Published in final edited form as:

Trends Genet. 2009 July ; 25(7): 298–307. doi:10.1016/j.tig.2009.05.005.

Complex human chromosomal and genomic rearrangements

Feng Zhang¹, Claudia M.B. Carvalho¹, and James R. Lupski^{1,2,3}

¹Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA

²Department of Pediatrics, Baylor College of Medicine, Houston, TX 77030, USA

³Texas Children's Hospital, Houston, TX 77030, USA

Abstract

Copy number variation (CNV) is a major source of genetic variation among humans. In addition to existing as benign polymorphisms, CNVs can also convey clinical phenotypes, including genomic disorders, sporadic diseases and complex human traits. CNV results from genomic rearrangements that can represent simple deletion or duplication of a genomic segment, or be more complex. Complex chromosomal rearrangements (CCRs) have been known for some time but their mechanisms have remained elusive. Recent technology advances and high-resolution human genome analyses have revealed that complex genomic rearrangements can account for a large fraction of non-recurrent rearrangements at a given locus. Various mechanisms, most of which are DNA-replication-based, for example fork stalling and template switching (FoSTeS) and microhomology-mediated break-induced replication (MMBIR), have been proposed for generating such complex genomic rearrangements and are probably responsible for CCR.

The nature of copy number variation and the link with human disease

In human and medical genetics we are often faced with a molecular observation or phenomenon that we seek to explain. Such phenomena often consist of a genetic alteration in the context of either a Mendelian or sporadic trait; usually a disease or birth defect. One major limitation of human genetics is the inability to perform the 'test cross' that is so experimentally useful in other genetic model organisms. It is a fact of life that one is often unable to execute either preferred or even desired matings in humans. Nevertheless, positive aspects of human genetics include a wealth of subjects with a myriad of biological variation, tremendous reagents and resources, ready ascertainment by virtue of the conveyed phenotype eliciting a visit to a physician and, perhaps most importantly, the study of phenomena directly relevant to the human organism. Whereas in other model organisms we might be able to experimentally manipulate them to surmise mechanism, the relevance to humans is by inference or analogy alone, not by direct observations. Here, we review the recent experimental observations revealing complex rearrangements of the human genome in association with different disease traits. We present the evidence for their existence and potential explanations (i.e. mechanisms) for their cause.

Whereas Watson–Crick base-pair changes have long been appreciated as a cause of mutation, and millions of single nucleotide polymorphisms (SNPs) are known to exist in human individuals and populations [1,2], it was not until recently that copy number variation (CNV; see Glossary) has been found to represent a major source for human genetic variation and genome diversity. In the past few years, a large number of CNVs have been identified by various genome-wide technologies, including array comparative genomic hybridization (aCGH), SNP genotyping arrays, and next-generation sequencing using ‘paired-end’ methods [3–13]. According to the latest statistics (March 2009), >38 000 entries of structural variations (SVs) have been included in the Database of Genomic Variants (DGV; <http://projects.tcag.ca/variation/>).

Many of the CNVs in DGV might represent benign polymorphisms; however, many CNVs can convey clinical phenotypes by various molecular mechanisms [14,15]. For example, deletion CNVs that disrupt functional genes can cause loss-of-function mutations and lead to human diseases such as color blindness [16], whereas CNVs that alter the copy number of dosage-sensitive genes can also result in disorders, for example, Charcot-Marie-Tooth disease type 1A (CMT1A; Mendelian Inheritance in Man [MIM] 118220; <http://www.ncbi.nlm.nih.gov/omim>) and hereditary neuropathy with liability to pressure palsies (HNPP; MIM 162500) [17]. Both disorders are peripheral neuropathies caused by CNV of the peripheral myelin protein 22 gene (PMP22). To date, many disease-associated CNVs have been reported [14,18]. The clinical phenotypes resulting from genomic structural changes owing to genomic instability and arising because of susceptible genome architecture have been referred to as genomic disorders [19,20]. In addition to being associated with Mendelian traits and sporadic diseases, CNVs have also been recently shown to cause human complex traits, such as susceptibility to HIV infection [21], autism [22,23], schizophrenia [24–26] and mental retardation [27–29]; for a review, see Ref. [30]. Thus, uncovering the mechanisms underlying CNV formation has tremendous implications for the diagnostics of CNV-associated diseases. Molecular mechanisms for simple rearrangements (e.g. simple deletion or duplication of a locus), either recurrent or non-recurrent, have been well studied (for recent reviews, see Refs [30,31]).

However, the complexity of selected human genomic rearrangements is becoming increasingly appreciated. In this article, we review the evidence for such complexity at the chromosomal and genomic level and discuss the potential mechanisms to explain these phenomena.

Complexity is not uncommon for human genome rearrangements

Complex chromosomal rearrangements

Complex chromosomal rearrangements (CCRs) describe genome structural rearrangements that involve at least three cytogenetically visible breakpoints and the exchange of genetic material between two or more chromosomes [32] (Table 1). These are rare, although clinically important to recognize because carriers can have phenotypes that include normal individuals, recurrent miscarriages in females, infertile males, mental retardation and congenital abnormalities [33–38]. The alterations can be *de novo* or familial; familial CCRs tend to involve fewer chromosomes and fewer breakpoints than *de novo* CCRs [39]. A

survey of 269 371 prenatal studies on amniotic fluid, chorionic villus and fetal blood samples reported a total of 246 apparently cytogenetically balanced anomalies; among them, 3% were *de novo* presumably balanced CCRs [38]. This study also showed a sex ratio bias for the occurrence of CCRs, females are carriers ~2 times more frequently than males, confirming previous data on familial cases [39]. Interestingly, the origin of *de novo* CCR and reciprocal translocation cases are frequently reported as paternal [39–41], although 77% of the prenatal cases with balanced rearrangements analysed by Giardino *et al.* [38] were referred owing to advanced maternal age.

