reflects genome architecture (Inoue and Lupski 2002; Stankiewicz and Lupski 2002*a,* 2002*b*). In many genomic disorders, the rearranged genomic segments are flanked by region-specific low-copy repeats (LCRs) (also known as "segmental duplications" or "duplicons"), and the rearrangements are mediated by nonallelic homologous recombination (NAHR) that uses the flanking LCRs as substrates—for example, Smith-Magenis syndrome (SMS [MIM 182290]) (Chen et al. 1997); DiGeorge syndrome (Edelmann et al. 1999; Shaikh et al. 2000); and Williams-Beuren syndrome (Peoples et al. 2000; Valero et al. 2000]).

SMS is a constellation of multiple congenital anomalies and mental retardation associated with an interstitial deletion on chromosome 17p11.2. The incidence of SMS is reported to be ~1:25,000 births, which is probably an underestimate, given the subtle clinical features, particularly early in life (Greenberg et al. 1991). Approximately 80%–90% of patients with SMS have an ~4-Mb heterozygous deletion, termed the "SMS common deletion," whereas the remainder of patients have smaller or larger deletions (Greenberg et al. 1991; Juyal et al. 1996*b;* Trask et al. 1996; Chen et al. 1997; Bi et al. 2002; Stankiewicz et al. 2003; Potocki et al., in press). Very rarely, patients with SMS have no de-

Received August 4, 2003; accepted for publication September 16, 2003; electronically published November 24, 2003.

Address for correspondence and reprints: Dr. James R. Lupski, Department of Molecular and Human Genetics, Baylor College of Medicine, Room 604B, One Baylor Plaza, Houston, TX 77030-3498. E-mail: jlupski@bcm.tmc.edu

* Present affiliation: Department of Clinical Pathology, Seoul National University Hospital, Seoul.

0002-9297/2003/7306-0010$15.00

letions detected by FISH; mutations in the retinoic acid inducible–1 gene (*RAI1*) were recently identified in three such patients (Slager et al. 2003).

Physical mapping and DNA sequence–based analyses indicated that the SMS common-deletion interval is flanked by two highly homologous LCR gene clusters that are termed "distal" (~176 kb) and "proximal" (~256 kb) SMS repeats (SMS-REPs) (Chen et al. 1997; Park et al. 2002). They share four homology segments (sequence blocks A, B, C, and D), and the sum of these high-sequence-similarity regions is ~170 kb, with homology >98% (Park et al. 2002). Two large fragments of ~47 kb and ~39 kb in the proximal SMS-REP are absent in the distal copy. Another LCR, the middle SMS-REP (~241 kb) is inverted with reference to the proximal and distal copies and is very similar in sequence and structure to the proximal copy. The directly oriented proximal and distal SMS-REPs act as substrates for NAHR, resulting in both the SMS common deletion and the reciprocal dup(17)(p11.2p11.2), the latter being responsible for a genomic disorder distinct from SMS (Chen et al. 1997; Potocki et al. 2000; Shaw et al. 2002).

Specific genome architectural features appear to be preferred for homologous recombination between nonallelic (i.e., paralogous) LCR copies. Analyses of recombinant breakpoints for rearrangements causing several genomic disorders indicate that the majority of strand exchanges occur in a restricted region within the LCRs (i.e., a recombination hotspot [Kiyosawa et al. 1995; Lopes et al. 1996; Reiter et al. 1996; Timmerman et al. 1997; Yamamoto et al. 1997; Han et al. 2000]). Charcot-Marie-Tooth type 1A (CMT1A) and hereditary neuropathy with liability to pressure palsies (HNPP) are two peripheral neuropathies resulting from DNA rearrangements that are the reciprocal products of NAHR between the misaligned CMT1A-REP LCRs on chromosome 17p12 (Pentao et al. 1992; Chance et al. 1994; Reiter et al. 1996). Most of the CMT1A patients have an ~1.4-Mb duplication on chromosome 17p12 known as the CMT1A duplication, and the majority of HNPP patients have a reciprocal deletion of the same interval. A 557-bp region, within the ~24-kb CMT1A-REPs in 17p12, is responsible for ~78% of crossover events in CMT1A (Reiter et al. 1998). The frequency of unequal crossovers between the proximal and distal CMT1A-REPs is almost identical to that of the average equal crossovers in the human genome, as determined by sperm DNA analysis (Han et al. 2000). Given that the CMT1A-REPs do not contain a genomic hotspot for equal crossover (a conclusion based on a comparison of genetic and physical maps [Inoue et al. 2001]), the hotspot in the CMT1A-REP should be defined as a hotspot for positional preference, not for recombination frequency (Han et al. 2000). Recently, "positional preference for recombination" (i.e., hotspots) were also identified for two other microdeletion syndromes: neurofibromatosis type 1 (NF1) and Williams-Beuren syndrome (WBS). The NF1 deletion was mediated by unequal crossover between two ~85 kb NF1-REPs, in which 46% of patients had recombination events occurring in a 2-kb region (Jenne et al. 2001; López-Correa et al. 2001). Studies of 74 patients with WBS indicated that in >37% of patients, the deletion breakpoints occurred in a 12-kb region that represented only 11.4% of the total sequence of homology blocks (Bayés et al. 2003).

The SMS-REPs are much larger and more complex than CMT1A-REPs and NF1-REPs (Park et al. 2002). We sought to examine the products of recombination in SMS-REPs and to determine whether there is also positional preference for NAHR that results in SMS and its reciprocal duplication syndrome. Through breakpoint analysis in somatic-cell hybrids that retain only the common deletion allele containing the recombinant SMS-REP (and using patient genomic DNAs), we have identified a positional hotspot in the *KER* gene cluster. As anticipated, the positional recombination hotspot occurred in both del(17)(p11.2p11.2) and dup(17)(p11.2p11.2) rearrangements.

Material and Methods

Subjects and Cell Lines

Patients with SMS who had common deletions and patients with dup(17)(p11.2(p11.2) who had a reciprocal duplication were distinguished according to the presence of an ~1.1-Mb junction fragment in pulsed field gel electrophoresis (PFGE) analysis (Chen et al. 1997) and/or by use of a FISH-based assay (Stankiewicz et al. 2003), under a program approved by the institutional review board of Baylor College of Medicine. Hybrids that retain only the recombinant allele of chromosome 17 with the SMS common deletion used in this study include: 254-1D, a human-mouse somatic hybrid (Guzzetta et al. 1992); 241-8D, 244-2D, 246-27D, 251-20D, 255-11D, 263-12D, 280-22D, 283-15D, 363-1D, 389-11D, 420-7D, 424-5D, 429-6D, 468-5D, and 1361-10, human-hamster somatic-cell hybrids (Guzzetta et al. 1992; Juyal et al. 1996*a*). The human-hamster somatic-cell hybrid 484-2D is derived from a patient with SMS who had a larger-than-average deletion, in which all three SMS-REPs were deleted (Zori et al. 1993; Chen et al. 1997).

Table 1

***Cis*-Morphism Markers Used for Mapping the Regions of Strand Exchange**

Marker	Position[a]	Primer Sequence		SMS-REP[b]	Restriction Enzyme	Size[c] (bp)	
		Forward	Reverse			Before	After
A1	13,204	GCAGCTATGCCCCTGAGTAA	TCCCACAATTTTACTGAAAGGTT	Prox	*Pvu*II	248	150/98
A2	21,086	CCATTATGCCAAGCTAATGTGA	TTGCTGAATGCTATGGATAAAAA	Prox	*Hin*dIII	253	169/84
A3	44,210	TGTGCAGGAGAGAAACGATG	TTTCCAAGTTGCACAGAGTTGT	Dist	*Eco*RI	205	117/88
A4	79,621	CAAAGGCATGTCTCCATAATCA	ACCCCACCCCCTACCTTT	Dist		204	
A5	97,890	TTGGCCATGCTCTCATTACA	CTTGAAGCCCAGGCCAAAT	Dist	*Rsa*I	204	116/88
A6	115,076	CCAGGTAAAGGCTCCCTCTT	GAGCCTGAAGACAGCAAAGC	Prox	*Ban*II	398	359/39
A7	115,508	GGAGAGAAATGCCATTATCCA	ACGTCTTCAATGCGAGGTTC	Prox	*Rsa*I	296	174/122
A8	118,670	CTGCTCAAAAGGCTGTGATG	GGCTTCAGAGGTGTGGAAGA	Dist	*Hha*I	329	231/98
A9	119,774	CAAGTGAGTTTGGCAAATGCT	GCAGAAGCAGGAAAGGTTTG	Dist	*Ban*II	321	264/57
A10	122,078	GATACGACAGCTGGGGAAAG	CTGAGGACTGAAAGCCAAGG	Dist	*Hha*I	290	170/80/40
A11	122,523	AAGGAGGGAGGGAACACCTA	GTTGGGTGCTCTCCAAGTCT	Dist	*Ban*II	289	223/66
AB1	139,367	CCAGGAAGCAGAGGTCAAAG	TTGCTGTGGACTTTCAGACG	Prox		180	
AB2	164,945	CAAGATGAGACCACGCAATG	ATAAGCAAGGGCTTCAGCAA	Prox		200	
C1	200,226	AGCAGCGGAGAGAGATTCCT	GGATTAAGGGTGGAGGGAAG	Dist	*Hin*cII	398	323/75
C2	202,755	GACAGAGGCCCTACCTTGCT	TGAGTGTGTCCGCTGTCCTA	Dist	*Bgl*I	528	268/260
CD1	209,751	TTGGAAGGGGATACAATCCA	TGGTTTTAAAATGACCTGAGTGAA	Prox		208	
CD2	210,915	GGGCTCATGACCTCTGCTTA	ACATATATGCAGCCCGGAAC	Prox		139	

[a] Nucleotide numbers correspond to the proximal SMS-REP (GenBank accession number BK001591). Only the number for marker A4 corresponds to the distal SMS-REP (GenBank accession number BK001589).

