Human Molecular Genetics, 2004, Vol. 13, Review Issue 1 **R57–R64**
DOI: 10.1093/hmg/ddh073
Advance Access published on February 5, 2004

Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease

Christine J. Shaw[1] and James R. Lupski[1,2,3,*]

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

Received December 19, 2003; Revised January 14, 2004; Accepted January 26, 2004

The term 'genomic disorder' refers to a disease that is caused by an alteration of the genome that results in complete loss, gain or disruption of the structural integrity of a dosage sensitive gene(s). In most of the common chromosome deletion/duplication syndromes, the rearranged genomic segments are flanked by large (usually >10 kb), highly homologous low copy repeat (LCR) structures that can act as recombination substrates. Recombination between non-allelic LCR copies, also known as non-allelic homologous recombination, can result in deletion or duplication of the intervening segment. Recent findings suggest that other chromosomal rearrangements, including reciprocal, Robertsonian and jumping translocations, inversions, isochromosomes and small marker chromosomes, may also involve susceptibility to rearrangement related to genome structure or architecture. In several cases, LCRs, AT-rich palindromes and pericentromeric repeats are located at such rearrangement breakpoints. Analysis of the products of recombination at the junctions of the rearrangements reveals both homologous recombination and non-homologous end joining as causative mechanisms. Thus, a more global concept of genomic disorders emerges in which susceptibility to rearrangements occurs due to underlying complex genomic architecture. Interestingly, this architecture plays a role not only in disease etiology, but also in primate genome evolution. In this review, we discuss recent advances regarding general mechanisms for the various rearrangements of our genome, and potential models for rearrangements with non-homologous breakpoint regions.

INTRODUCTION

Genomic disorders previously have been defined as disorders in which the clinical phenotype is a consequence of abnormal dosage of a gene(s) located within a rearranged segment of the genome (1–3). This group of disorders is distinguished from conventional Mendelian disease in that the phenotype does not result from a point mutation, but rather from larger alterations of the genome. These alterations include deletions, duplications, inversions and translocations. Such rearrangements occur via recombination mechanisms whereas point mutations usually result from DNA replication or repair errors. The number of recognized genomic disorders continues to expand, with the recent additions of Sotos syndrome (SoS), split hand–split foot malformation 3 (SHFM3), and Kabuki syndrome (KS) (4–6).

Chromosome rearrangement breakpoints have been located throughout the genome; however, they predominate in the pericentromeric and subtelomeric regions, particularly in intervals containing complex genomic architecture, such as low-copy repeats (LCRs) or AT-rich palindromes. Non-allelic homologous recombination (NAHR) is usually the mechanism responsible for rearrangements with breakpoints clustering in LCRs. Other mechanisms such as non-homologous end joining (NHEJ) have been observed (7), particularly for rearrangements with scattered breakpoints (Fig. 1). Nevertheless, regardless of recombination mechanism, genomic architectural features have been associated with many rearrangement breakpoints. This suggests that chromosomal rearrangements are not random events, but result from predisposition to rearrangement due to the existence of complex genomic architecture that may create instability in the genome.

*To whom correspondence should be addressed at: One Baylor Plaza, Room 604B, Houston, TX 77030, USA. Tel: +1 7137986530; Fax: +1 7137985073; Email: jlupski@bcm.tmc.edu

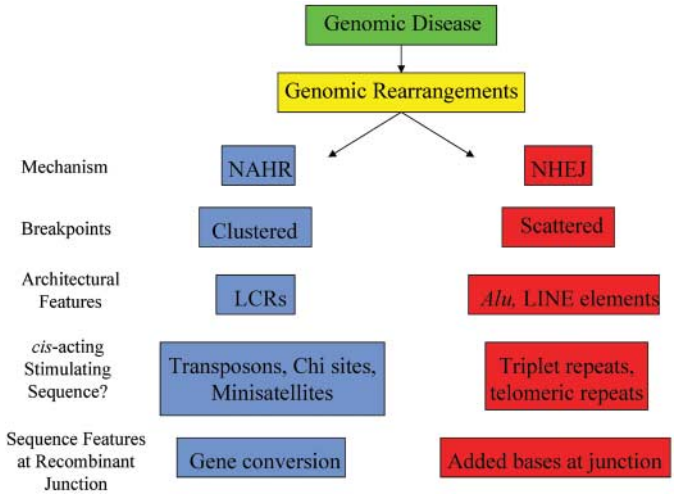

Figure 1. Mechanisms of genomic rearrangements. Two primary recombination mechanisms, NAHR (blue) and NHEJ (red), are shown. Features associated with NAHR or NHEJ are shown in blue and red, respectively.

RECURRENT REARRANGEMENTS RESULTING FROM LCR-MEDIATED NAHR

NAHR is the most common mechanism underlying disease-associated genome rearrangements. LCRs, usually on the same but sometimes on different chromosomes, can act as substrates for NAHR. NAHR between LCRs in direct orientation on the same chromosome results in reciprocal deletions and duplications, whereas NAHR between LCRs in inverted orientation on the same chromosome results in inversions. NAHR also can occur between LCRs located on different chromosomes, resulting in reciprocal translocations.

LCRs are usually 10–500 kb in size and >95% identical (2). NAHR between LCRs results in a clustering of rearrangement breakpoints within the LCRs, allowing detection of a rearrangement-specific common junction fragment by pulsed-field gel electrophoresis (PFGE) analysis. These junction fragments are key to narrowing the strand exchange interval and uncovering the precise recombination mechanism in recurrent rearrangements (8–12). Previous PFGE and sequencing studies on Charcot–Marie–Tooth disease type 1A (CMT1A) and hereditary neuropathy with liability to pressure palsies (HNPP) revealed a 557 bp recombination hotspot within 24 kb LCRs (CMT1A-REPs) in patients with either the CMT1A duplication or the HNPP deletion (13–15). Evidence for gene conversion between the CMT1A-REPs was observed, and a *mariner*-like transposable element was identified near the hotspot, along with evidence for a double strand break (DSB) mechanism (13–15). A 2 kb hotspot containing a chi-like sequence also was identified within the neurofibromatosis type 1 LCRs (NF1-REPs), along with evidence for gene conversion (16). A model in which *cis*-acting sequences stimulate increased potential for double strand breaks was proposed as the etiology of the observed preference for strand exchange within the LCRs, and prompted further studies of the crossover sites in other NAHR-mediated rearrangements (13–16).

As was demonstrated with CMT1A/HNPP and NF1, recent work on Williams–Beuren (WBS), Smith-Magenis (SMS), and dup(17)(p11.2p11.2) syndromes, and Y chromosome deletions associated with azoospermia and male infertility provide further evidence of a positional preference for strand exchange within LCRs, despite several hundred kilobases of highly homologous (>98% identical) sequence (17–19).