The occurrence of *de novo* CCR is associated with an increased risk of mental retardation and the degree of severity correlates with a higher number of breakpoints [33,42]. The development of the clinical phenotype is believed to be caused by the alteration of the regulation of genes located nearby (i.e. a position effect [43]), by disruption of a gene at or near the breakpoint or by imbalances at the breakpoints or elsewhere in the genome [41]. Despite the importance of refining the multiple rearrangement breakpoints at the sequence level in CCR cases, virtually no breakpoints have been sequenced and no molecular mechanisms have been proposed for how they might occur. To date, most of the breakpoints have been mapped using conventional cytogenetic G-banded karyo-typing, multi-subtelomeric fluorescence in situ hybridization (FISH), whole chromosome painting FISH, M-FISH or SKY or multicolor banding (MCB) [34,35,44–46].

More recent studies have used CGH to uncover cryptic rearrangements [40,47–49], but the resolution of the breakpoints is to the level of hundreds or thousands of base pairs. Deletions at the breakpoint regions are a common finding, but duplications or inversions are also detected [35,36,40,50]. Importantly, when the resolution of the analysis methods used to analyse CCR increases, the initially identified number of breakpoints tends to increase [40,48,50–54]. This observation suggests that many, and possibly the majority of, CCRs detected to date might actually be more complex than initially thought. In fact, De Gregori *et al.* [40] reported that 40% of patients reported as ‘balanced translocations’ were unbalanced and, remarkably, 18% of the reciprocal translocations were, instead, complex rearrangements. Interestingly, some patients present alterations not contiguous to the translocation breakpoint [52]; sometimes they are even located in different chromosomes [41,50,53].

The non-random occurrences of CCR breakpoints have been shown in different studies; however, other studies did not find an association with specific chromosomes or cytogenetic bands or genomic regions [40]. Giardino *et al.* [38] reported the preferential involvement of chromosomes 22, 7, 21, 3 and 9 in balanced chromosomal anomalies of all types (including reciprocal translocations, inversions and CCRs) in prenatal cases. Non-random distribution of the breakpoints was also noted in the same study, with most breakpoints co-localized with fragile sites. Vermueulen *et al.* [36] reviewed the literature and showed that 30% of the CCRs had breakpoints on chromosome 7, which was shown to be a segmental-duplication-enriched chromosome (7% according to Ref. [55]). Vissers *et al.* [47] reported two patients in which 17 total breaks were found scattered throughout chromosomes 17 and 21; four of them mapped within low copy repeats (LCRs) usually associated with different recurrent chromosome aberrations, including those causing CMT1A, HNPP, Smith–Magenis

syndrome (SMS; MIM 182290) and Potocki–Lupski syndrome (PTLS; MIM 610883). Batanian and Eswara [34] reviewed 100 CCR cases and noted that band 18q21 has an assigned breakpoint in 11/14 cases where the chromosome 18 was involved. Giardino *et al.* [51] detected the preferential involvement of chromosomes 3, 4, 6 and 10 in 20 prenatally ascertained CCRs; particularly, they reported a patient in which the breakpoints observed on chromosomes 4 and 10 were coincident with bands containing members of the olfactory receptor family. According to them, such preferential involvement is not observed in simple translocations. In fact, after reviewing 226 CCRs reported in the literature to date, it is possible to observe a clear chromosome preference in CCR events (Figure 1). Chromosomes 2, 3, 4, 7 and 11 are the most frequently involved with similar frequencies of approximately 10–12% and chromosomes 17, 22, X and Y are rarely involved, with a frequency of <1% of the cases. The reason for this preference is not obvious.

The mechanism(s) underlying CCR formation remains elusive; some studies propose models based upon the principle of parsimony and the minimum amount of breaks required for the formation of the CCR [35,47]. Li *et al.* [56] propose a translocation–deletion–inversion process involving multiple breakage–fusion events to explain a patient carrying nine breakpoints. A similar model was used by Lindstrand *et al.* [57] to explain a patient with 14 breakpoints along chromosome 1. Patsalis [58] also cites the ‘major’ catastrophe within the gamete (spermatogenesis) hypothesis and remarks that some reports found an association of parental exposure to potential mutagens and *de novo* CCRs. Thus, the phenomena of CCRs with breakpoints that involve many different regions of the human genome can occur. However, because the breakpoint sequences of CCRs have not yet been experimentally determined, the potential association between genomic architecture and the formation of the CCRs, along with the ability to infer or surmise underlying mechanism for formation from rearrangement products, is not possible.

Complex genomic rearrangements

Accumulating data have revealed frequent complex genomic rearrangements of smaller scale at many different loci that are not resolved by conventional G-banded karyotyping and FISH. Actually, complex genomic rearrangements can be revealed at different levels of genome resolution (i.e. resolved at levels of BAC aCGH, oligonucleotide aCGH, or DNA sequencing; Figure 2) and often multiple molecular methods must be used to visualize the complexity.

Employing both oligonucleotide aCGH and breakpoint sequence analyses enabled higher genome resolution for mechanistic studies. Lee *et al.* [59] studied the large non-recurrent genomic duplications of the proteolipid protein 1 (*PLP1*) locus mapped at Xq22 associated with Pelizaeus–Merzbacher disease (PMD; MIM 312080). Interestingly, the PMD patients with three or more copies of *PLP1*, as detected by the locus-specific multiplex ligation-dependent probe amplification (MLPA) assay and suggestive of either triplication or more complex rearrangements, were reported to have a more severe phenotype [60]. The approach by Lee *et al.* [59] combined methods of oligonucleotide aCGH (average resolution of 0.5 kb) and breakpoint sequencing analysis (resolution of 1 bp) and revealed an unanticipated high frequency of complexity, that is, ~65% of non-recurrent *PLP1* duplications were shown

to have interspersed stretches of DNA of normal copy number or triplication and, thus, were shown not to be simple tandem duplications [59] (Table 3). The sequencing data from successfully amplified breakpoint intervals also identified 2–5 bp microhomologies at breakpoint junctions or join points, suggesting a microhomology-mediated rearrangement process [59]. Interestingly, the sequencing of the breakpoint junctions in an aCGH-based ‘simple’ duplication revealed more sequence complexity with microhomologies of GC, AGT and TT at each breakpoint [59] (Figure 2d). Such complex rearrangements are neither rare instances nor limited to certain loci, but can be frequent and occur throughout the human genome (Tables 2,3).