[b] The distal SMS-REP (Dist) and the proximal (Prox) SMS-REPs in which the corresponding restriction enzyme is unique or wherein PCR amplification is specific.

[c] The sizes of PCR amplified fragments before and after restriction enzyme digestion.

Refining Recombinant Breakpoints in the SMS Common Deletion

Restriction enzyme sites in the distal SMS-REPs were mapped and compared with those in the proximal SMS-REPs, using Sequencher 3.1 (Gene Codes), to identify *cis*-morphisms (paralogous sequence variations [PSVs]), to enable distinction between the copies. Primers for PCR were designed to amplify the regions containing restriction *cis*-morphism, using Primer 3. The amplified products from somatic-cell hybrids were digested by restriction enzymes and were electrophoresed on a 2% agarose gel. To define a more precise location for the breakpoint within the *KER* gene cluster, DNA fragments were amplified from the recombinant SMS-REPs and sequenced. To avoid amplification outside the SMS-REP, we used only those primers that did not amplify any products from a control hybrid, 484-2D, deleted for all three SMS-REPs. The eight primer pairs for specific amplification are as follows: Kerat1F, 5′- GACTAGGAACCCTGGGGCTA-3′, and Kerat1R, 5′-TAAGCATCTTGGTGGCTGAG-3′; Kerat4F, 5′-GGGGTGAAGAGGGAAATGAT-3′, and Kerat4R, 5′-AGAATCACTGCAAGCCAAGG-3′; Kerat9F, 5′-CAAGTGCTAAGCCCCAAGAG-3′, and Kerat9R, 5′-TCAGGTCTGTCGGACTCGTA-3′; Kerat11F, 5′-CTAAGCCCAGCAACCTTCAG-3′, and Kerat11R, 5′-CTCTTTGGTGAGGGCTCCTT-3′; Kerat12F, 5′-GGCCTGTTTGTTCCTGACTC-3′, and Kerat12R, 5′-CTCCATCAAGGGCTCCTCT-3′; and Kerat13F, 5′-AAATGGGTTGCAATGTCAGG-3′, and Kerat13R, 5′-TTCCCCCATCCTATAGCTGA-3′. The primer sequences for Kerat16 and Kerat14 correspond to those shown in table 1 for markers C1 and C2, respectively.

Sequencing of the Region of Unequal Crossover in Hybrids

For sequencing the junction fragments, 25 pairs of primers were designed throughout the hotspot in *KER* gene clusters. A 16.3-kb or an 8.4-kb fragment amplified from hybrids was used as template for the sequencing of the junction fragment in hybrids. Primers for amplification of the 16.3-kb fragment were JP1F (5′-CCCCAGACTTACTCTTTATGCTGGTCTAAG-3′) and JP1R (5′-CATAGTGACCAACTGTTACGGACAGACT-3′). Primers for the 8.4-kb fragment were JP5F (5′-CATCTTCTCATACTGCTCACGCATCTCA-3′) and Kerat9R. A 50-μl PCR analysis was performed using 2.5 U TaKaRa LA *Taq* (TAKARA Bio) with 1 × PCR buffer, 0.4 mM dNTP, 10 pmol each primer, 1 μl DMSO, and 50 ng DNA template. The PCR conditions were as follows: 98°C for 30 s, 32 cycles at 94°C for 1 min, 65°C for 20 s, and 68°C for 20 min, followed by 68°C for 10 min. The PCR products were purified, using a Gel extraction kit (Qiagen), and were then used as template in PCR analyses with HotStar*Taq* (Qiagen), using the 25 pairs of primers. For sequencing, 5 μl from 25-μl PCRs was

treated with 2 μl of ExoSAP (USB Corp.) at 37°C for 15 min, and then ExoSAP was inactivated by heating at 80°C for 15 min. Sequencing was performed with an ABI PRISM BigDye Terminators version 3.0 Cycle Sequencing Kit (PE Applied Biosystems).

Southern Blotting

DNA was isolated from peripheral blood samples, somatic hybrid cell lines, and Epstein-Barr virus–transformed lymphoblastoid cell lines, using a Puregene kit (Gentra Systems). BAC DNA was prepared using a plasmid Midi kit with minor modification (Qiagen). After digestion overnight with *Bam*HI and *Spe*I, 7 μg genomic DNA or 10 ng BAC DNA was separated by electrophoresis on a 0.5% agarose gel in 0.5 × TBE buffer at 30 V. The DNA was transferred to a Hybond-XL nylon membrane (Amersham) and was hybridized to a DNA probe labeled by use of Rediprime random priming (Amersham). The probe used for detection of junction fragments in patients with SMS and dup(17)(p11.2p11.2) is a fragment amplified from genomic DNA, using primers 5′-CGTCTTCACGTCCAGAAGGA-3′ and 5′-TGAGTGTGTCCGGTGTCCTA-3′.

Results

Refining SMS Deletion Breakpoint Intervals

The SMS-REPs are extremely large (~200 kb) and highly homologous (>98%) LCRs (Park et al. 2002). In addition, the presence of multiple SMS-REPs posed the challenge of distinguishing the recombinant SMS-REP on the rearranged chromosome from the three SMS-REPs remaining on the normal nondeleted chromosome 17 homologue. To circumvent these difficulties, we initiated our breakpoint analysis by using somatic-cell hybrids that retain only the del(17)(p11.2p11.2) chromosome, which effectively isolates the recombinant SMS-REP from the other copies (Guzzetta et al. 1992; Juyal et al. 1996*a*). We confirmed that the hybrids retain only a single chromosome 17 homologue harboring the SMS deletion; this was accomplished by performing PCR aimed at identifying the presence of genes on chromosome 17 (*PMP22* and *GASTRIN*) and the absence of genes within the SMS common deletion (*COPS3, ALDH2A1,* and *PRPSAP2*) (Bi et al. 2002). Sixteen selected hybrids from patients with SMS who had the common ~4-Mb deletions were used to refine the crossover.

Two unique sequence segments in the proximal SMS-REP are absent in the distal copy. One of them, an ~47-kb region between homology blocks A and B, contains pseudogenes *NOS2, LGALS9,* and *MIP;* and the second, an ~39-kb interval between blocks C and D, harbors *DKFZp434O047* and *USP6* (fig. 1). To determine whether the crossovers occurred centromeric or telomeric to these segments, hybrids were subjected to PCR with two pairs of primers specific for each region. Among the 16 hybrids examined for recombinant breakpoints, eight strand exchanges occurred within block C, which contains the *KER* gene cluster, whereas seven were mapped telomeric to the ~47-kb region, which is potentially within block A. One breakpoint (255-11D) was located centromeric to the ~39-kb unique sequence region between blocks C and D (fig. 1).

Further refinement of breakpoints required distinctions between the distal and proximal SMS-REP sequences within homology segments (Lupski 2003). To determine from which substrate copy (proximal or distal) the site within recombinant SMS-REP was derived, we used a PCR–restriction-enzyme-digestion assay based on *cis*-morphisms, namely, restriction sites present in the distal SMS-REP but absent in the proximal copy or vice versa (table 1). Eleven primer pairs (A1–A11) were implemented for the block A homology segment. This analysis revealed that crossovers occurred throughout the entirety of block A. It is remarkable that the recombinant SMS-REPs from three hybrids (254-1D, 246-27D, and 389-11D) had their strand exchanges in a 1.1-kb region within the *NOS2A* pseudogene, which represents only 0.6% of the 170-kb homology (fig. 1). The 1.1-kb potential recombination hotspot is within block A and is only 3.6 kb from the boundary between block A and the ~47-kb proximal unique sequence.

The same approach was difficult to apply to the block C homology segment that contains the *KER* gene cluster, because paralogous copies of this region are also located outside the SMS interval (Park et al. 2002). To avoid the coamplification of *KER* sequences from regions other than the SMS-REPs, we used hybrid 484-2D, which was deleted for the 17p11.2-p12 region (Zori et al. 1993; Chen et al. 1997), as a control for the specificity of amplification. Primers for PCR were designed throughout the *KER* gene clusters, and eight primer pairs that yielded SMS-REP–specific amplification were used to refine the crossover events within block C. DNA sequences of the products amplified from the recombinant SMS-REP revealed that each of the eight strand-exchange events in block C occurred centromeric to one *cis*-morphic nucleotide at position 154,075 bp in the distal SMS-REP (GenBank accession number BK001589), located 12.0 kb away from the boundary (166,092 bp in the distal SMS-REP) between block C and the ~39-kb proximal unique sequence (fig. 1). Thus, half the hybrids (8 of 16) have their deletion-associated recombinant breakpoints within a 12.0-kb region in the *KER* gene cluster.