In a study of 30 WBS patients with a common 1.55 Mb deletion of 7q11.23 between centromeric and medial WBS LCRs (each composed of blocks A, B and C), breakpoints were found to cluster in block B (~143 kb), which has the highest sequence identity at 99.6% (17). Microsatellite analysis of recombinant B blocks in 19/30 WBS patients revealed that 7–12/19 (37–63%) recombinations occurred in a 12 kb region within the *GTF2I/GTF2IP1* gene, representing only 11.4% of the total sequence of block B (17). Interestingly, it was found that 11/30 (37%) of the WBS patients studied harbored an inversion between B blocks of the medial and telomeric LCRs, which are inverted with respect to one another. Additionally, in a larger sample, 21/74 (28%) of the transmitting progenitors were heterozygous for an inversion between centromeric and telomeric LCRs. Sequence analysis of block B revealed the total percentage of repetitive elements to encompass 49.7% of the block, which is significantly higher than the average 34% predicted for DNA with similar GC content (20).

Analysis of large deletions of the Y chromosome [including azoospermia factor b (*AZFb*) and *AZFc* loci] associated with spermatogenic failure has shown that large palindromes on Yq (named P1–P5) serve as substrates for NAHR (19,21). In a study of 11 Yq deletions, 10/11 (91%) of proximal breakpoints clustered within 30 kb of the center of P5, and 11/11 distal breakpoints clustered within 25 kb of either of two mini-palindromes within P1 (19). Four deletions were found to be a result of NAHR between two copies of a 933 bp sequence located within the palindromes (19). An additional deletion occurred via NAHR between a second set of the 933 bp sequences, also located within the palindromes. Although 7/9 (78%) of the deletions for which junctions were sequenced were due to NAHR, two of the deletions had no homologous sequence at the junction, despite three of the breakpoints mapping within the proximal and distal recombination hotspots (19). This suggests that a non-homologous recombination mechanism stimulated by the palindromic structure may be responsible for generation of these latter deletions.

In the case of SMS/dup(17)(p11.2p11.2), the same positional preference was identified for the strand exchanges resulting in either deletion or duplication (18), demonstrating, as had been done for HNPP/CMT1A (14,15), the reciprocity of the crossover event. A study on patients with the common SMS deletion or dup(17)(p11.2p11.2) revealed clustering of breakpoints within the *KER* (keratin) gene cluster of the proximal and distal SMS-REPs (18). Analysis of 16 somatic cell hybrids showed that 50% of the recombinant junctions occurred in a 12 kb region within the *KER* gene clusters, despite 170 kb of high similarity (>98% identity) between the proximal and distal SMS-REP copies. Sequencing of this hotspot in seven of the recombinant SMS-REPs further narrowed the crossovers to an 8 kb interval. Four of the seven breakpoints occurred in a 1688 bp region rich in polymorphic nucleotides, potentially reflecting frequent gene conversion. Genomic Southern analysis of 27 SMS patients revealed a junction fragment in four additional cases, corresponding to crossovers in a 6.9 kb region of the 12 kb hotspot, totaling 5/34 (15%) of SMS

patients with crossovers in this interval (18). Patients with the common reciprocal duplication were also analyzed by Southern analysis, and 3/13 (23%) of the cases studied had strand exchanges that occurred in the 12 kb hotspot within the *KER* gene cluster, documenting reciprocity at the strand exchange level. Sequence analysis of the SMS-REPs identified an AT-rich 2.1 kb inverted repeat near the 12 kb hotspot, which could mediate a hairpin loop formation, potentially predisposing to DSBs (18).

LCRs and palindromes are also implicated in recurrent somatic rearrangements, such as idic(17q), the most common chromosomal rearrangement observed in neoplasia (22). The breakpoint of the idic(17q) chromosome is located in a complex LCR consisting of five segments of ~40 kb each in 17p11.2, two of which are located in a palindromic structure (22). There is also evidence for potential involvement of LCRs in the genesis of the Philadelphia chromosome t(9;22) (23).

INVERSION POLYMORPHISMS MAY PREDISPOSE TO FUTURE REARRANGEMENTS

Several inversion polymorphisms have been identified in association with genomic disorders. The inversions occur via NAHR using LCRs that are positioned in the genome in an inverted orientation, as substrates for recombination. In addition to the inversion associated with WBS, 4/6 (67%) of mothers of Angelman syndrome (AS) patients with class II (BP2/3) deletions and 4/44 (9%) of control subjects were found to carry a heterozygous inversion of the same region deleted in AS patients (24). Likewise, inversions between olfactory receptor–gene clusters in 4p16 and 8p23 recently have been shown to mediate the recurrent t(4;8)(p16;p23) (25). In this latter case, inversions of both 4p16 and 8p23 were identified in 5/5 mothers of translocation carriers, as well as in 2.5% of control subjects. Inversions of 4p16 and 8p23 were detected in 12.5 and 26% of control subjects, respectively (25). In patients with KS, a BAC probe located just distal to the duplicated segment within 8p23 showed an inverted signal in 6/6 of KS patients and in 2/2 of the KS patients' mothers (6). In a control population, 1/20 (5%) of individuals carried a larger inversion of 8p22–8p23.1, that contained the BAC inverted in KS patients (6).

The presence of an inversion between LCR copies may stimulate aberrant recombination between chromosomes or chromatids, resulting in the aforementioned deletions, duplications and translocations. Given the prevalence of such inversions in the normal population, it appears that a minority of individuals may be at a greater risk of having children with genomic disorders.

DERIVATION OF LCRS, THE SUBSTRATES FOR NAHR

The implication of LCRs in disease-associated rearrangements is growing, as novel LCRs are identified at breakpoint sites throughout the genome (4,26,27). LCRs (also known as segmental duplications or duplicons) result from segmental duplications of the genome and may represent genes, pseudogenes, gene fragments, repeat gene clusters and other chromosomal segments. The genome-wide frequency of LCRs (>1 kb; >95% identity) has been estimated, by computational analysis, at 5–10% (28). However, they are unevenly distributed, with clustering in particular regions of the genome, such as pericentromeric and subtelomeric areas. Recent analysis of proximal 17p revealed that LCRs constitute >23% of the 7.5 Mb genome sequence analyzed in that interval (27). Interestingly, several of the LCRs went unidentified until patient deletion breakpoints were mapped by FISH, and the sequences at the breakpoints analyzed. This suggests that additional as-yet-unidentified LCRs may exist throughout the genome, and may only be revealed through focused studies of the sequence surrounding rearrangement breakpoints.