Bauters *et al.* [61] reported two complex duplications out of 16 non-recurrent rearrangements at the methyl CpG-binding protein 2 (*MECP2*) locus on Xq28 (Table 3). Carvalho *et al.* [62] also identified structural complexity in eight out of 30 genomic duplications of the *MECP2* locus (Table 3), of which two show a pattern of two distinct duplication intervals (duplication–normal–duplication) and six are triplications embedded within duplications (duplication–triplication–duplication or triplication–duplication). In addition, complex genomic rearrangements have also been found at other X-linked loci [63–65] (Table 2).

Similar findings also come from the observations of disease-associated *de novo* genomic rearrangements on autosomes (Tables 2,3). For example, complex duplications were identified that involved the platelet-activating factor acetylhydrolase, isoform Ib, subunit 1 (*PAFAH1B1*) gene encoding the LIS1 protein in chromosome 17p13.3, deletions and point mutations of which are associated with Miller–Dieker syndrome (MDS; MIM 247200) and isolated lissencephaly. Bi *et al.* [66] identified three complex rearrangements out of seven cases with copy number changes of *PAFAH1B1*, two of which were *de novo* complex genomic rearrangements, one duplication–normal–duplication and one duplication–triplication–duplication. Interestingly, an additional *PAFAH1B1* duplication was an insertional translocation [66]. However, its breakpoint sequence was not obtained and further complexity could not be evaluated. Complex genomic rearrangements have been identified in other 17p loci associated with PTLIS (57% of non-recurrent duplications) and non-recurrent rearrangements of both CMT1A and HNPP [67]. Notably, a complex exonic rearrangement of *PMP22* has also been identified [67]. This latter complex rearrangement recurred in affected siblings of unaffected parents; the mother of whom was mosaic in her blood for the complex rearrangement, documenting a mitotic event as predicted for a DNA replication mechanism [67].

Examples of complex genomic rearrangements with sequenced breakpoints are listed in Tables 2 and 3. The majority of these breakpoints reveal microhomology of 1–10 bp, whereas some have longer microhomologies (21–39 bp) shared between *Alu* elements. These observations suggest that these complexities are caused by serial microhomology-mediated processes. Similar microhomology-mediated complexities have also been identified in polymorphic CNVs among humans. Perry and colleagues sequenced breakpoint junctions of 23 CNVs uncovered in a study using HapMap individuals; five out of 23 CNVs (22%) showed sequence complexities [12]. Interestingly, microhomologies of 2–6 bp were identified at all the breakpoints of these five complex CNVs. Furthermore, sequence

complexities have also been shown at many human–gibbon (*Nomascus leucogenys*) breakpoints of synteny [68].

Replication-based mechanisms for complex rearrangements

FoSTeS model

The observed human complex chromosomal and genomic rearrangements cannot be generated by a single recombination event (e.g. non-allelic homologous recombination [NAHR] or non-homologous end-joining [NHEJ]). Thus, a mechanism to more parsimoniously explain the multiple breakpoints involved in the complex rearrangements, the often large linear distance between them and the microhomology observed at the breakpoints must be invoked for their formation. To parsimoniously explain the sequence complexities (e.g. duplication–triplication–duplication) and microhomologies at breakpoints, a novel replication-based FoSTeS model was proposed [59].

During DNA replication, it is proposed that the DNA replication fork can stall, the lagging strand disengages from the original template and invades another replication fork and restarts DNA synthesis on the new fork by priming it via the microhomology between the switched template site and the original fork [59]. The new template strand is not necessarily adjacent to the original replication fork in primary sequence, but probably in 3D physical proximity. It could be on the same chromosome, the homologous chromosome or a non-homologous different chromosome; the proximity in 3D space potentially reflecting a replication factory [69]. Upon annealing, the transferred strand primes its own template-driven extension at the transferred fork. Depending on the direction of the fork progression and if the lagging or leading strand in the new fork was invaded and copied, the erroneously incorporated fragment from the new replication fork could be in direct or inverted orientation with respect to its original position. Furthermore, depending on if the new fork is located downstream or upstream of the original fork, the template switching results in either a deletion or a duplication. This procedure of disengaging, invading and synthesizing can occur multiple times in series (FoSTeS $\times$ 2, FoSTeS $\times$ 3, etc.), potentially reflecting the poor processivity of the DNA polymerase used and resulting in the observed complex rearrangements.

According to the FoSTeS model [59], two ‘hallmarks’ can be identified for FoSTeS-mediated rearrangements: (i) sequence complexity that results from serial template-driven juxtaposition of sequences from different genomic locations; and (ii) DNA sequence microhomologies at the join points. From the phenomenology observed in human disease-associated complex rearrangements, one could not determine if a single-stranded or double-stranded DNA was the intermediate involved in the invasion during template switching. Nor could it be determined if a stalled fork or broken (collapsed) fork was the initiating event.

A recent experimental study of stress-induced CNVs in normal human cells also revealed structural complexities to the DNA rearrangements and microhomologies at most sequenced breakpoints in the CNVs induced by treatment with aphidicolin, which can directly inhibit DNA polymerase and cause replication fork stalling [70]. This suggests the potential involvement of FoSTeS in CNV formation in mitotic human cells.

Based on the aforementioned experimental evidence from human subjects, the FoSTeS mechanism not only underlies complex genomic rearrangements involving large segments of megabases in size, but also can result in small rearrangements involving only one gene or even just a single exon. This implicates FoSTeS in gene duplication and potentially in exon shuffling, both important to human gene and genome evolution [67,71,72].

Related replication-based rearrangement mechanisms

Complex rearrangements are not limited to human subjects. Actually, the FoSTeS mechanism was proposed based on the observations of the stress-induced gene amplification by template-switching that occurs during DNA replication in *E. coli* [73].