Sequence Analysis of the Crossover Breakpoints in Hybrids

DNA sequencing was applied for the precise determination of the sites of strand exchange in the 12.0-kb

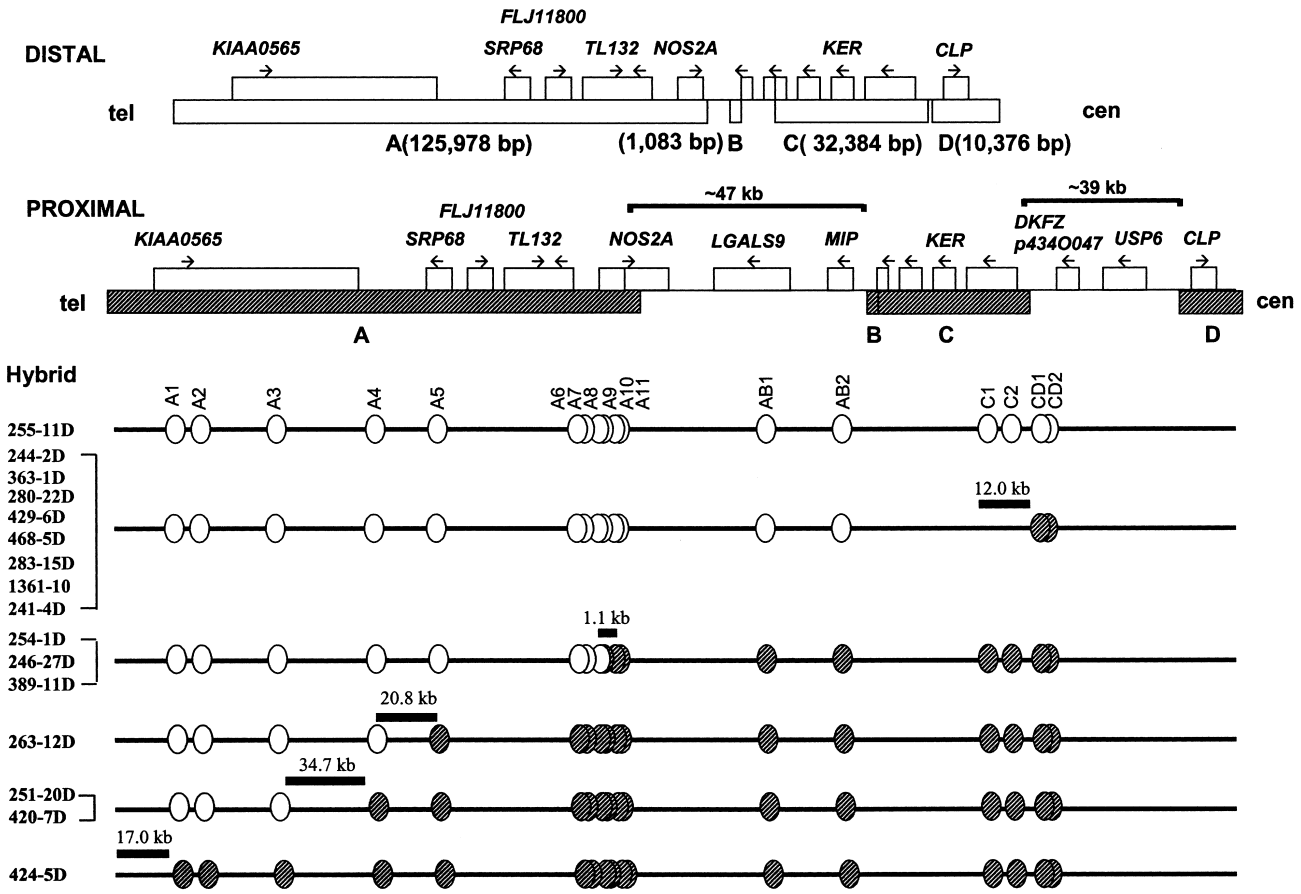

Figure 1 Refining the regions of unequal crossover in somatic-cell hybrids. *Top*, the genomic structures of SMS-REPs, with the distal and proximal copies in direct orientation, telomere (tel) (*left*) and centromere (cen) (*right*). Shown above each SMS-REP copy are the pseudogenes (*unblackened horizontal rectangles*) and the direction of transcription (*arrows*). Shown below the SMS-REPs are the homology segments. As described elsewhere (Park et al. 2002), the A, B, C, and D sequence blocks have >98% identity between the distal (*unblackened rectangles*) and the proximal (*hatched rectangles*) SMS-REP. Two segments (*horizontal brackets*) in the proximal SMS-REP (~47 kb between blocks A and B, and ~39 kb between blocks C and D) are absent in the distal copy. *Bottom*, regions of strand exchange within the recombinant SMS-REP of the deletion chromosome isolated in hybrids. Restriction *cis*-morphism markers enabling distinction between the distal and proximal SMS-REP copies are indicated (*circles*), with their positions corresponding to the proximal SMS-REP at the top of the figure. Some sites in the recombinant SMS-REP were derived from the distal SMS-REP (*unblackened circles*), and others were derived from the proximal copy (*hatched circles*). For each somatic-cell hybrid, the region of strand exchange within the recombinant SMS-REP and its size are indicated (*blackened horizontal bar*). The recombination event in hybrid 255-11D is centromeric to the *CLP* region (*D*) in the distal SMS-REP.

region within which 50% of the crossovers occurred. To avoid the coamplification of paralogous genomic regions, we first generated a 16.3-kb or an 8.4-kb PCR product that specifically encompassed the recombination junction fragments. Twenty-five primer pairs were then applied for subsequent PCR and DNA sequencing.

Nucleotide differences between the proximal and distal SMS-REPs were determined on the basis of the published sequences, and the DNA sequences from the recombinant SMS-REPs were compared, to identify the transition from "distal-like" to "proximal-like," thereby mapping the location where the strand exchange occurred (fig. 2). Within the 8.7-kb region analyzed, there were 80 nucleotide differences between the proximal SMS-REP and distal SMS-REP that could potentially be used to determine the region of exchange in the recombinant SMS-REP from the hybrids. However, 16 sites that differ between the two published clones for the distal and proximal SMS-REPs turned out not to be *cis*-morphisms, and 25 sites appeared to be polymorphic, as determined by different nucleotides found at an allelic position when recombinant SMS-REPs were sequenced from at least six different hybrids. Therefore, the latter sites were not used for mapping strand-exchange events. The remaining 39 sites, including one site that was the same in the published clones, represented true *cis*-morphisms and were thus valuable for mapping crossovers.

Seven strand exchanges within the *KER* gene cluster occurred in four intervals in an ~8-kb region (bp