The generation and structure of LCRs appears to be associated with *Alu* elements. *Alu* sequences have been identified at the junctions of genes/pseudogenes within LCRs on 22q11 (29). An additional study of the junctions of segmental duplications across the human genome revealed a highly significant ($P < 0.0001$) enrichment of *Alu* sequences near or within the junctions (30). Intriguingly, the *Alu* elements at the junctions showed higher levels of divergence, consistent with *Alu–Alu*-mediated recombination. This *Alu* enrichment was due exclusively to the younger subfamilies *Alu*Y and *Alu*S, whereas the oldest primate subfamily, *Alu*J, showed no enrichment (30). This discovery lead to the proposal that the primate-specific burst of *Alu* retroposon activity (35–40 million years ago) sensitized the ancestral genome for *Alu–Alu*-mediated recombination events, that, in turn, might have initiated expansion of gene-rich segmental duplications and their subsequent role in NAHR (30). LCRs have >95% sequence identity, suggesting they have evolved over the last 35 million years, consistent with the high level of *Alu* enrichment occurring at the same time (31). Interestingly, essentially all of the LCRs involved in genomic disorders, that have been examined to date, have evolved as segmental duplications during primate speciation (32).

NON-RECURRENT REARRANGEMENTS ASSOCIATED WITH OTHER GENOME ARCHITECTURAL FEATURES

In addition to recurrent rearrangements, non-recurrent rearrangement breakpoints are also associated with LCRs. A study of unusually sized interstitial deletions and reciprocal translocations involving proximal 17p showed that 21/33 (64%) of deletion breakpoints within 17p11.2 occurred in LCRs, whereas only 1/8 (13%) of translocation breakpoints were within an LCR (27). However, 5/8 (63%) of translocation breakpoints in this region occurred either within or immediately adjacent to the centromere. Interestingly, 4/8 (50%) of partner chromosome breakpoints mapped within the most telomeric sub-bands (27). Recently, a constitutional jumping translocation between donor chromosome 21q21.3-qter and recipients 13qter and 18qter was reported in which a novel 550 kb complex LCR flanked the 21q breakpoint (33).

Translocation breakpoints also cluster in the LCR22s in the DiGeorge/Velocardiofacial syndrome (DGS/VCFS) region in 22q11.2 (34,35). The recurrent t(11;22)(q23;q11) breakpoint is

mediated by double strand breaks in AT-rich palindromes on both chromosomes 11 and 22 (34,36,37). In addition, Spiteri *et al.* (35) mapped 8/14 (57%) of non-recurrent translocation breakpoints involving 22q11.2 within LCR22s. All 14 partner chromosome breakpoints were located in the telomeric bands (35). Additional t(17;22), t(4;22), and t(1;22) breakpoints were also mapped within an LCR22, and palindromic AT-rich repeats (PATRRs) were found at the breakpoints on the derivative chromosomes, suggesting a stem–loop structure formation (38–40).

The breakpoints of the most common constitutional recurrent marker chromosomes, deriving from chromosomes 15 [inv dup (15)] and 22 [inv dup (22)/cat eye syndrome], are also associated with LCRs and sometimes the centromere (41–43). Non-recurrent marker chromosome breakpoints predominate at or near the centromere, and a recent study showed a marker chromosome derived from 17p11.2 had breakpoints within an LCR and the centromere (44). The involvement of the centromere may be due to the variation in condensation of the heterochromatin, which may create instability. These data provide evidence that genomic architecture other than LCRs, such as centromeres, pericentromeric repeats, and telomeres, may be involved in the origin of both non-recurrent and recurrent rearrangements. However, nucleotide sequence of most non-recurrent recombinant junctions, and therefore the mechanisms by which they occur, have yet to be elucidated.

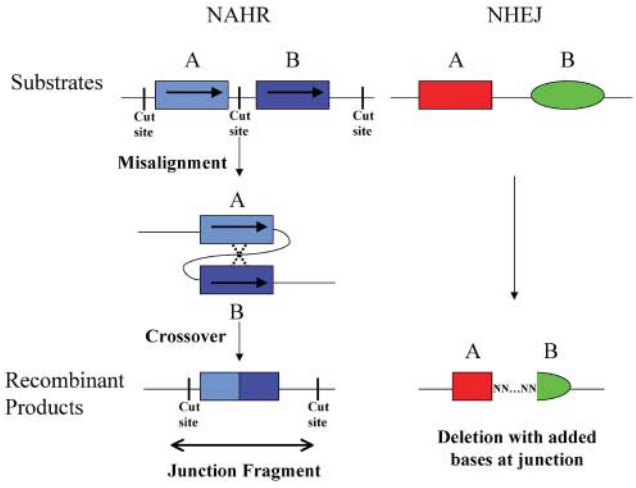

Figure 2. Generation of deletion rearrangement by NAHR and NHEJ. The substrates and products of recombination are shown. NAHR (left), utilizes two non-allelic LCRs (A and B) as substrates for recombination. The LCRs are depicted as blue rectangles, due to high homology, but are different shades of blue, signifying the few *cis*-morphisms, or paralogous sequence variants, that distinguish them. LCRs A and B, directly oriented (shown by arrows) misalign, and subsequent homologous recombination results in a deletion with a single recombinant LCR, shown as a two-tone blue rectangle. Restriction enzyme consensus sequences (cut sites) are depicted as vertical lines on either side of the recombinant LCR, with deletion of the consensus sequence between the two substrate LCRs. Digestion using this enzyme results in the isolation of a recombination-specific junction fragment, shown below. NHEJ (right), utilizes two non-homologous sequences [red rectangle (A) and green oval (B)] as substrates for recombination. The two sequences are joined via NHEJ, with deletion of the intervening fragment. Additional bases (NN . . . NN) are added at the deletion junction.

NON-HOMOLOGOUS END JOINING

LCR-mediated NAHR does not explain all cases of genomic rearrangement. Several genomic disorders are associated with rearrangements whose breakpoints do not cluster within LCRs, but often occur within apparently unique sequence. Sequencing of deletion junctions in the *dystrophin* gene associated with Duchenne muscular dystrophy (DMD) showed a scattering of breakpoints throughout the gene, with *Alu* and long tandem repeat (LTR) elements present at 3/10 (33%) of breakpoints, and unique sequence located at the remainder (45). The sequence TTTAAA, known to be able to curve the DNA molecule (46), was found at or near three of the junctions studied. Taken together, these data suggest a NHEJ mechanism of deletion formation, possibly stimulated by DSBs in the curved DNA structure.