Break-induced replication (BIR), a double-strand break (DSB) repair model of replication restarting at broken forks that has been developed from experimental observations in yeast [74–76], might also be able to cause human complex rearrangement associated with hemophilia A [77]. The classic BIR model is a homologous recombination-based mechanism, which is dependent upon recombination proteins and requires long homology for template invasion [74–76] (Table 4). The microhomologies identified at the breakpoints of human complex rearrangements are not of sufficient length and cannot be employed in classic BIR. Recently, an investigation into the formation of segmental duplications in yeast reveals a genetic requirement for Pol32, a replicative polymerase subunit, whereas the recombination proteins tested therein (Rad52, Dnl4, Rad1) are not required [78]. A template-switching model of microhomology or microsatellite-induced replication (MMIR) was proposed for DSB repair in this yeast study [78] (Table 4), supporting the contention that a replication-based mechanism can be involved in CNV formation.

Based upon the evidence regarding genomic rearrangements in human, bacteria and yeast, a generalized, replicative, template-switch model has been proposed to explain structural variations from all domains of life; the microhomology-mediated break-induced replication (MMBIR) model [79]. Both FoSTeS and MMBIR are consistent with the features of complexity and microhomology at the junctions that are identified in most human complex rearrangements and alternative to some extent; although MMBIR is break-induced (i.e. generated by a collapsed fork), FoSTeS is initiated by replication fork stalling (i.e. no DSB is required). However, there is currently no experimental data that determines whether the complex rearrangements causing human diseases are triggered by DSB, a collapsed replication fork, or a single-strand DNA transfer. FoSTeS was proposed based upon the observed phenomena in humans, whereas MMBIR provides mechanistic details from experimental observations in model organisms [59,79]. We propose that FoSTeS and/or MMBIR can be responsible for human complex rearrangements.

Other related replication-based models, for example ‘serial replication slippage (SRS)’ [80] and ‘strand misalignment–realignment’ [81], have also been proposed for complex rearrangements. In the SRS model, the primer strand can serially disassociate with template strand, reanneal by misalignment and restart DNA replication to cause human complex rearrangements [80,82,83]. Because the hypothesized slippages take place within replication forks, SRS might not be able to cause large genomic rearrangements involving hundreds of kb or longer.

The replication-based mechanisms can readily explain complex rearrangements and the microhomologies at breakpoints. However, the breakpoint sequence analyses of complex rearrangements also show absence of microhomology in some cases, which could be interpreted as NHEJ-specific nucleotide insertion or deletion [84]. Such observations suggest that the involvement of NHEJ in one rearrangement step of complex rearrangements or the entire complexity cannot be excluded [85]. However, because the NHEJ-specific nucleotide insertion at breakpoints are short (1–4 nt [86]), longer nucleotide insertions at breakpoints might not be caused by NHEJ, but by template-switching from other genomic loci via replication mechanisms.

Simple rearrangement that can be explained by FoSTeS and/or MMBIR

During the replication-based FoSTeS and/or MMBIR mechanism, template switching can occur in different ways and cause diverse chromosomal structural changes [79]. In contrast to the complex rearrangements, a simple rearrangement can be generated if the template switching occurs only once during DNA replication. Such FoSTeS and/or MMBIR-mediated simple rearrangements share some features (simple rearrangement pattern and microhomology at breakpoint) with those that are generated by NHEJ or microhomology-mediated end-joining (MMEJ) but do not have an ‘information scar’ [59,86,87]. Therefore, these rearrangements can be either explained by NHEJ, MMEJ or a single template-switching event. However, it is presently impossible to distinguish which exact mechanism, recombination-based NHEJ or MMEJ, or a replication-based FoSTeS or MMBIR, actually acts in each of these cases. In deletion cases, the analysis can be even more complicated because of the inability to assay the genomic regions lost during the rearrangement process. For example, there is still another interpretation for the simple rearrangement; that is, multiple FoSTeS or MMBIR events can also lead to a simple deletion CNV if the last template switching occurs to a template ahead of the fork and removes all evidence for sequence complexity.

In cases in which *Alu* elements were identified at both of the breakpoint ends, the formation mechanism has been interpreted as an *Alu–Alu*-mediated NAHR [87–91]. However, for NAHR to take place, there must be segments of a minimal length sharing extremely high similarity or identity between the homologous segments, termed minimal efficient processing segments (MEPS) [92]. In fact, the breakpoint sequencing of complex rearrangements reveals that many of the *Alu–Alu* rearrangements do not provide the MEPS requirements for homologous recombination via NAHR, especially those involving *Alu* from different subfamilies. Thus, if they were involved as substrates in a homologous recombination reaction, any Holliday structure that formed might be broken down by the mismatch repair apparatus. Bi *et al.* [66] reported two *AluSg* elements of 79% identity at the breakpoint junctions of one patient carrying a non-recurrent *PAFAH1B1* tandem duplication. Matejas *et al.* [93] reported a non-recurrent deletion involving *PMP22* at 17p11.2 region on which an *AluYb* and an *AluSq* with 86% of similarity were found at the breakpoint junctions. The microhomology was restricted to 27 and 29 bp in each case.

Interestingly, Bauters *et al.* [61] reported sequencing of three MECP2 duplicated breakpoint junctions in three patients, two of which were complex rearrangements and presented *Alu–*

Alu at the juxtaposition regions; one patient had an *AluY–AluY* with 39-bp microhomology and another *AluJo–AluSg* with 23-bp microhomologies. In the aggregate, these data indicate that NAHR might not be the mechanism underlying many of the complex rearrangements showing *Alu–Alu* switching at one of the breakpoints. Instead, the long microhomology (~20–40 bp) shared between *Alu* elements can potentially facilitate template switching and prime DNA replication to a new fork during DNA replication. Therefore, we propose that template switching between *Alu* elements can also cause simple rearrangements. It is noteworthy that a prevalence of *Alu* elements has been observed at the ends of LCRs [88,94] and at breakpoints of the subunits of complex LCRs [95]; therefore, FoSTeS and/or MMBIR involving *Alu* elements might also be responsible for LCR or SD formation [79].