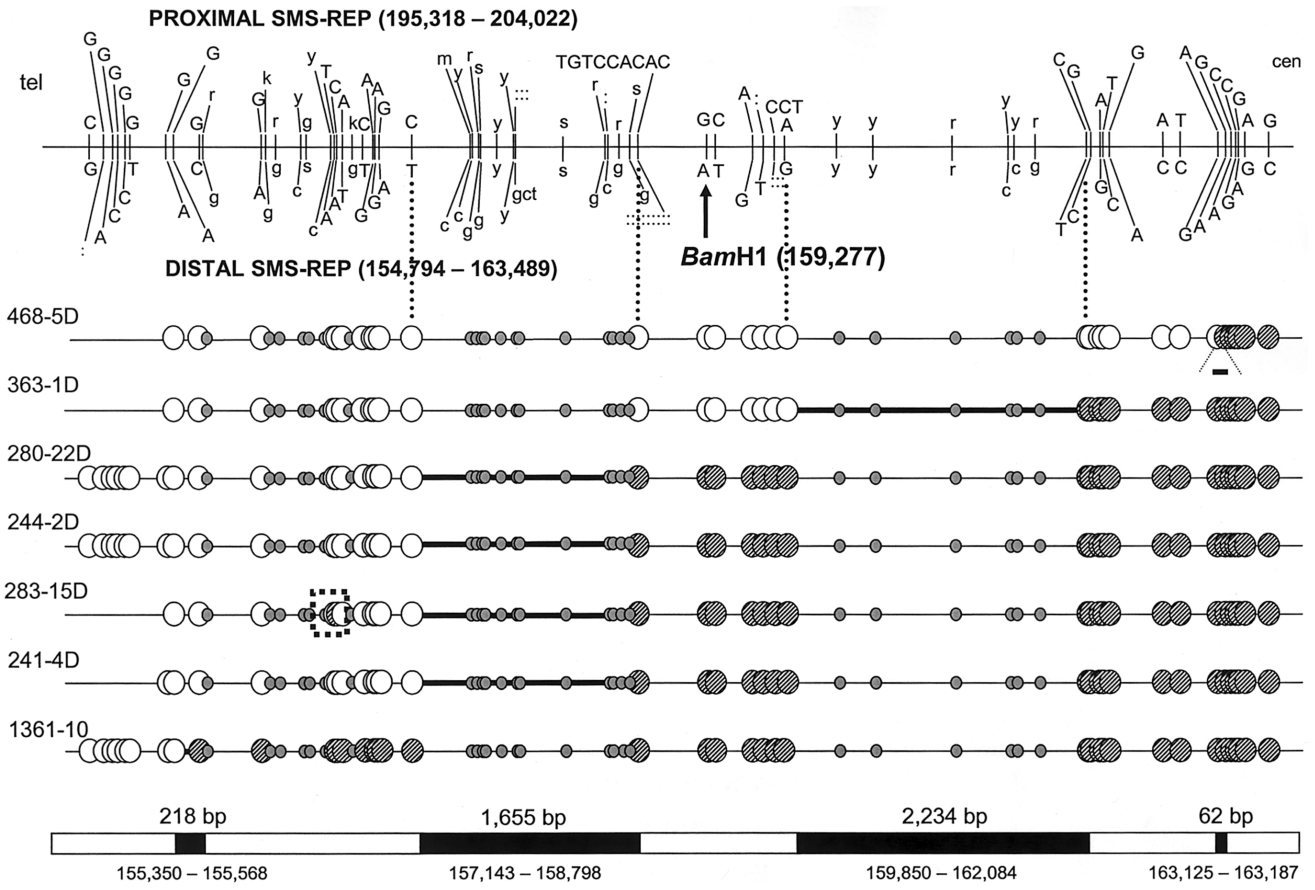

Figure 2 Sites of strand exchanges in recombinant SMS-REPs in somatic-cell hybrids. *Top,* the nucleotide differences between the proximal and distal SMS-REPs. The positions of the first and the last nucleotide difference in the 8.7-kb recombinant region of both proximal and distal SMS-REP copies are shown (see text for details). The absence of the corresponding nucleotide is denoted by a vertical dotted line. A total of 39 differences in DNA sequence (*uppercase letters*) between proximal and distal copies of SMS-REP were used for mapping the sites of strand exchange. These represent PSVs, also known as "*cis*-morphisms" (Lupski 2003), between proximal and distal copies and thus enable distinctions between the copies. Twenty-five sites are polymorphic (*lowercase letters*) and thus noninformative (*small shaded circle*) for mapping recombinants. The recombinant SMS-REP of each individual hybrid (*horizontal line*) is shown, with the hybrid number given at left. The DNA sequence differences between the distal and proximal SMS-REP are represented by circles for ease of analysis; white circles denote matches to the distal SMS-REP, and hatched circles denote matches to the proximal SMS-REP. The *cis*-morphic restriction site (*Bam*HI) used for the identification of novel junction fragments and its genomic position (in parentheses) in the distal SMS-REP are indicated. regions of strand exchange in each hybrid are highlighted using a thick black line. The dotted square on the recombinant chromosome from 283-15D reveals an apparent gene-conversion event. Vertical dotted lines demarcate regions of increased polymorphic variation, potentially occurring because of gene conversion accompany crossovers, that overlap with a positional preference for unequal crossovers. *Bottom,* summary of regions of strand exchange in each hybrid (blackened rectangles), with the nucleotide length given above. The genomic sequence positions of these strand exchange regions are marked below according to the nucleotide numbers in the distal SMS-REP.

155,350–163,187 in the distal SMS-REP) (fig. 2). Four of them (241D-4D, 244-2D, 280-22D, and 283-15D) were mapped in a 1,655-bp region that harbored 12 polymorphic bases. Sequencing comparison between the distal and proximal SMS-REPs indicated that there were only three >200-bp stretches of perfect sequence identity in the 1,655-bp region. However, one nucleotide difference that was located between a 240-bp and a 164-bp stretch of perfect identity was actually not a *cis*-morphism in all the sequenced hybrids, resulting in 404 bp of perfect identity. One nucleotide difference (363-1D) was located in a 2,234-bp interval with six polymorphic nucleotides. Recombination breakpoints for 468-5D and 1361-10 were mapped in 62-bp and 218-bp intervals, respectively. A gene-conversion event adjacent to the crossover site was detected in the recombinant SMS-REP from one hybrid (283-15D) (fig. 2).

Identification of Novel Junction Fragments

To verify whether the majority of strand exchanges occurred in the 12.0-kb region, as indicated by the recombinant studies in the SMS hybrids, we attempted genomic Southern analysis to detect novel junction fragments from

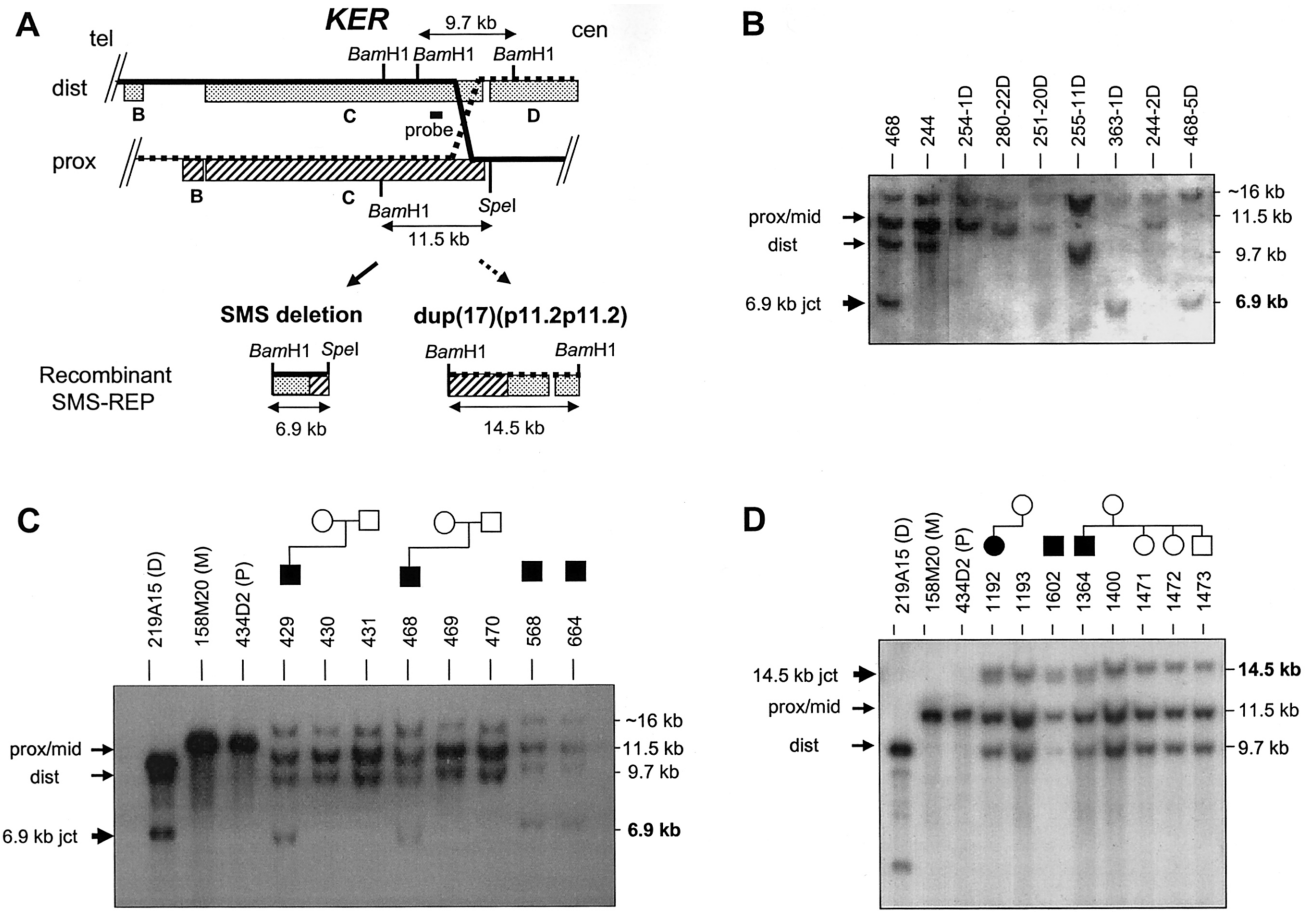

Figure 3 Junction fragments identified in patients with SMS and dup(17)(p11.2p11.2). *A,* sequence prediction of novel junction fragments from recombinant SMS-REPs. Portions of the *KER* gene cluster of distal (dist) and proximal (prox) SMS-REPs are shown with the horizontal dotted and hatched rectangles depicting homology segments. A 9.7-kb fragment from the distal SMS-REP and an 11.5-kb fragment from the proximal SMS-REP are predicted when genomic DNA is double-digested with *Bam*HI and *Spe*I and hybridized with a 906-bp probe (*short horizontal bar*). In addition, a 6.9-kb junction fragment from patients with the SMS deletion and a 14.5-kb junction fragment in its reciprocal duplication from patients with dup(17)(p11.2p11.2) are predicted. *B,* DNA samples from patients (468 and 244) and del(17)p11.2 retaining hybrids were digested with *Bam*HI and *Spe*I. A predicted 6.9-kb novel junction fragment band was observed in patient 468 and in hybrids 363-1D and 468-5D. Hybrid 255-11D had only the band from the distal SMS-REP, whereas hybrids 254-1D, 280-22D, 251-20D, and 244-2D had the band from the proximal SMS-REP. *C,* The 6.9-kb novel junction fragment cosegregates with the SMS phenotype. Only the affected individuals (blackened symbols) and neither parent (in the case of families) had the novel 6.9-kb junction fragment. The three control lanes contain BAC DNA and indicate that the 9.7-kb band was from the distal (RP11-219A15) SMS-REP, whereas both the middle (RP11-158M20) and proximal (RP11-434D2) SMS-REPs generated an 11.5-kb band. *D,* Identification of a novel junction fragment in patients with dup(17)(p11.2p11.2). DNA samples from three families were digested with *Bam*HI and *Spe*1. A 14.5-kb novel junction fragment was observed only in patients, not in their unaffected family members. Note an ~16-kb band is nonspecific to SMS-REPs.