Studies of the products of recombination in three patients with different sized *PLP*1 deletions in Xq22 implicated NHEJ as a causal mechanism. Sequence analysis of three deletion junctions revealed no homologous sequence at the breakpoint junctions; however, two of the distal breakpoints were embedded in a novel 32 kb LCR, termed LCR-PMDB (26). In both cases, a sequence of either 12 bp or 34 bp of unknown origin was located at the deletion junction, which is common to rearrangements generated via NHEJ (Fig. 2). Additionally, duplications of the same region in Xq22, which are more frequently observed in patients than are deletions, also vary in size and have scattered breakpoints (47,48). Sequencing of the duplication junction in three patients with different sized duplications localized the telomeric breakpoints within different repetitive elements, L1PA7, *Alu*Sp, and L1ME3B (Fig. 1) (49). One of these breakpoints was mapped in an X-chromosome specific LCR-rich region; however, there was no homology between the centromeric and telomeric breakpoint flanking sequences (49). The presence of LCRs and *Alu* elements at some of the breakpoints indicates that genome architecture may stimulate, but not necessarily mediate, non-recurrent rearrangements (27).

Recently, breakpoint mapping studies of 60 deletions involving 1p36 (43 terminal, four interstitial, three complex, and 10 derivative chromosomes) revealed a scattering of breakpoints throughout the distal 10.5 Mb of chromosome 1p, with no common breakpoints (50). Somatic cell hybrid analysis of three of the terminal deletions demonstrated that one deletion was stabilized by telomeric repeat sequences and two deletions were associated with cryptic interrupted inverted duplications at the end of the chromosomes (51). Sequencing of the breakpoint junctions of these two deletions revealed a structure identical to chromosomes that have gone through breakage–fusion–bridge (BFB) cycles in which uncapped sister chromatids are fused by NHEJ (52,53).

NHEJ also may be responsible for the recently identified duplications of 10q24, associated with SHFM3 (5). Seven patients with SHFM3 were found to have a ~0.5 Mb duplication, with proximal and distal breakpoints clustering within 130 and 80 kb regions, respectively (5). Although the duplication breakpoints cluster, as would be expected if NAHR occurred, there is no evidence (as of build 34 of the human

genome) of the presence of LCRs at the breakpoint regions. An increased density (20%) of *Alu* elements was observed in both breakpoint regions, although sequence analysis of the breakpoint junctions in two patients did not detect any *Alu* elements within 50 bp of the junctions (5). This observation suggests that NHEJ, possibly mediated by the abundant repetitive elements, may be responsible for these rearrangements.

MOLECULAR DIAGNOSIS OF GENOMIC REARRANGEMENTS

During the last two decades, technology developments have enabled a higher resolution analysis of the human genome. The diagnosis of genomic rearrangements has seen a shift from cytogenetic techniques such as G-banding to locus-specific fluorescent *in situ* hybridization (FISH), chromosome painting, and telomere FISH (54). PFGE, used to detect a rearrangement-specific junction fragment for common rearrangements, is now considered time and labor-intensive compared with new technologies. Recently, array-CGH using BAC and PAC clones has been successfully used to identify genomic deletions and duplications (55–58). This technology is higher throughput than FISH and PFGE, and may be especially useful in identifying new genomic disorders, or in detecting submicroscopic rearrangements not visible by routine chromosome analysis (59). The array-CGH technology may also detect reciprocal duplications of common microdeletions, which are presumably under-ascertained due to the mild degree or lack of appreciable phenotypes. Although the reciprocal duplications are expected to occur at the same frequency as deletions, only CMT1A/HNPP, SMS/dup(17)(p11.2p11.2) syndrome, DGS/VCFS and the newly described dup(22)(q11.2q11.2) syndrome, and Y chromosome *AZF*a deletions/duplications have been identified as reciprocal deletion/duplication syndromes (9,12,60,61).

GENOME REARRANGEMENTS AND PRIMATE EVOLUTION

Genomic architectural features such as LCRs and *Alu* elements have evolved only recently in the primate lineage. Comparative genomic analysis between humans and chimpanzees, our closest ancestor, has shown 98.8% identity (62). Karyotype analysis of the respective genomes reveals tremendous similarity; several chromosomal rearrangements (nine pericentric inversions and an acrocentric fusion) have occurred that define the human karyotype (62,63). The role of genomic architecture in these rearrangements is apparent, as both the evolutionary t(4;19) translocation in gorilla and two pericentric inversion breakpoints in chimpanzee have been localized to LCRs in the orthologous chromosomal regions (64–66). In addition to karyotypic differences, smaller indel events appear to be a major source of variation between the primates (67,68). Thus, it seems that the driving force of evolution may be genomic rearrangements rather than single nucleotide changes. This is supported by genomic disorders, as the rearrangements of our genome are apparent from generation to generation.

POTENTIAL MECHANISMS FOR REARRANGEMENTS ASSOCIATED WITH DISEASE

Rearrangement breakpoints are associated with LCRs far more frequently than would be expected if the rearrangements occurred randomly. Despite large stretches of high sequence identity, it appears that 'hotspots' exist for the majority of the crossovers that occur within LCRs (13–19). Previous work has shown that positional preferences also exist for allelic homologous recombination, resulting in transmission of haplotype 'blocks' (69). Taken together, these data suggest that both allelic and non-allelic recombination possibly may take place at the same hotspots throughout the genome, wherein programmed DSBs occur and initiate recombination in meiosis. Resolution of Holliday structures formed between non-allelic LCR copies could result in rearrangements or gene conversion events. The latter could potentially be responsible for increased polymorphic variation at crossover preference regions or may further homogenize LCRs (70). Additionally, meiotic recombination is known to be elevated near telomeres, possibly suggesting a role for the frequent involvement of this region in rearrangements (71,72). In mammalian cells, interstitial telomeric sequences (also present in humans) (73) have been shown to increase rearrangements by up to 30-fold (74).

The positional preference for strand exchange seen in NAHR may suggest the presence of additional architectural features at the hotspots that make the region more prone to recombination. AT-rich palindromes are located near several of the hotspots, suggesting that a predisposition to DSBs may possibly influence the location of strand exchange (18,34,38,39). In support of this, studies in mice have shown that large palindromes in the germ line are extremely unstable and undergo stabilizing rearrangements at frequencies up to 56%, often through deletions (75–77). Elevated gene conversion events in and adjacent to the palindromes were also documented (76,77). Palindromes and other sequence features, such as triplet repeats, are able to form hairpin structures, potentially exposing DNA to an increased frequency of spontaneous DSBs and subsequent rearrangements, as seen in mammalian cells and patients with either Fragile X or Jacobsen syndromes (Fig. 2) (78–83).