In summary, in spite of simple CNVs that can be caused by *Alu*-mediated NAHR and NHEJ or MMEJ, the potential involvement of a replication-based template-switching mechanism in these cases cannot be excluded.

Breakpoint grouping and the influence of genomic architecture on complex rearrangement

FoSTeS and/or MMBIR-mediated complex rearrangements at the loci of *PLP1* [59] and *MECP2* [62] preferentially occur within or adjacent to complex genomic architecture that undergoes frequent polymorphic variation. Obvious breakpoint grouping was revealed at both loci (Figure 3). Monte Carlo simulation, which was performed to evaluate the breakpoint distribution of the *MECP2* rearrangements, strongly supports the contention that the segments of *MECP2* rearrangements are non-randomly distributed and that breakpoints are grouping around the shifted average breakpoint 153.0888 Mb, coincident with the LCR-laden region [62] (Figure 3b).

It is possible that the complex architecture flanking both *MECP2* and *PLP1* genes can form cruciforms or other non-B DNA structures [96,97], which can potentially cause fork stalling. Moreover, such structures could lead to single-stranded DNA regions that could produce collapsed DNA replication forks and generate one-ended double-stranded DNA, the presumed substrates for MMBIR [79]. Alternatively, in some cases, such structures can approximate participating DNA replication forks that are otherwise separated by sizeable linear distances and provide single-stranded regions for template switching.

Concluding remarks

Various replication-based molecular mechanisms, including FoSTeS and/or MMBIR, can be involved in complex genomic rearrangements, most of which can be consistent with the features of complexity and microhomology at breakpoints. CCR might also occur by replication-based mechanisms given their genomic complexities and preferential paternal origin; that is, there are many additional mitoses in sperm versus egg formation. However, it remains a technical challenge to determine which exact mechanism is responsible for selected disease-associated rearrangements because it is not feasible to distinguish between replication-based and recombination-based or between DSB-induced and fork-stalling-caused processes in some human subjects. Notably, accumulating evidence reveals that

many CNVs represent complex rearrangements and, thus, can be more parsimoniously explained by replication-based microhomology-mediated mechanisms, including FoSTeS and/or MMBIR. The complex nature of many LCRs or SDs and the conveyed clinical phenotypes for *de novo* complex rearrangements implicate replication-based mechanism in human genome and gene evolution and, thus, in human health and disease.

Acknowledgments

Our work has been sponsored by the National Institute of Neurological Disorders and Stroke (NINDS, NIH), the March of Dimes, the Muscular Dystrophy Association, the Foundation Fighting Blindness and the Charcot-Marie-Tooth Association. We apologize, but take full responsibility for, omissions of citations owing to space limitations.

Glossary

Copy number variation (CNV)	a DNA segment that is present at a variable copy number in comparison with a reference genome, for example, deletion and duplication [98]
Non-recurrent rearrangement	occurs at a single locus, can vary in size and have different breakpoints. This can be caused by several mechanisms, including NHEJ [84], retrotransposition [7,10,99] and NAHR between repetitive sequences such as L1 [100] and <i>Alu</i> elements [88,90,91]
Recurrent rearrangement	occurs in multiple unrelated individuals, shares a common rearrangement interval and size, and has breakpoints that cluster. This is caused by NAHR between relatively large (>10 kb) DNA repeats, that is, LCRs or segmental duplications (SDs) [31,101]
Structural variation (SV)	includes CNV and other balanced variations such as inversion and translocation. Structural variations are caused by genome rearrangements, which can be categorized into two major groups:: recurrent or non-recurrent rearrangements

References

1. The International HapMap Consortium. A haplotype map of the human genome. *Nature*. 2005; 437:1299–1320. [PubMed: 16255080]
2. The International HapMap Consortium. A second generation human haplotype map of over 3.1 million SNPs. *Nature*. 2007; 449:851–861. [PubMed: 17943122]
3. Iafrate AJ, et al. Detection of large-scale variation in the human genome. *Nat Genet*. 2004; 36:949–951. [PubMed: 15286789]
4. Sebat J, et al. Large-scale copy number polymorphism in the human genome. *Science*. 2004; 305:525–528. [PubMed: 15273396]
5. Conrad DF, et al. A high-resolution survey of deletion polymorphism in the human genome. *Nat Genet*. 2006; 38:75–81. [PubMed: 16327808]
6. Redon R, et al. Global variation in copy number in the human genome. *Nature*. 2006; 444:444–454. [PubMed: 17122850]
7. Korb J, et al. Paired-end mapping reveals extensive structural variation in the human genome. *Science*. 2007; 318:420–426. [PubMed: 17901297]
8. Levy S, et al. The diploid genome sequence of an individual human. *PLoS Biol*. 2007; 5:e254. [PubMed: 17803354]