the recombinant SMS-REPs in patients. DNA sequence analysis predicted a unique *Bam*HI site within the 12.0-kb region of the distal SMS-REP and a *Spe*I site adjacent and centromeric to the 12.0-kb region and within the ~39-kb unique sequence interval specific to the proximal and middle SMS-REPs (between homology blocks C and D) (fig. 3*A*). We observed a predicted 6.9-kb novel junction fragment from the recombinant chromosomes in hybrids 363-1D and 468-5D (fig. 3*B*), confirming the interpretation of the recombinant SMS-REP DNA sequencing studies (fig. 2). As anticipated, these two hybrids had no bands from the distal, middle, or proximal SMS-REPs, because they retained only the recombinant chromosome 17. Hybrids 254-1D, 280-22D, 251-20D, and 244-2D revealed only the 11.5-kb fragment from the distal SMS-REP (fig. 3*B*), indicating that the recombination events occurred telomeric to the *Bam*HI site that is unique to the distal SMS-REP. These data were consistent with our breakpoint results, which were determined by restriction enzyme *cis*-morphism analyses and direct sequencing. Hybrid 255-11D had only the 9.7-kb fragment from the distal SMS-REP (fig. 3*B*), indicating that the strand-

exchange event occurred centromeric to the *Spe*I site unique to the proximal SMS-REP (fig. 1). Further analysis revealed that, although patient 255 is considered to carry a common deletion (because of the presence of the ~1.1-Mb junction fragment in PFGE), the unequal crossover event actually occurred outside of SMS-REP, in an ~33-kb repeat termed "LCR17pF" (Stankiewicz et al. 2003), which was ~130 kb centromeric to the defined distal SMS-REP (data not shown).

By use of genomic Southern analysis of 27 patients who had SMS with common deletions, we identified the junction fragment in 4 patients: 429, 468, 568, and 664 (fig. 3*C*). The combined results from analysis of patient and/or hybrid samples indicated that the strand exchanges resulting in the common deletion occurred within the ~6.9-kb region in ~15% (5/34) of the patients with SMS. The novel junction fragment was specific to the de novo SMS deletion, because it was observed in patients 429 and 468, but not in their parents (fig. 3*C*). The lack of a proximal SMS-REP–specific endonuclease restriction site in the recombinant hotspot telomeric to the distal SMS-REP–specific *Bam*HI site precludes genomic Southern analysis detection of a novel junction fragment for crossovers in this region.

Reciprocal Recombination in SMS-REP Resulting in Duplication

The reciprocal product of the SMS deletion—dup(17)(p11.2p11.2)—was identified recently by FISH and PFGE analyses that showed the presence of an ~1.1-Mb junction fragment (Potocki et al. 2000). If the duplication results from the reciprocal recombination in these patients with dup(17)(p11.2p11.2), and the strand exchange occurs within the 12.0-kb interval in the *KER* gene cluster, a 14.5-kb junction fragment is predicted to be found when genomic DNA from patients with dup(17)(p11.2p11.2) is double digested with *Bam*HI and *Spe*I (fig. 3*A*). This predicted junction fragment was indeed observed in 3 of 13 dup(17)(p11.2p11.2) patients analyzed (fig. 3*D*). In one family, the 14.5-kb junction fragment was present in patient 1364 but absent in his mother and three of his siblings (fig. 3*D*). Although data from his father were not available, marker genotype studies indicated that both of his father's alleles were independently transmitted to the unaffected siblings (Shaw et al. 2002). Moreover, the duplication of 17p11.2-p11.2 in this patient resulted from a maternal intrachromosomal recombination (Shaw et al. 2002). Therefore, the de novo junction fragment was associated with the dup(17)(p11.2p11.2) syndrome in this family. In conclusion, 23% of strand exchanges resulting in reciprocal duplication of chromosome 17 band p11.2 occurred within the same 12.0-kb hotspot region identified in our hybrid analysis of del(17)(p11.2p11.2) chromosomes.

Sequence Analysis of the Crossovers

DNA sequence analysis in the few genomic disorders studied to date suggest that strand exchange occurs preferentially in regions of perfect identity (>300–500-bp) within LCR sequence and that minimal efficient processing segments (MEPS) appear to be required for homologous recombination (Vnencak-Jones and Phillips 1990; Reiter et al. 1998; López-Correa et al. 2001). To examine the correlation between the positional hotspot and the regions of perfect identity, we performed a DNA sequence comparison between the distal and proximal SMS-REPs. As illustrated in figure 4, regions of perfect identity (>200 bp) were present throughout the 170-kb homology. The longest stretch of perfect sequence identity was 754 bp within the *KIAA0565* pseudogene in block A. The 1.1-kb potential hotspot region contained 501 bp of perfect sequence identity (fig. 4). Fourteen fragments of perfect identity sequences that were 200–300 bp and four fragments >300 bp were located in the 12.0-kb hotspot interval (fig. 4). Remarkably, there were no >300-bp perfect stretches in the 1,655-bp region where four breakpoints were mapped (figs. 3 and 4).

The average GC content of both the distal and proximal SMS-REPs is ~45%, whereas that of the 12.0-kb hotspot is ~57%. Similarly, the 2-kb positional hotspot for NF1 has a GC content 12% higher than the average of the NF1REPs (~50%) (López-Correa et al. 2001). However, the GC content of the 1.1-kb hotspot is the same as the average.

A search for sequences stimulating recombination (Badge et al. 2000) within the ~170-kb homology regions between the distal and proximal SMS-REPs identified three human minisatellite core sequences (5′-GGGCAGGARG-3′), one hypervariable minisatellite (5′-GGAGGTGGGCAGGARG-3′), and 11 χ elements from *Escherichia coli* (5′-GCTGGTGG-3′) (fig. 4). All of these elements, except one minisatellite core sequence and two χ sites, were located within block C. One minisatellite core sequence and two χ sites were located inside the 1,655-bp breakpoint (fig. 4). No recombination-prone sequences were identified in or around the 1.1-kb hotspot in block A, except for a single human minisatellite core sequence located 6.6 kb centromeric within block B (fig. 4). The latter sequence was specific to the distal SMS-REP copy.

Inverted Repeats Flank the KER *Gene Clusters in the Proximal SMS-REP*

Analysis of the sequence of SMS-REPs by the program Dotter and by pairwise BLAST showed that a 2.1-kb inverted repeat flanks homology blocks B and C in the proximal and middle SMS-REPs, but not in the distal copy (fig. 5). One repeat was adjacent and telomeric to the block B homology region. The other repeat