Other potentially *cis*-acting elements such as a *mariner* transposon-like element, minisatellite-like sequences, and chi-like sequences have been identified near the CMT1A/HNPP and NF1 hotspots (Fig. 1) (13,14,16). These sequences have not previously been implicated in human recombination events, although it is possible that their presence also increases the likelihood for DSBs, which then must be resolved by patch repair and heteroduplex resolution, potentially within the hotspot (13,84). These same architectural features may be associated with rearrangements resulting from NHEJ, which are thought to be initiated by DSBs. Few NHEJ recombinant junctions have been studied at the nucleotide sequence level, thus further investigations of the sequences near scattered breakpoints are warranted.

Rather than a *cis*-acting nucleotide sequence stimulating recombination, another possibility for the positional preference of crossovers associated with rearrangements is a constraint on access to the DNA because of the chromatin structure of the

region. An open chromatin structure may expose DNA to DSBs or other damage, that is then repaired in an aberrant fashion, yielding rearrangements.

The group of genomic disorders has evolved to encompass not only deletions, duplications, inversions and translocations, but also somatic rearrangements associated with malignancies. The aforementioned investigations have shown that the majority of genomic rearrangements are not random events, but in fact represent potential mechanical errors inherent in the maintenance of a genome complicated by complex architecture. Perhaps the same genome flexibility that has enabled us to evolve relatively rapidly also makes us as a species more susceptible to rearrangements associated with disease.

ACKNOWLEDGEMENTS

We appreciate the critical reviews of Drs W. Bi, K. Inoue, P. Stankiewicz, and J.H. Wilson. This work has been generously supported by the National Institute for Neurological Disorders and Strokes (RO1 NS27042), and the National Institute for Child Health and Development (PO1 HD39420).

REFERENCES

1. Lupski, J.R. (1998) Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet.*, **14**, 417–422.
2. Stankiewicz, P. and Lupski, J.R. (2002) Genome architecture, rearrangements and genomic disorders. *Trends Genet.*, **18**, 74–82.
3. Lupski, J.R. (2003) 2002 Curt Stern Award Address. Genomic disorders recombination-based disease resulting from genomic architecture. *Am. J. Hum. Genet.*, **72**, 246–252.
4. Kurotaki, N., Harada, N., Shimokawa, O., Miyake, N., Kawame, H., Uetake, K., Makita, Y., Kondoh, T., Ogata, T., Hasegawa, T. *et al.* (2003) Fifty microdeletions among 112 cases of Sotos syndrome: Low copy repeats possibly mediate the common deletion. *Hum. Mutat.*, **22**, 378–387.
5. de Mollerat, X.J., Gurrieri, F., Morgan, C.T., Sangiorgi, E., Everman, D.B., Gaspari, P., Amiel, J., Bamshad, M.J., Lyle, R., Blouin, J.L. *et al.* (2003) A genomic rearrangement resulting in a tandem duplication is associated with split hand-split foot malformation 3 (SHFM3) at 10q24. *Hum. Mol. Genet.*, **12**, 1959–1971.
6. Milunsky, J.M., Huang, X.L. (2003) Unmasking Kabuki syndrome: chromosome 8p22–8p23.1 duplication revealed by comparative genomic hybridization and BAC-FISH. *Clin. Genet.*, **64**, 509–516.
7. Roth, D.B. and Wilson, J.H. (1986) Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Mol. Cell. Biol.*, **6**, 4295–4304.
8. Lupski, J.R., de Oca-Luna, R.M., Slaugenhaupt, S., Pentao, L., Guzzetta, V., Trask, B.J., Saucedo-Cardenas, O., Barker, D.F., Killian, J.M., Garcia, C.A. *et al.* (1991) DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell*, **66**, 219–232.
9. Chance, P.F., Abbas, N., Lensch, M.W., Pentao, L., Roa, B.B., Patel, P.I. and Lupski, J.R. (1994) Two autosomal dominant neuropathies result from reciprocal DNA duplication/deletion of a region on chromosome 17. *Hum. Mol. Genet.*, **3**, 223–228.
10. Chen, K.S., Manian, P., Koeuth, T., Potocki, L., Zhao, Q., Chinault, A.C., Lee, C.C. and Lupski, J.R. (1997) Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nat. Genet.*, **17**, 154–163.
11. Perez Jurado, L.A., Wang, Y.K., Peoples, R., Coloma, A., Cruces, J. and Francke, U. (1998) A duplicated gene in the breakpoint regions of the 7q11.23 Williams–Beuren syndrome deletion encodes the initiator binding protein TFII-I and BAP-135, a phosphorylation target of BTK. *Hum. Mol. Genet.*, **7**, 325–334.
12. Potocki, L., Chen, K.S., Park, S.S., Osterholm, D.E., Withers, M.A., Kimonis, V., Summers, A.M., Meschino, W.S., Anyane-Yeboa, K., Kashork, C.D. *et al.* (2000) Molecular mechanism for duplication 17p11.2—the homologous recombination reciprocal of the Smith–Magenis microdeletion. *Nat. Genet.*, **24**, 84–87.
13. Reiter, L.T., Murakami, T., Koeuth, T., Pentao, L., Muzny, D.M., Gibbs, R.A. and Lupski, J.R. (1996) A recombination hotspot responsible for two inherited peripheral neuropathies is located near a mariner transposon-like element. *Nat. Genet.*, **12**, 288–297.
14. Lopes, J., Tardieu, S., Silander, K., Blair, I., Vandenberghe, A., Palau, F., Ruberg, M., Brice, A. and LeGuern, E. (1999) Homologous DNA exchanges in humans can be explained by the yeast double-strand break repair model: a study of 17p11.2 rearrangements associated with CMT1A and HNPP. *Hum. Mol. Genet.*, **8**, 2285–2292.
15. Reiter, L.T., Hastings, P.J., Nelis, E., De Jonghe, P., Van Broeckhoven, C. and Lupski, J.R. (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.
16. Lopez-Correa, C., Dorschner, M., Brems, H., Lazaro, C., Clementi, M., Upadhyaya, M., Dooijes, D., Moog, U., Kehrer-Sawatzki, H., Rutkowski, J.L. *et al.* (2001) Recombination hotspot in NF1 microdeletion patients. *Hum. Mol. Genet.*, **10**, 1387–1392.
17. Bayes, M., Magano, L.F., Rivera, N., Flores, R. and Perez Jurado, L.A. (2003) Mutational mechanisms of Williams–Beuren syndrome deletions. *Am. J. Hum. Genet.*, **73**, 131–151.
18. Bi, W., Park, S.-S., Shaw, C.J, Withers, M.A., Patel, P.I., Lupski, J.R. (2003) Reciprocal crossovers and a positional preference for strand exchange in recombination events resulting in deletion/duplication 17p11.2. *Am. J. Hum. Genet.*, **73**, 1302–1315.
19. Repping, S., Skaletsky, H., Lange, J., Silber, S., Van Der Veen, F., Oates, R.D., Page, D.C. and Rozen, S. (2002) Recombination between palindromes P5 and P1 on the human Y chromosome causes massive deletions and spermatogenic failure. *Am. J. Hum. Genet.*, **71**, 906–922.
20. Smit, A.F. (1996) The origin of interspersed repeats in the human genome. *Curr. Opin. Genet. Dev.*, **6**, 743–748.
21. Kuroda-Kawaguchi, T., Skaletsky, H., Brown, L.G., Minx, P.J., Cordum, H.S., Waterston, R.H., Wilson, R.K., Silber, S., Oates, R., Rozen, S. *et al.* (2001) The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nat. Genet.*, **29**, 279–286.
22. Barbouti, A., Stankiewicz, P., Birren, B., Nusbaum, C., Cuomo, C., Hoglund, M., Johansson, B., Hagemeijer, A., Park, S.-S., Mitelman, F., Lupski, J.R., Fioretos, T. (2004) 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.*, **74**, 1–10.
23. Saglio, G., Storlazzi, C.T., Giugliano, E., Surace, C., Anelli, L., Rege-Cambrin, G., Zagaria, A., Jimenez Velasco, A., Heiniger, A., Scaravaglio, P. *et al.* (2002) A 76-kb duplicon maps close to the BCR gene on chromosome 22 and the ABL gene on chromosome 9: possible involvement in the genesis of the Philadelphia chromosome translocation. *Proc. Natl Acad. Sci. USA*, **99**, 9882–9887.
24. Gimelli, G., Pujana, M.A., Patricelli, M.G., Russo, S., Giardino, D., Larizza, L., Cheung, J., Armengol, L., Schinzel, A., Estivill, X. *et al.* (2003) Genomic inversions of human chromosome 15q11-q13 in mothers of Angelman syndrome patients with class II (BP2/3) deletions. *Hum. Mol. Genet.*, **12**, 849–858.
25. Giglio, S., Calvari, V., Gregato, G., Gimelli, G., Camanini, S., Giorda, R., Ragusa, A., Guerneri, S., Selicorni, A., Stumm, M. *et al.* (2002) Heterozygous submicroscopic inversions involving olfactory receptor–gene clusters mediate the recurrent t(4;8)(p16;p23) translocation. *Am. J. Hum. Genet.*, **71**, 276–285.
26. Inoue, K., Osaka, H., Thurston, V.C., Clarke, J.T., Yoneyama, A., Rosenbarker, L., Bird, T.D., Hodes, M.E., Shaffer, L.G. and Lupski, J.R. (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.
27. Stankiewicz, P., Shaw, C.J., Dapper, J.D., Wakui, K., Shaffer, L.G., Withers, M., Elizondo, L., Park, S.S. and Lupski, J.R. (2003) Genome architecture catalyzes nonrecurrent chromosomal rearrangements. *Am. J. Hum. Genet.*, **72**, 1101–1116.
28. Bailey, J.A., Yavor, A.M., Massa, H.F., Trask, B.J. and Eichler, E.E. (2001) Segmental duplications: organization and impact within the current human genome project assembly. *Genome Res.*, **11**, 1005–1017.