9. Cooper GM, et al. Systematic assessment of copy number variant detection via genome-wide SNP genotyping. *Nat Genet.* 2008; 40:1199–1203. [PubMed: 18776910]
10. Kidd JM, et al. Mapping and sequencing of structural variation from eight human genomes. *Nature.* 2008; 453:56–64. [PubMed: 18451855]
11. McCarroll SA, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet.* 2008; 40:1166–1174. [PubMed: 18776908]
12. Perry GH, et al. The fine-scale and complex architecture of human copy-number variation. *Am J Hum Genet.* 2008; 82:685–695. [PubMed: 18304495]
13. Wheeler DA, et al. The complete genome of an individual by massively parallel DNA sequencing. *Nature.* 2008; 452:872–876. [PubMed: 18421352]
14. Lupski JR, Stankiewicz P. Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet.* 2005; 1:e49. [PubMed: 16444292]
15. Hurler ME, et al. The functional impact of structural variation in humans. *Trends Genet.* 2008; 24:238–245. [PubMed: 18378036]
16. Nathans J, et al. Molecular genetics of inherited variation in human color vision. *Science.* 1986; 232:203–210. [PubMed: 3485310]
17. Lupski JR.; Chance, PF. Hereditary motor and sensory neuropathies involving altered dosage or mutation of PMP22: The CMT1A duplication and HNPP deletion. In: Dyck, PJ.; Thomas, PK., editors. *Peripheral Neuropathy.* Elsevier Science; 2005. p. 1659-1680.
18. Lee JA, Lupski JR. Genomic rearrangements and gene copy-number alterations as a cause of nervous system disorders. *Neuron.* 2006; 52:103–121. [PubMed: 17015230]
19. Lupski JR. Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet.* 1998; 14:417–422. [PubMed: 9820031]
20. Lupski JR. Genomic disorders ten years on. *Genome Med.* 2009; 1:42. [PubMed: 19439022]
21. Gonzalez E, et al. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science.* 2005; 307:1434–1440. [PubMed: 15637236]
22. Sebat J, et al. Strong association of *de novo* copy number mutations with autism. *Science.* 2007; 316:445–449. [PubMed: 17363630]
23. Weiss LA, et al. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med.* 2008; 358:667–675. [PubMed: 18184952]
24. Lupski JR. Schizophrenia: Incriminating genomic evidence. *Nature.* 2008; 455:178–179. [PubMed: 18784712]
25. Stefansson H, et al. Large recurrent microdeletions associated with schizophrenia. *Nature.* 2008; 455:232–236. [PubMed: 18668039]
26. The International Schizophrenia Consortium. Rare chromosomal deletions and duplications increase risk of schizophrenia. *Nature.* 2008; 455:237–241. [PubMed: 18668038]
27. Stankiewicz P, Beaudet AL. Use of array CGH in the evaluation of dysmorphology, malformations, developmental delay, and idiopathic mental retardation. *Curr Opin Genet Dev.* 2007; 17:182–192. [PubMed: 17467974]
28. Froyen G, et al. Submicroscopic duplications of the hydroxysteroid dehydrogenase HSD17B10 and the E3 ubiquitin ligase HUWE1 are associated with mental retardation. *Am J Hum Genet.* 2008; 82:432–443. [PubMed: 18252223]
29. Sharp AJ, et al. A recurrent 15q13.3 microdeletion syndrome associated with mental retardation and seizures. *Nat Genet.* 2008; 40:322–328. [PubMed: 18278044]
30. Zhang F, et al. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet.* (in press).
31. Gu W, et al. Mechanisms for human genomic rearrangements. *PathoGenetics.* 2008; 1:4. [PubMed: 19014668]
32. Pai GS, et al. Complex chromosome rearrangements. Report of a new case and literature review. *Clin Genet.* 1980; 18:436–444. [PubMed: 7449183]
33. Warburton D. *De novo* balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. *Am J Hum Genet.* 1991; 49:995–1013. [PubMed: 1928105]

34. Batanian JR, Eswara MS. *De novo* apparently balanced complex chromosome rearrangement (CCR) involving chromosomes 4, 18 and 21 in a girl with mental retardation: report and review. *Am J Med Genet.* 1998; 78:44–51. [PubMed: 9637422]
35. Astbury C, et al. Delineation of complex chromosomal rearrangements: evidence for increased complexity. *Hum Genet.* 2004; 114:448–457. [PubMed: 14767757]
36. Vermeulen S, et al. Molecular cytogenetic analysis of complex chromosomal rearrangements in patients with mental retardation and congenital malformations: delineation of 7q21.11 breakpoints. *Am J Med Genet A.* 2004; 124A:10–18. [PubMed: 14679581]
37. Lee IW, et al. Constitutional complex chromosomal rearrangements in azoospermic men—case report and literature review. *Urology.* 2006; 68:1343.e5–8. [PubMed: 17169659]
38. Giardino D, et al. *De novo* balanced chromosome rearrangements in prenatal diagnosis. *Prenat Diagn.* 2009; 29:257–265. [PubMed: 19248039]
39. Batista DA, et al. Molecular analysis of a complex chromosomal rearrangement and a review of familial cases. *Am J Med Genet.* 1994; 53:255–263. [PubMed: 7856662]
40. De Gregori M, et al. Cryptic deletions are a common finding in “balanced” reciprocal and complex chromosome rearrangements: a study of 59 patients. *J Med Genet.* 2007; 44:750–762. [PubMed: 17766364]
41. Baptista J, et al. Breakpoint mapping and array CGH in translocations: comparison of a phenotypically normal and an abnormal cohort. *Am J Hum Genet.* 2008; 82:927–936. [PubMed: 18371933]
42. Madan K, et al. Recombination in a balanced complex translocation of a mother leading to a balanced reciprocal translocation in the child. Review of 60 cases of balanced complex translocations. *Hum Genet.* 1997; 99:806–815. [PubMed: 9187678]
43. Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet.* 2005; 76:8–32. [PubMed: 15549674]
44. Schrock E, et al. Multicolor spectral karyotyping of human chromosomes. *Science.* 1996; 273:494–497. [PubMed: 8662537]
45. Speicher MR, et al. Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat Genet.* 1996; 12:368–375. [PubMed: 8630489]
46. Karadeniz N, et al. Further delineation of complex chromosomal rearrangements in fertile male using multicolor banding. *Mol Cytogenet.* 2008; 1:17. [PubMed: 18687140]
47. Vissers LE, et al. Complex chromosome 17p rearrangements associated with low-copy repeats in two patients with congenital anomalies. *Hum Genet.* 2007; 121:697–709. [PubMed: 17457615]
48. Surace C, et al. Masked complex chromosome rearrangement in a child thought to have del(8qter) as the sole cytogenetic abnormality. *Cytogenet. Genome Res.* 2008; 121:215–221. [PubMed: 18758162]
49. Haj R, et al. Identification of a rare *de novo* three-way complex t(5;20;8)(q31;p11.2;p21) with microdeletions on 5q31.2, 5q31.3, and 8p23.2 in a patient with hearing loss and global developmental delay: case report. *Mol Cytogenet.* 2009; 2:2. [PubMed: 19128483]
50. Sismani C, et al. Cryptic genomic imbalances in patients with *de novo* or familial apparently balanced translocations and abnormal phenotype. *Mol Cytogenet.* 2008; 1:15. [PubMed: 18644119]
51. Giardino D, et al. Prenatal diagnosis of a *de novo* complex chromosome rearrangement (CCR) mediated by six breakpoints, and a review of 20 prenatally ascertained CCRs. *Prenat Diagn.* 2006; 26:565–570. [PubMed: 16683274]
52. Shetty S, et al. Cytogenetic and molecular characterization of a *de-novo* cryptic deletion of 7p21 associated with an apparently balanced translocation and complex craniosynostosis. *Clin Dysmorphol.* 2007; 16:253–256. [PubMed: 17786117]
53. Jiang YH, et al. *De novo* and complex imbalanced chromosomal rearrangements revealed by array CGH in a patient with an abnormal phenotype and apparently “balanced” paracentric inversion of 14(q21q23). *Am J Med Genet A.* 2008; 146A:1986–1993. [PubMed: 18627051]
54. Ballarati L, et al. Cytogenetic, FISH and array-CGH characterization of a complex chromosomal rearrangement carried by a mentally and language impaired patient. *Eur J Med Genet.* 2009; 10.1016/j.ejmg.2009.02.004