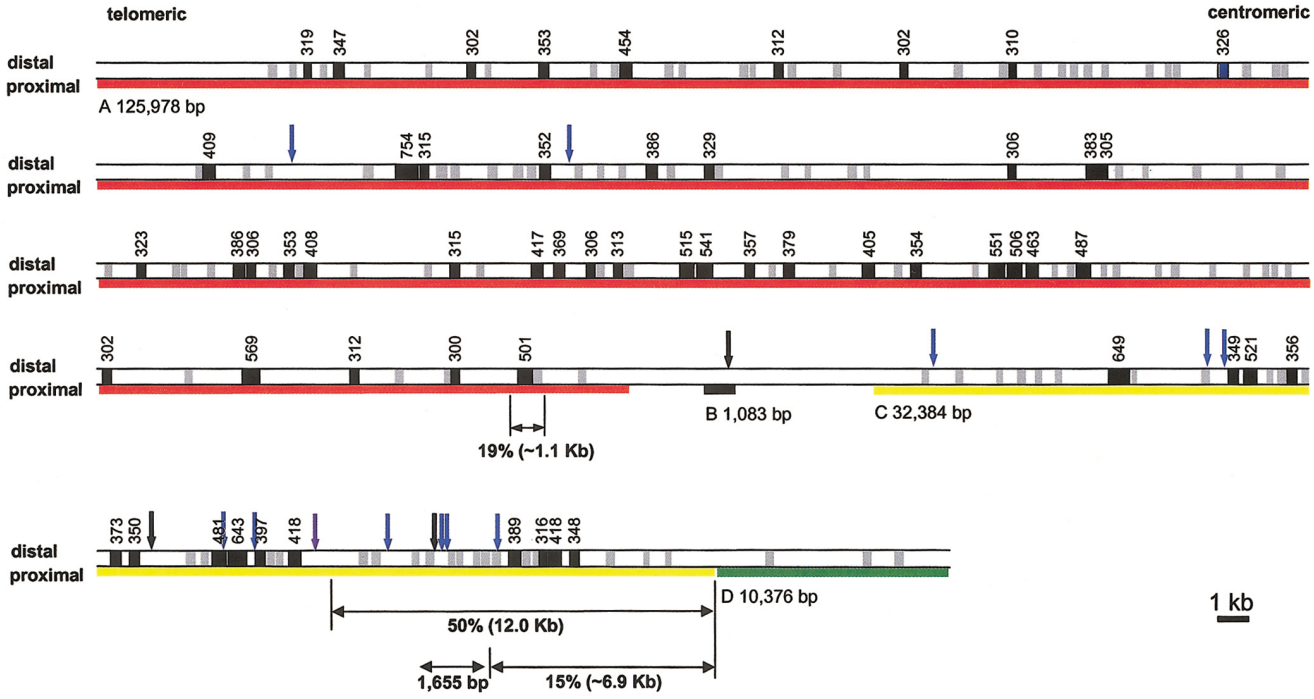

Figure 4 Distribution of sequence blocks with perfect identity between the distal and proximal SMS-REPs. Regions with 200–300 bp perfect identity (*shaded blocks*) and regions with >300 bp perfect identity (*blackened blocks*) are shown with the size of identity in base pairs. The four homology blocks between the distal and proximal SMS-REPs are denoted as blocks A (*red*), B (*black*), C (*yellow*), and D (*green*). The two unequal crossover preferential sites are delineated by horizontal double-end arrows, with the occurrence rate and size indicated. One of them is a 12-kb region at the centromeric end of the homology block C, where 50% (8/16) of strand exchanges occurred. By use of both hybrid analyses and Southern blotting, 15% (5/34) exchanges were mapped in an ~6.9-kb interval. The 1,655-bp region within the 12-kb hotspot is responsible for four strand exchanges among 16 hybrids. Another hotspot is an ~1.1-kb region at the centromeric end of the homology block A that accounts for 19% (3/16) of the unequal crossovers. The locations of χ sites (*blue arrows*), human minisatellite core sequences (*black arrows*), and a human hypervariable minisatellite (*purple arrow*) are indicated.

centromeric to the block C was separated from the homology region, with a 91-bp segment that shares 98% homology with a segment within the inverted repeat and an additional 17 nt, including 4 continuous adenines (fig. 5). The 4 adenines were within a tract of 19 continuous adenines, of which 15 were located within block C, the homologous region between the distal and proximal SMS-REPs. The latter repeat was in greater proximity to the hotspot for strand exchange; the tract of 19 continuous adenines may be more susceptible to double-strand breaks. The inverted repeat was AT rich (37.8% GC content). Within each repeat, there was an internal 646-bp inverted repeat, which, according to the DNA secondary structure prediction program MFold, may enable each repeat to form a hairpin structure.

Discussion

Positional Recombination Hotspots in the SMS-REP

We have identified a 12-kb positional hotspot for strand exchange during NAHR between the proximal and distal SMS-REPs. Through hybrid analysis and identification of novel junction fragments, we demonstrated that the 12.0-kb hotspot within the *KER* gene cluster is a preferential recombination site for the unequal crossover events resulting in both the SMS deletion (50%) and duplication in dup(17)(p11.2p11.2) (>23%) patients. A relative-risk calculation indicates that, for SMS, the strand exchange event occurs in this 12-kb region with a preference of 13:1 over the remaining ~158 kb of homology. Positional recombination hotspots have been identified in DNA rearrangements responsible for each of the genomic disorders studied at the nucleotide sequence level (Reiter et al. 1998; López-Correa et al. 2001; Bayés et al. 2003). Our data show that SMS is another genomic disorder that has a positional recombination hotspot. The existence of a preferential site for NAHR in each of these genomic disorders suggests that there are specific sequences or structural features, in or around the positional hotspot, that promote strand exchanges in this hotspot region. It is interesting that multiple polymorphic variations were documented within a

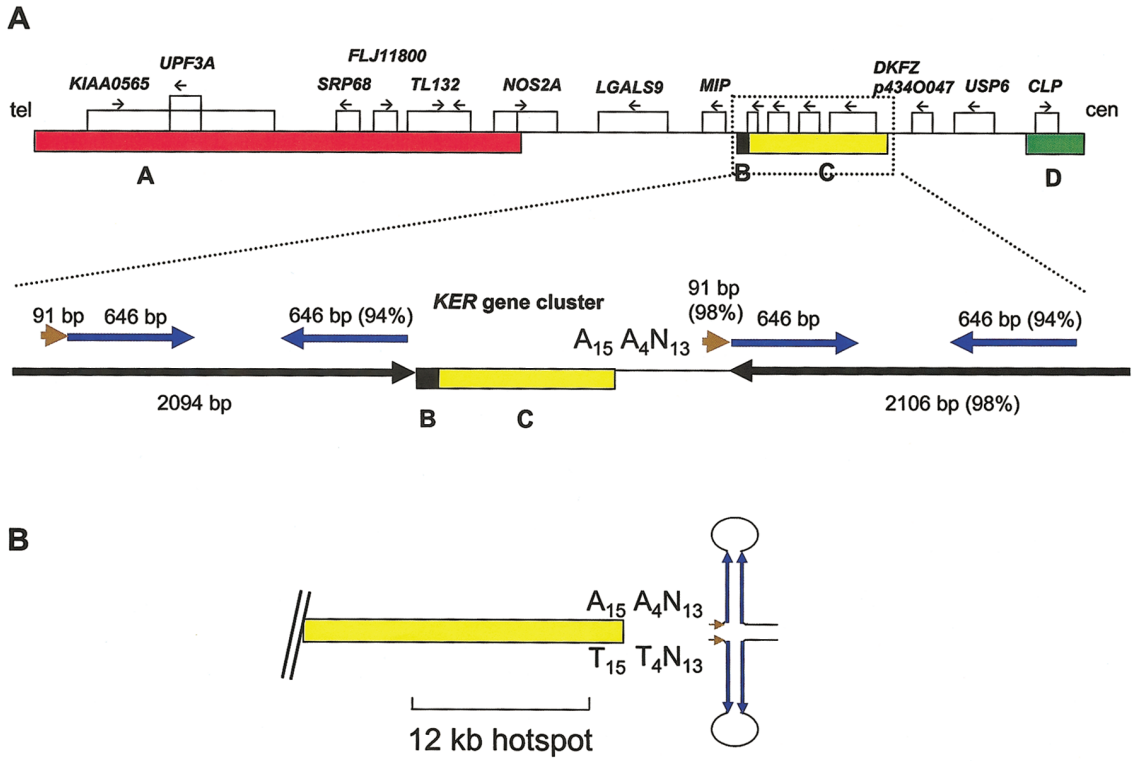

Figure 5 Repeats flanking the *KER* gene cluster. *A*, The 2.1-kb AT-rich (38% GC) inverted repeats (*black arrows*) with 98% homology flank the sequence blocks B and C, including the *KER* gene cluster, in both the proximal and the middle SMS-REPs, but not in the distal SMS-REP. Note that the sizes of the homology blocks and the flanking inverted repeats are not drawn to scale. Within the inverted repeat, there are 646-bp internal inverted repeats with 93% homology (*blue arrows*). One repeat centromeric to the block C is separated from block C with a 91-bp segment (*brown arrows*) that shares 98% homology with a segment within the inverted repeat and 17 additional nucleotides, including 4 adenines (indicated as A_4N_{13}). The 4 adenines are within a tract of 19 continuous adenines, in which 15 are located within block C, the homologous region between the distal and proximal SMS-REPs. *B*, Each of the inverted repeats can theoretically form hairpin structures as indicated. The 12-kb hotspot region is adjacent to the inverted repeat centromeric to the block C. The long adenine tract and the hairpin structures may be more sensitive to DSBs, which might serve as the initiation event for NAHR.

couple of these positions (e.g., 1,655 bp and 2,234 bp) for preference of strand exchanges. This observation suggests the possibility that positional preference for equal exchange (i.e., allelic homologous recombination), as evidenced by gene conversion yielding frequent polymorphism, may overlap with positional preference for unequal strand exchange that accompanies NAHR.

Patients with dup(17)(p11.2p11.2) harbor a reciprocal duplication of the same region deleted in patients with SMS. According to the reciprocal-recombination model, every deletion mediated by NAHR between LCRs has a reciprocal duplication of the same region deleted. One such well-studied event is the unequal crossover between CMT1A-REPs that results in two reciprocal products, the CMT1A duplication and HNPP deletion (Chance et al. 1994). Strand-exchange events in both CMT1A and HNPP preferentially occur in an ~1.7-kb positional recombination hotspot in CMT1A-REP identified through detection of novel junction fragments (Reiter et al. 1996). Similarly, our results indicated that the deletion in patients with SMS and the duplication in patients with dup(17)(p11.2p11.2) shared the same 12.0-kb positional hotspot in the SMS-REP. Thus, it is highly likely that all microdeletion syndromes that result from NAHR between flanking LCRs have reciprocal microduplication syndromes, given the reciprocity of the strand exchange that results in both deletion and duplication rearrangements. Correlating a specific phenotype with a predicted microduplication remains a challenge for human and medical geneticists.

In addition to the 12.0-kb positional hotspot, we identified a 1.1-kb region that accounts for ~19% (3/16) of unequal crossover events that result in SMS deletion. It is of interest that both the 12.0-kb hotspot and the 1.1-kb hotspot were adjacent to the boundary between homology and nonhomology sequences. This finding suggests that the homology sequences adjacent to nonhomology regions may be more vulnerable to either DSBs or

resolution of Holliday junctions than is the sequence embedded inside the homology sequences.