29. Babcock, M., Pavlicek, A., Spiteri, E., Kashork, C.D., Ioshikhes, I., Shaffer, L.G., Jurka, J., Morrow, B.E. (2003) Shuffling of genes within low-copy repeats on 22q11 (LCR22) by Alu–mediated recombination events during evolution. *Genome Res.*, **13**, 2519–2532.
30. Bailey, J.A., Liu, G. and Eichler, E.E. (2003) An alu transposition model for the origin and expansion of human segmental duplications. *Am. J. Hum. Genet.*, **73**, 823–834.
31. Bailey, J.A., Yavor, A.M., Viggiano, L., Misceo, D., Horvath, J.E., Archidiacono, N., Schwartz, S., Rocchi, M. and Eichler, E.E. (2002) Human-specific duplication and mosaic transcripts: the recent paralogous structure of chromosome 22. *Am. J. Hum. Genet.*, **70**, 83–100.
32. Stankiewicz, P. and Lupski, J.R. (2002) Molecular-evolutionary mechanisms for genomic disorders. *Curr. Opin. Genet. Dev.*, **12**, 312–319.
33. Stankiewicz, P., Cheung, S.W., Shaw, C.J., Saleki, R., Szigeti, K. and Lupski, J.R. (2003) The donor chromosome breakpoint for a jumping translocation is associated with large low-copy repeats in 21q21.3. *Cytogenet Genome Res.*, **101**, 118–123.
34. Edelmann, L., Spiteri, E., Koren, K., Pulijaal, V., Bialer, M.G., Shanske, A., Goldberg, R. and Morrow, B.E. (2001) AT-rich palindromes mediate the constitutional t(11;22) translocation. *Am. J. Hum. Genet.*, **68**, 1–13.
35. Spiteri, E., Babcock, M., Kashork, C.D., Wakui, K., Gogineni, S., Lewis, D.A., Williams, K.M., Minoshima, S., Sasaki, T., Shimizu, N. *et al.* (2003) Frequent translocations occur between low copy repeats on chromosome 22q11.2 (LCR22s) and telomeric bands of partner chromosomes. *Hum. Mol. Genet.*, **12**, 1823–1837.
36. Kurahashi, H., Shaikh, T.H., Hu, P., Roe, B.A., Emanuel, B.S. and Budarf, M.L. (2000) Regions of genomic instability on 22q11 and 11q23 as the etiology for the recurrent constitutional t(11;22). *Hum. Mol. Genet.*, **9**, 1665–1670.
37. Kurahashi, H. and Emanuel, B.S. (2001) Long AT-rich palindromes and the constitutional t(11;22) breakpoint. *Hum. Mol. Genet.*, **10**, 2605–2617.
38. Kurahashi, H., Shaikh, T., Takata, M., Toda, T. and Emanuel, B.S. (2003) The constitutional t(17;22): another translocation mediated by palindromic AT-rich repeats. *Am. J. Hum. Genet.*, **72**, 733–738.
39. Nimmakayalu, M.A., Gotter, A.L., Shaikh, T.H. and Emanuel, B.S. (2003) A novel sequence-based approach to localize translocation breakpoints identifies the molecular basis of a t(4;22). *Hum. Mol. Genet.*, **12**, 2817–2825.
40. Gotter, A.L., Shaikh, T.H., Budarf, M.L., Rhodes, C.H. and Emanuel, B.S. (2004) A palindrome-mediated mechanism distinguishes translocations involving LCR-B of chromosome 22q11.2. *Hum. Mol. Genet.*, **13**, 103–115.
41. Huang, B., Crolla, J.A., Christian, S.L., Wolf-Ledbetter, M.E., Macha, M.E., Papenhausen, P.N. and Ledbetter, D.H. (1997) Refined molecular characterization of the breakpoints in small inv dup(15) chromosomes. *Hum. Genet.*, **99**, 11–17.
42. Wandstrat, A.E., Leana-Cox, J., Jenkins, L. and Schwartz, S. (1998) Molecular cytogenetic evidence for a common breakpoint in the largest inverted duplications of chromosome 15. *Am. J. Hum. Genet.*, **62**, 925–936.
43. McTaggart, K.E., Budarf, M.L., Driscoll, D.A., Emanuel, B.S., Ferreira, P. and McDermid, H.E. (1998) Cat eye syndrome chromosome breakpoint clustering: identification of two intervals also associated with 22q11 deletion syndrome breakpoints. *Cytogenet. Cell Genet.*, **81**, 222–228.
44. Stankiewicz, P., Parka, S.S., Holder, S.E., Waters, C.S., Palmer, R.W., Berend, S.A., Shaffer, L.G., Potocki, L. and Lupski, J.R. (2001) Trisomy 17p10-p12 resulting from a supernumerary marker chromosome derived from chromosome 17: molecular analysis and delineation of the phenotype. *Clin. Genet.*, **60**, 336–344.
45. Nobile, C., Toffolatti, L., Rizzi, F., Simionati, B., Nigro, V., Cardazzo, B., Patarnello, T., Valle, G. and Danieli, G.A. (2002) Analysis of 22 deletion breakpoints in dystrophin intron 49. *Hum. Genet.*, **110**, 418–421.
46. Singh, G.B., Kramer, J.A. and Krawetz, S.A. (1997) Mathematical model to predict regions of chromatin attachment to the nuclear matrix. *Nucl. Acids Res.*, **25**, 1419–1425.
47. Inoue, K., Osaka, H., Imaizumi, K., Nezu, A., Takanashi, J., Arii, J., Murayama, K., Ono, J., Kikawa, Y., Mito, T. *et al.* (1999) Proteolipid protein gene duplications causing Pelizaeus–Merzbacher disease: molecular mechanism and phenotypic manifestations. *Ann. Neurol.*, **45**, 624–632.
48. Hobson, G., Cundall, M., Sperle, K., Kamholz, J., Garbern, J., Heng, H., Sistermans, E., Malcolm, S., Woodward, K. (2003) Fine mapping of duplication endpoints in Pelizaeus–Merzbacher disease. *Am. J. Hum. Genet.*, **73** (suppl.), 342.