55. Zhang L, et al. Patterns of segmental duplication in the human genome. *Mol Biol Evol.* 2005; 22:135–141. [PubMed: 15371527]
56. Li P, et al. Karyotype-phenotype insights from 11q14.1-q23. 2 interstitial deletions: FZD4 haploinsufficiency and exudative vitreoretinopathy in a patient with a complex chromosome rearrangement. *Am J Med Genet A.* 2006; 140:2721–2729. [PubMed: 17103440]
57. Lindstrand A, et al. Molecular cytogenetic characterization of a constitutional, highly complex intrachromosomal rearrangement of chromosome 1, with 14 breakpoints and a 0.5 Mb submicroscopic deletion. *Am J Med Genet A.* 2008; 146A:3217–3222. [PubMed: 19006217]
58. Patsalis PC. Complex chromosomal rearrangements. *Genet Couns.* 2007; 18:57–69. [PubMed: 17515301]
59. Lee JA, et al. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell.* 2007; 131:1235–1247. [PubMed: 18160035]
60. Wolf NI, et al. Three or more copies of the proteolipid protein gene PLP1 cause severe Pelizaeus-Merzbacher disease. *Brain.* 2005; 128:743–751. [PubMed: 15689360]
61. Bauters M, et al. Nonrecurrent MECP2 duplications mediated by genomic architecture-driven DNA breaks and break-induced replication repair. *Genome Res.* 2008; 18:847–858. [PubMed: 18385275]
62. Carvalho CM, et al. Complex rearrangements in patients with duplications of MECP2 can occur by fork stalling and template switching. *Hum Mol Genet.* 2009; 18:2188–2203. [PubMed: 19324899]
63. Small K, et al. Emerin deletion reveals a common X-chromosome inversion mediated by inverted repeats. *Nat Genet.* 1997; 16:96–99. [PubMed: 9140403]
64. Muhle C, et al. Recurrent inversion with concomitant deletion and insertion events in the coagulation factor VIII gene suggests a new mechanism for X-chromosomal rearrangements causing hemophilia. *A Hum Mutat.* 2007; 28:1045.
65. Madden HR, et al. Characterization of a complex Duchenne muscular dystrophy-causing dystrophin gene inversion and restoration of the reading frame by induced exon skipping. *Hum Mutat.* 2009; 30:22–28. [PubMed: 18570328]
66. Bi W, et al. Increased LIS1 expression affects human and mouse brain development. *Nat Genet.* 2009; 41:168–177. [PubMed: 19136950]
67. Zhang F, et al. The DNA replication FoSTeS/MMBIR mechanism can generate human genomic, genic, and exonic complex rearrangements. *Nat Genet.* (in press).
68. Girirajan S, et al. Sequencing human-gibbon breakpoints of synteny reveals mosaic new insertions at rearrangement sites. *Genome Res.* 2009; 19:178–190. [PubMed: 19029537]
69. Kitamura E, et al. Live-cell imaging reveals replication of individual replicons in eukaryotic replication factories. *Cell.* 2006; 125:1297–1308. [PubMed: 16814716]
70. Arlt MF, et al. Replication stress induces genome-wide copy number changes in human cells that resemble polymorphic and pathogenic variants. *Am J Hum Genet.* 2009; 84:339–350. [PubMed: 19232554]
71. Ohno, S. *Evolution by Gene Duplication.* Springer-Verlag; 1970.
72. Gilbert W. Why genes in pieces? *Nature.* 1978; 271:501. [PubMed: 622185]
73. Slack A, et al. On the mechanism of gene amplification induced under stress in *Escherichia coli*. *PLoS Genet.* 2006; 2:e48. [PubMed: 16604155]
74. Morrow DM, et al. “Break copy” duplication: a model for chromosome fragment formation in *Saccharomyces cerevisiae*. *Genetics.* 1997; 147:371–382. [PubMed: 9335579]
75. Lydeard JR, et al. Break-induced replication and telomerase-independent telomere maintenance require Pol32. *Nature.* 2007; 448:820–823. [PubMed: 17671506]
76. Smith CE, et al. Template switching during break-induced replication. *Nature.* 2007; 447:102–105. [PubMed: 17410126]
77. Sheen CR, et al. Double complex mutations involving F8 and FUNDC2 caused by distinct break-induced replication. *Hum Mutat.* 2007; 28:1198–1206. [PubMed: 17683067]
78. Payen C, et al. Segmental duplications arise from Pol32-dependent repair of broken forks through two alternative replication-based mechanisms. *PLoS Genet.* 2008; 4:e1000175. [PubMed: 18773114]