Requirements for Efficient Homologous Recombination

The factors that promote the preferential occurrence of unequal crossover events in the positional recombination hotspots during NAHR are still unknown. Direct sequencing of recombination junction fragments suggests that a certain length of perfect identity (300–500 bp) or a MEPS in humans may be required for an efficient NAHR (Reiter et al. 1998; López-Correa et al. 2001). Our sequencing of the junctions in the recombinant SMS-REPs supports this observation. There is a 501-bp stretch of perfect identity present in the 1.1-kb hotspot and four stretches of identity >300 bp in the 12-kb hotspot. A 404-bp stretch of perfect identity was present in the 1,655-bp region that was responsible for four strand-exchange events among 16 hybrids (fig. 2). Moreover, polymorphic nucleotides were abundant in this region (fig. 2), which may provide better sequence identity between misaligned SMS-REPs.

Recombination-promoting sequences have been shown to stimulate homologous recombination in different species. Human minisatellite core sequences stimulate homologous recombination in human cells (Jeffreys et al. 1985; Wahls et al. 1990). χ elements stimulate recombination in *E. coli* (Smith 1997). Five human minisatellite-like sequence clusters and one χ element have been identified within the ~741-bp hotspot in CMT1A-REP (Lopes et al. 1998). In addition, one χ element was also identified in the NF1 recombination hotspot (López-Correa et al. 2001). Within the 12-kb hotspot, in both distal and proximal SMS-REPs, we identified a human minisatellite core sequence and four χ elements. One minisatellite core sequence was located ~6.6 kb from the 1.1-kb hotspot. However, the involvement of any potential *cis*-acting, recombination-promoting elements in NAHR-mediated genomic disorders remains to be experimentally demonstrated.

Complex Genome Architecture May Stimulate Recombination

LCRs play an important role in human genomic disorders. The majority of human genomic disorders occur in LCR-rich regions; these paralogous regions serve as substrates for NAHR (Stankiewicz and Lupski 2002*a*). Some DNA rearrangements are mediated by nonhomologous recombination initiated by DNA breakage and followed by reunion of DNA segments (Roth et al. 1985; Karran 2000). Several lines of evidence suggest that LCRs may also be involved in nonhomologous recombination. Recently, inverted LCRs were identified adjacent to the region in which nonhomologous recombination occurred. In two families with Pelizaeus-Merzbacher disease, the *PLP1* gene was deleted, and no homologous sequences flanked the breakpoints (Inoue et al. 2002). Close to the *PLP1* gene, two LCRs were identified that share 99.3% homology for >20 kb. The breakpoints for isodicentric chromosome idic(17)(p11) found in malignancies (e.g., leukemia and medulloblastoma) were mapped within the SMS region to a single BAC that consists entirely of a complex LCR with large palindromes (Gao et al. 1997; Fioretos et al. 1999; Barbouti et al. in press).

Long inverted repeats (LIR) in which the repeat unit is >100 bp promote genetic instability by forming hairpin structures and stimulating deletions and recombination in yeast and mammals (Gordenin and Resnick 1998). Two 1.0-kb inverted repeats stimulated interchromosomal homologous recombination by 17,000-fold in yeast (Lobachev et al. 1998). An LIR composed of two 1.1-kb *URA3* sequences separated by a 200-bp spacer sequence stimulated a ~5-fold increase in intrachromosomal homologous recombination in mammalian cells (Waldman et al. 1999). Breakpoints of deletions stimulated by LIRs were within or adjacent to the LIRs (Gordenin et al. 1993). In addition, palindromic AT-rich repeats have been suggested to promote DSB that led to nonhomologous recombination events associated with genomic disorders such as the constitutional t(11;22) translocation (Edelmann et al. 2001; Kurahashi and Emanuel 2001).

We identified a 2.1-kb inverted repeat, flanking the ~34-kb *KER* gene clusters in which the 12-kb hotspot was located. The inverted repeat was located in the proximal and middle SMS-REPs, but not in the distal SMS-REP and may be involved in the evolution of blocks B and C in the SMS-REPs from their progenitor (Park et al. 2002). It is of interest that internal 646-bp repeats were present in the 2.1-kb inverted repeats, suggesting that the 2.1-kb repeat may be able to form a hairpin structure like an LIR. One copy of the 2.1-kb inverted repeats was adjacent and centromeric to the 12-kb hotspot that contains several stretches of >300 bp perfect identity. We speculate that the hairpin structure formed by the inverted repeat and its adjacent long adenine tract are more sensitive to DSBs, which might serve as the initiation event for homologous recombination. The homologous sequences close to the inverted repeat enable pairing between either allelic or nonallelic SMS-REP copies. When strand-exchange events occur between the latter, the resultant products are deletion/duplication rearrangements. Alternatively, positional preference for strand exchange during homologous recombination may reflect chromosome environment or other *trans*-acting factors that remain to be identified.

In summary, we have identified a 12-kb positional recombination hotspot in SMS-REP, which accounts for half of the strand exchange events that occurred in the

~170-kb homology between SMS-REPs. This positional preference for crossovers in SMS-REP is associated with events that lead to either del(17)(p11.2p11.2) or dup(17)(p11.2p11.2). The stimulation of crossover events in the hotspot by the inverted repeats flanking the preferential region for strand exchange needs to be further investigated.

Acknowledgments

We thank Drs. K. Inoue and P. Stankiewicz, for their critical reviews, and J. Cook, for the excellent technical assistance. S.-S. P. was supported by a fellowship from the South Korean government. This research was supported, in part, by the Muscular Dystrophy Association and National Institute of Child Health and Human Development grants P01 HD38420, HD 2406407 (to Baylor College of Medicine Mental Retardation Research Center), and M01 RR00188 (to the Texas Children's Hospital General Clinical Research Center).

Electronic-Database Information

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

BLAST 2 Sequences, http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html (for pairwise sequence comparison)

Dotter program, http://www.cgr.ki.se/cgr/groups/sonnhammer/Dotter.html (for sequence comparison)

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for the distal SMS-REP [accession number BK001589], the middle SMS-REP [accession number BK001590], and the proximal SMS-REP [accession number BK001591])

MFold program, http://www.biology.wustl.edu/gcg/mfold.html (for DNA secondary structure prediction)

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for SMS)

Primer 3, http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi (for primer design)

References

Badge RM, Yardley J, Jeffreys AJ, Armour JA (2000) Crossover breakpoint mapping identifies a subtelomeric hotspot for male meiotic recombination. Hum Mol Genet 9:1239–1244

Barbouti A, Stankiewicz P, Birren B, Nusbaum C, Cuomo C, Cook A, Höglund M, Johansson B, Hagemeijer A, Park S-S, Mitelman F, Lupski JR, Fioretos T. The breakpoint region of the most common isochromosome, i(17q), in human neoplasia is characterized by a complex genomic architecture with large palindromic low-copy repeats. Am J Hum Genet (in press)

Bayés M, Magano LF, Rivera N, Flores R, Pérez Jurado LA (2003) Mutational mechanisms of Williams-Beuren syndrome deletions. Am J Hum Genet 73:131–151

Bi W, Yan J, Stankiewicz P, Park S-S, Walz K, Boerkoel CF, Potocki L, Shaffer LG, Devriendt K, Nowaczyk MJM, Inoue K, Lupski JR (2002) Genes in a refined Smith-Magenis syndrome critical deletion interval on chromosome 17p11.2 and the syntenic region of the mouse. Genome Res 12:713–728

Chance PF, Abbas N, Lensch MW, Pentao L, Roa BB, Patel PI, Lupski JR (1994) Two autosomal dominant neuropathies result from reciprocal DNA duplication/deletion of a region on chromosome 17. Hum Mol Genet 3:223–228

Chen K-S, Manian P, Koeuth T, Potocki L, Zhao Q, Chinault AC, Lee CC, Lupski JR (1997) Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. Nat Genet 17:154–163

Edelmann L, Pandita RK, Morrow BE (1999) Low-copy repeats mediate the common 3-Mb deletion in patients with velo-cardio-facial syndrome. Am J Hum Genet 64:1076–1086

Edelmann L, Spiteri E, Koren K, Pulijaal V, Bialer MG, Shanske A, Goldberg R, Morrow BE (2001) AT-rich palindromes mediate the constitutional t(11;22) translocation. Am J Hum Genet 68:1–13

Fioretos T, Strömbeck B, Sandberg T, Johansson B, Billström R, Borg A, Nilsson P-G, Van Den Berghe H, Hagemeijer A, Mitelman F, Höglund M (1999) Isochromosome 17q in blast crisis of chronic myeloid leukemia and in other hematologic malignancies is the result of clustered breakpoints in 17p11 and is not associated with coding *TP53* mutations. Blood 94:225–232

Gao L, Frey MR, Matera AG (1997) Human genes encoding U3 snRNA associate with coiled bodies in interphase cells and are clustered on chromosome 17p11.2 in a complex inverted repeat structure. Nucleic Acids Res 25:4740–4747

Gordenin DA, Lobachev KS, Degtyareva NP, Malkova AL, Perkins E, Resnick MA (1993) Inverted DNA repeats: a source of eukaryotic genomic instability. Mol Cell Biol 13:5315–5322

Gordenin DA, Resnick MA (1998) Yeast ARMs (DNA at-risk motifs) can reveal sources of genome instability. Mutat Res 400:45–58

Greenberg F, Guzzetta V, Montes de Oca-Luna R, Magenis RE, Smith ACM, Richter SF, Kondo I, Dobyns WB, Patel PI, Lupski JR (1991) Molecular analysis of the Smith-Magenis syndrome: a possible contiguous-gene syndrome associated with del(17)(p11.2). Am J Hum Genet 49:1207–1218