49. Iwaki, A., Kondo, J., Ototsuji, M., Kurosawa, K., Fukumaki, Y. (2003) Characterization of the breakpoints of PLP1 duplication in three cases of Pelizaeus–Merzbacher disease. *Am. J. Hum. Genet.*, **73** (suppl.), 549.
50. Heilstedt, H.A., Ballif, B.C., Howa L.A., Lewis, R.A., Stal, S., Kashork, C.D., Bacino, C.A., Shapira, S.K. and Shaffer, L.G. (2003) Physical map of 1p36, placement of breakpoints in monosomy 1p36, and clinical characterization of the syndrome. *Am. J. Hum. Genet.*, **72**, 1200–1212.
51. Ballif, B.C., Yu, W., Shaw, C.A., Kashork, C.D. and Shaffer, L.G. (2003) Monosomy 1p36 breakpoint junctions suggest pre-meiotic breakage fusion–bridge cycles are involved in generating terminal deletions. *Hum. Mol. Genet.*, **12**, 2153–2165.
52. Fouladi, B., Sabatier, L., Miller, D., Pottier, G. and Murnane, J.P. (2000) The relationship between spontaneous telomere loss and chromosome instability in a human tumor cell line. *Neoplasia*, **2**, 540–554.
53. Lo, A.W., Sprung, C.N., Fouladi, B., Pedram, M., Sabatier, L., Ricoul, M., Reynolds, G.E. and Murnane, J.P. (2002) Chromosome instability as a result of double-strand breaks near telomeres in mouse embryonic stem cells. *Mol. Cell. Biol.*, **22**, 4836–4850.
54. Flint, J. and Knight, S. (2003) The use of telomere probes to investigate submicroscopic rearrangements associated with mental retardation. *Curr. Opin. Genet. Dev.*, **13**, 310–316.
55. Bruder, C.E., Hirvela, C., Tapia-Paez, I., Fransson, I., Segraves, R., Hamilton, G., Zhang, X.X., Evans, D.G., Wallace, A.J., Baser, M.E. *et al.* (2001) High resolution deletion analysis of constitutional DNA from neurofibromatosis type 2 (NF2) patients using microarray-CGH. *Hum. Mol. Genet.*, **10**, 271–282.
56. Veltman, J.A., Jonkers, Y., Nuijten, I., Janssen, I., van der Vliet, W., Huys, E., Vermeesch, J., Van Buggenhout, G., Fryns, J.P., Admiraal, R. *et al.* (2003) Definition of a critical region on chromosome 18 for congenital aural atresia by arrayCGH. *Am. J. Hum. Genet.*, **72**, 1578–1584.
57. Yu, W., Ballif, B.C., Kashork, C.D., Heilstedt, H.A., Howard, L.A., Cai, W.W., White, L.D., Liu, W., Beaudet, A.L., Bejjani, B.A. *et al.* (2003) Development of a comparative genomic hybridization microarray and demonstration of its utility with 25 well-characterized 1p36 deletions. *Hum. Mol. Genet.*, **12**, 2145–2152.
58. Shaw, C.J., Shaw, C.A., Yu, W., Stankiewicz, P., White, L.D., Beaudet, A.L. and Lupski, J.R. (2004) Comparative genomic hybridization using a proximal 17p BAC/PAC array detects rearrangements responsible for four genomic disorders. *J Med Genet.*, **41**, 113–119.
59. Vissers, L.E., De Vries, B.B., Osoegawa, K., Janssen, I.M., Feuth, T., Choy, C.O., Straatman, H., Van Der Vliet, W., Huys, E.H., Van Rijk, A. *et al.* (2003) Array-based comparative genomic hybridization for the genomewide detection of submicroscopic chromosomal abnormalities. *Am. J. Hum. Genet.*, **73**, 1261–1270.
60. Ensenauer, R.E., Adeyinka, A., Flynn, H.C., Michels, V.V., Lindor, N.M., Dawson, D.B., Thorland, E.C., Lorentz, C.P., Goldstein, J.L., McDonald, M.T. *et al.* (2003) Microduplication 22q11.2, an emerging syndrome: clinical, cytogenetic, and molecular analysis of thirteen patients. *Am. J. Hum. Genet.*, **73**, 1027–1040.
61. Bosch, E. and Jobling, M.A. (2003) Duplications of the AZFa region of the human Y chromosome are mediated by homologous recombination between HERVs and are compatible with male fertility. *Hum. Mol. Genet.*, **12**, 341–347.
62. Navarro, A. and Barton, N.H. (2003) Chromosomal speciation and molecular divergence-accelerated evolution in rearranged chromosomes. *Science*, **300**, 321–324.
63. Warburton, D., Firschein, I.L., Miller, D.A. and Warburton, F.E. (1973) Karyotype of the chimpanzee, Pan troglodytes, based on measurements and banding pattern: comparison to the human karyotype. *Cytogenet. Cell Genet.*, **12**, 453–461.
64. Stankiewicz, P., Park, S.S., Inoue, K. and Lupski, J.R. (2001) The evolutionary chromosome translocation 4;19 in Gorilla gorilla is associated with microduplication of the chromosome fragment syntenic to sequences surrounding the human proximal CMT1A-REP. *Genome Res.*, **11**, 1205–1210.
65. Nickerson, E. (2000) Molecular definition of pericentric inversions distinguishing the genomes of humans and chimpanzees since divergence from a common ancestor. Ph.D. dissertation, Baylor College of Medicine, Houston.
66. Locke, D.P., Archidiacono, N., Misceo, D., Cardone, M.F., Deschamps, S., Roe, B., Rocchi, M. and Eichler, E.E. (2003) Refinement of a chimpanzee pericentric inversion breakpoint to a segmental duplication cluster. *Genome Biol.*, **4**, R50.