Guzzetta V, Franco B, Trask BJ, Zhang H, Saucedo-Cardenas O, Montes de Oca-Luna R, Greenberg F, Chinault AC, Lupski JR, Patel PI (1992) Somatic cell hybrids, sequence-tagged sites, simple repeat polymorphisms, and yeast artificial chromosomes for physical and genetic mapping of proximal 17p. Genomics 13:551–559

Han L-L, Keller MP, Navidi W, Chance PF, Arnheim N (2000) Unequal exchange at the Charcot-Marie-Tooth disease type 1A recombination hot-spot is not elevated above the genome average rate. Hum Mol Genet 9:1881–1889

Inoue K, Dewar K, Katsanis N, Reiter LT, Lander ES, Devon KL, Wyman DW, Lupski JR, Birren B (2001) The 1.4-Mb CMT1A duplication/HNPP deletion genomic region reveals unique genome architectural features and provides insights into the recent evolution of new genes. Genome Res 11:1018–1033

Inoue K, Lupski JR (2002) Molecular mechanisms for genomic disorders. Annu Rev Genomics Hum Genet 3:199–242

Inoue K, Osaka H, Thurston VC, Clarke JTR, 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

Jeffreys AJ, Wilson V, Thein SL (1985) Hypervariable "minisatellite" regions in human DNA. Nature 314:67–73

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

Juyal RC, Figuera LE, Hauge X, Elsea SH, Lupski JR, Greenberg F, Baldini A, Patel PI (1996*a*) Molecular analyses of 17p11.2 deletions in 62 Smith-Magenis syndrome patients. Am J Hum Genet 58:998–1007

3.3.CO;2-9Juyal RC, Kuwano A, Kondo I, Zara F, Baldini A, Patel PI (1996*b*) Mosaicism for del(17)(p11.2p11.2) underlying the Smith-Magenis syndrome. Am J Med Genet 66: 193–196

Karran P (2000) DNA double strand break repair in mammalian cells. Curr Opin Genet Dev 10:144–150

Kiyosawa H, Lensch MW, Chance PF (1995) Analysis of the CMT1A-REP repeat: mapping crossover breakpoints in CMT1A and HNPP. Hum Mol Genet 4:2327–2334

Kurahashi H, Emanuel BS (2001) Long AT-rich palindromes and the constitutional t(11;22) breakpoint. Hum Mol Genet 10:2605–2617

Lobachev KS, Shor BM, Tran HT, Taylor W, Keen JD, Resnick MA, Gordenin DA (1998) Factors affecting inverted repeat stimulation of recombination and deletion in *Saccharomyces cerevisiae*. Genetics 148:1507–1524

Lopes J, LeGuern E, Gouider R, Tardieu S, Abbas N, Birouk N, Gugenheim M, Bouche P, Agid Y, Brice A, French CMT Collaborative Research Group (1996) Recombination hot spot in a 3.2-kb region of the Charcot-Marie-Tooth type 1A repeat sequences: new tools for molecular diagnosis of hereditary neuropathy with liability to pressure palsies and of Charcot-Marie-Tooth type 1A. Am J Hum Genet 58:1223–1230

López-Correa C, Dorschner M, Brems H, Lázaro C, Clementi M, Upadhyaya M, Dooijes D, Moog U, Kehrer-Sawatzki H, Rutkowski JL, Fryns J-P, Marynen P, Stephens K, Legius E (2001) Recombination hotspot in *NF1* microdeletion patients. Hum Mol Genet 10:1387–1392

Lupski JR (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. Trends Genet 14:417–422

——— (2003) Genomic disorders: recombination-based disease resulting from genomic architecture. 2002 Curt Stern Award Address. Am J Hum Genet 72:246–252

Park S-S, Stankiewicz P, Bi W, Shaw C, Lehoczky J, Dewar K, Birren B, Lupski JR (2002) Structure and evolution of the Smith-Magenis syndrome repeat gene clusters, SMS-REPs. Genome Res 12:729–738

Pentao L, Wise CA, Chinault AC, Patel PI, Lupski JR (1992) Charcot-Marie-Tooth type 1A duplication appears to arise from recombination at repeat sequences flanking the 1.5 Mb monomer unit. Nat Genet 2:292–300

Peoples R, Franke Y, Wang Y-K, Pérez-Jurado L, Paperna T, Cisco M, Francke U (2000) A physical map, including a BAC/PAC clone contig, of the Williams-Beuren syndrome–deletion region at 7q11.23. Am J Hum Genet 66:47–68

Potocki L, Chen K-S, Park S-S, Osterholm DE, Withers MA, Kimonis V, Summers AM, Meschino WS, Anyane-Yeboa K, Kashork CD, Shaffer LG, Lupski JR (2000) Molecular mechanism for duplication 17p11.2: the homologous recombination reciprocal of the Smith-Magenis microdeletion. Nat Genet 24:84–87

Potocki L, Shaw CJ, Stankiewicz P, Lupski JR (2003) Variability in clinical phenotype despite common chromosomal deletion in Smith-Magenis syndrome [del(17)(p11.2p11.2)]. Genet Med 5:430–434

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

Reiter LT, Murakami T, Koeuth T, Pentao L, Muzny DM, Gibbs RA, Lupski JR (1996) A recombination hotspot responsible for two inherited peripheral neuropathies is located near a *mariner* transposon-like element. Nat Genet 12:288–297

Roth DB, Porter TN, Wilson JH (1985) Mechanisms of nonhomologous recombination in mammalian cells. Mol Cell Biol 5:2599–2607

Shaikh TH, Kurahashi H, Saitta SC, O'Hare AM, Hu P, Roe BA, Driscoll DA, McDonald-McGinn DM, Zackai EH, Budarf ML, Emanuel BS (2000) Chromosome 22–specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. Hum Mol Genet 9:489–501

Shaw CJ, Bi W, Lupski JR (2002) Genetic proof of unequal meiotic crossovers in reciprocal deletion and duplication of 17p11.2. Am J Hum Genet 71:1072–1081

Slager RE, Newton TL, Vlangos CN, Finucane B, Elsea SH (2003) Mutations in *RAI1* associated with Smith-Magenis syndrome. Nat Genet 33:466–468

Smith GR (1997) Chi sites and their consequences. In: de Bruijn FJ, Lupski JR, Weinstock GM (eds) Bacterial genomes: physical structure and analysis. International Thomson Publishing, New York, pp 49–66

Stankiewicz P, Lupski JR (2002*a*) Genome architecture, rearrangements and genomic disorders. Trends Genet 18:74–82

——— (2002*b*) Molecular-evolutionary mechanisms for genomic disorders. Curr Opin Genet Dev 12:312–319

Stankiewicz P, Shaw CJ, Dapper JD, Wakui K, Shaffer LG, Withers M, Elizondo L, Park S-S, Lupski JR (2003) Genome architecture catalyzes nonrecurrent chromosomal rearrangements. Am J Hum Genet 72:1101–1116

Timmerman V, Rautenstrauss B, Reiter LT, Koeuth T, Lofgren A, Liehr T, Nelis E, Bathke KD, De Jonghe P, Grehl H, Martin J-J, Lupski JR, Van Broeckhoven C (1997) Detection of the CMT1A/HNPP recombination hotspot in unrelated patients of European descent. J Med Genet 34:43–49

Trask BJ, Mefford H, van den Engh G, Massa HF, Juyal RC, Potocki L, Finucane B, Abuelo DN, Witt DR, Magenis E, Baldini A, Greenberg F, Lupski JR, Patel PI (1996) Quan-

tification by flow cytometry of chromosome-17 deletions in Smith-Magenis syndrome patients. Hum Genet 98:710–718

Valero MC, de Luis O, Cruces J, Pérez Jurado LA (2000) Fine-scale comparative mapping of the human 7q11.23 region and the orthologous region on mouse chromosome 5G: the low-copy repeats that flank the Williams-Beuren syndrome deletion arose at breakpoint sites of an evolutionary inversion(s). Genomics 69:1–13

Vnencak-Jones CL, Phillips JA III (1990) Hot spots for growth hormone gene deletions in homologous regions outside of Alu repeats. Science 250:1745–1748

Wahls WP, Wallace LJ, Moore PD (1990) Hypervariable minisatellite DNA is a hotspot for homologous recombination in human cells. Cell 60:95–103

Waldman AS, Tran H, Goldsmith EC, Resnick MA (1999) Long inverted repeats are an at-risk motif for recombination in mammalian cells. Genetics 153:1873–1883

Yamamoto M, Yasuda T, Hayasaka K, Ohnishi A, Yoshikawa H, Yanagihara T, Ikegami T, Yamamoto T, Ohashi H, Nishimura T, Mitsuma T, Kiyosawa H, Chance PF, Sobue G (1997) Locations of crossover breakpoints within the CMT1A-REP repeat in Japanese patients with CMT1A and HNPP. Hum Genet 99:151–154

Zori RT, Lupski JR, Heju Z, Greenberg F, Killian JM, Gray BA, Driscoll DJ, Patel PI, Zackowski JL (1993) Clinical, cytogenetic, and molecular evidence for an infant with Smith-Magenis syndrome born from a mother having a mosaic 17p11.2p12 deletion. Am J Med Genet 47:504–511