67. Britten, R.J., Rowen, L., Williams, J. and Cameron, R.A. (2003) Majority of divergence between closely related DNA samples is due to indels. *Proc. Natl Acad. Sci. USA*, **100**, 4661–4665.
68. Frazer, K.A., Chen, X., Hinds, D.A., Pant, P.V., Patil, N. and Cox, D.R. (2003) Genomic DNA insertions and deletions occur frequently between humans and nonhuman primates. *Genome Res.*, **13**, 341–346.
69. Daly, M.J., Rioux, J.D., Schaffner, S.F., Hudson, T.J. and Lander, E.S. (2001) High-resolution haplotype structure in the human genome. *Nat. Genet.*, **29**, 229–232.
70. Inoue, K. and Lupski, J.R. (2002) Molecular mechanisms for genomic disorders. *A. Rev. Genomics Hum. Genet.*, **3**, 199–242.
71. Ashley, T. (1994) Mammalian meiotic recombination: a reexamination. *Hum. Genet.*, **94**, 587–593.
72. Kipling, D., Wilson, H.E., Thomson, E.J., Lee, M., Perry, J., Palmer, S., Ashworth, A. and Cooke, H.J. (1996) Structural variation of the pseudoautosomal region between and within inbred mouse strains. *Proc. Natl Acad. Sci. USA*, **93**, 171–175.
73. Azzalin, C.M., Mucciolo, E., Bertoni, L. and Giulotto, E. (1997) Fluorescence *in situ* hybridization with a synthetic $(T2AG3)_n$ polynucleotide detects several intrachromosomal telomere-like repeats on human chromosomes. *Cytogenet. Cell Genet.*, **78**, 112–115.
74. Kilburn, A.E., Shea, M.J., Sargent, R.G. and Wilson, J.H. (2001) Insertion of a telomere repeat sequence into a mammalian gene causes chromosome instability. *Mol. Cell. Biol.*, **21**, 126–135.
75. Akgun, E., Zahn, J., Baumes, S., Brown, G., Liang, F., Romanienko, P.J., Lewis, S. and Jasin, M. (1997) Palindrome resolution and recombination in the mammalian germ line. *Mol. Cell. Biol.*, **17**, 5559–5570.
76. Lewis, S., Akgun, E. and Jasin, M. (1999) Palindromic DNA and genome stability. Further studies. *Ann. NY Acad. Sci.*, **870**, 45–57.
77. Zhou, Z.H., Akgun, E. and Jasin, M. (2001) Repeat expansion by homologous recombination in the mouse germ line at palindromic sequences. *Proc. Natl Acad. Sci. USA*, **98**, 8326–8333.
78. Meservy, J.L., Sargent, R.G., Iyer, R.R., Chan, F., McKenzie, G.J., Wells, R.D. and Wilson, J.H. (2003) Long CTG tracts from the myotonic dystrophy gene induce deletions and rearrangements during recombination at the APRT locus in CHO cells. *Mol. Cell. Biol.*, **23**, 3152–3162.
79. Meijer, H., de Graaff, E., Merckx, D.M., Jongbloed, R.J., de Die-Smulders, C.E., Engelen, J.J., Fryns, J.P., Curfs, P.M. and Oostra, B.A. (1994) A deletion of 1.6 kb proximal to the CGG repeat of the FMR1 gene causes the clinical phenotype of the fragile X syndrome. *Hum. Mol. Genet.*, **3**, 615–620.
80. de Graaff, E., Rouillard, P., Willems, P.J., Smits, A.P., Rousseau, F. and Oostra, B.A. (1995) Hotspot for deletions in the CGG repeat region of FMR1 in fragile X patients. *Hum. Mol. Genet.*, **4**, 45–49.
81. Mila, M., Castellvi-Bel, S., Sanchez, A., Lazaro, C., Villa, M. and Estivill, X. (1996) Mosaicism for the fragile X syndrome full mutation and deletions within the CGG repeat of the FMR1 gene. *J. Med. Genet.*, **33**, 338–340.
82. Schmucker, B., Ballhausen, W.G. and Pfeiffer, R.A. (1996) Mosaicism of a microdeletion of 486 bp involving the CGG repeat of the FMR1 gene due to misalignment of GTT tandem repeats at chi-like elements flanking both breakpoints and a full mutation. *Hum. Genet.*, **98**, 409–414.
83. Jones, C., Mullenbach, R., Grossfeld, P., Auer, R., Favier, R., Chien, K., James, M., Tunnacliffe, A. and Cotter, F. (2000) Co-localisation of CCG repeats and chromosome deletion breakpoints in Jacobsen syndrome: evidence for a common mechanism of chromosome breakage. *Hum. Mol. Genet.*, **9**, 1201–1208.
84. Stankiewicz, P., Inoue, K., Bi, W., Walz, K., Park, S.-S., Kurotaki, N., Shaw, C.J., Fonseca, P., Yan, J., Lee, J.A., Khajavi, M. and Lupski, J.R. (2004) Genomic disorders—genome architecture results in susceptibility to DNA rearrangements causing common human traits. *Cold Spring Harb. Symp. Quant. Biol.*, **68**, in press.