79. Hastings PJ, et al. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.* 2009; 5:e1000327. [PubMed: 19180184]
80. Chen JM, et al. Complex gene rearrangements caused by serial replication slippage. *Hum Mutat.* 2005; 26:125–134. [PubMed: 15977178]
81. Gordon AJ, Halliday JA. Inversions with deletions and duplications. *Genetics.* 1995; 140:411–414. [PubMed: 7635304]
82. Chen JM, et al. Intrachromosomal serial replication slippage in trans gives rise to diverse genomic rearrangements involving inversions. *Hum Mutat.* 2005; 26:362–373. [PubMed: 16110485]
83. Chen JM, et al. Meta-analysis of gross insertions causing human genetic disease: novel mutational mechanisms and the role of replication slippage. *Hum Mutat.* 2005; 25:207–221. [PubMed: 15643617]
84. Lieber MR. The mechanism of human nonhomologous DNA end joining. *J Biol Chem.* 2008; 283:1–5. [PubMed: 17999957]
85. Gajecka M, et al. Unexpected complexity at breakpoint junctions in phenotypically normal individuals and mechanisms involved in generating balanced translocations t(1;22)(p36;q13). *Genome Res.* 2008; 18:1733–1742. [PubMed: 18765821]
86. McVey M, Lee SE. MMEJ repair of double-strand breaks (director's cut): deleted sequences and alternative endings. *Trends Genet.* 2008; 24:529–538. [PubMed: 18809224]
87. Shaw CJ, Lupski JR. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. *Hum Mol Genet.* 2004; 13:R57–R64. [PubMed: 14764619]
88. Bailey JA, et al. An *Alu* transposition model for the origin and expansion of human segmental duplications. *Am J Hum Genet.* 2003; 73:823–834. [PubMed: 14505274]
89. Shaw CJ, Lupski JR. Non-recurrent 17p11.2 deletions are generated by homologous and non-homologous mechanisms. *Hum Genet.* 2005; 116:1–7. [PubMed: 15526218]
90. Sen SK, et al. Human genomic deletions mediated by recombination between *Alu* elements. *Am J Hum Genet.* 2006; 79:41–53. [PubMed: 16773564]
91. Han K, et al. *Alu* recombination-mediated structural deletions in the chimpanzee genome. *PLoS Genet.* 2007; 3:1939–1949. [PubMed: 17953488]
92. Waldman AS, Liskay RM. Dependence of intrachromosomal recombination in mammalian cells on uninterrupted homology. *Mol Cell Biol.* 1988; 8:5350–5357. [PubMed: 2854196]
93. Matejas V, et al. Identification of *Alu* elements mediating a partial PMP22 deletion. *Neurogenetics.* 2006; 7:119–126. [PubMed: 16570190]
94. Kim PM, et al. Analysis of copy number variants and segmental duplications in the human genome: Evidence for a change in the process of formation in recent evolutionary history. *Genome Res.* 2008; 18:1865–1874. [PubMed: 18842824]
95. Babcock M, et al. Shuffling of genes within low-copy repeats on 22q11 (LCR22) by *Alu*-mediated recombination events during evolution. *Genome Res.* 2003; 13:2519–2532. [PubMed: 14656960]
96. Wells RD. Non-B DNA conformations, mutagenesis and disease. *Trends Biochem Sci.* 2007; 32:271–278. [PubMed: 17493823]
97. Voineagu I, et al. Replication stalling at unstable inverted repeats: interplay between DNA hairpins and fork stabilizing proteins. *Proc Natl Acad Sci U S A.* 2008; 105:9936–9941. [PubMed: 18632578]
98. Feuk L, et al. Structural variation in the human genome. *Nat Rev Genet.* 2006; 7:85–97. [PubMed: 16418744]
99. Xing J, et al. Mobile elements create structural variation: analysis of a complete human genome. *Genome Res.* 2009; 19:1101/gr.091827.109
100. Han K, et al. L1 recombination-associated deletions generate human genomic variation. *Proc Natl Acad Sci U S A.* 2008; 105:19366–19371. [PubMed: 19036926]
101. Stankiewicz P, Lupski JR. Genome architecture, rearrangements and genomic disorders. *Trends Genet.* 2002; 18:74–82. [PubMed: 11818139]

102. Cazeneuve C, et al. A new complex homozygous large rearrangement of the PINK1 gene in a Sudanese family with early onset Parkinson's disease. *Neurogenetics*. 2009;10:1007/s10048-009-0174-4
103. Chen JM. The 10-Mb paracentric inversion of chromosome arm 2p in activating MSH2 and causing hereditary nonpolyposis colorectal cancer: re-annotation and mutational mechanisms. *Genes Chromosomes Cancer*. 2008; 47:543–545. [PubMed: 18335504]
104. McVety S, et al. Novel genomic insertion–deletion in *MLH1*: possible mechanistic role for non-homologous end-joining DNA repair. *Clin Genet*. 2005; 68:234–238. [PubMed: 16098012]
105. Vervoort VS, et al. Genomic rearrangements of *EYAI* account for a large fraction of families with BOR syndrome. *Eur J Hum Genet*. 2002; 10:757–766. [PubMed: 12404110]
106. Demura M, et al. Regional rearrangements in chromosome 15q21 cause formation of cryptic promoters for the *CYP19* (aromatase) gene. *Hum Mol Genet*. 2007; 16:2529–2541. [PubMed: 17584767]
107. Rugless MJ, et al. A large deletion in the human alpha-globin cluster caused by a replication error is associated with an unexpectedly mild phenotype. *Hum Mol Genet*. 2008; 17:3084–3093. [PubMed: 18632685]
108. Arnoldi A, et al. A clinical, genetic, and biochemical characterization of SPG7 mutations in a large cohort of patients with hereditary spastic paraplegia. *Hum Mutat*. 2008; 29:522–531. [PubMed: 18200586]
109. Inoue K, et al. Genomic rearrangements resulting in PLP1 deletion occur by nonhomologous end joining and cause different dysmyelinating phenotypes in males and females. *Am J Hum Genet*. 2002; 71:838–853. [PubMed: 12297985]
110. Yatsenko SA, et al. Molecular mechanisms for subtelomeric rearrangements associated with the 9q34.3 microdeletion syndrome. *Hum Mol Genet*. 2009; 18:1924–1936. [PubMed: 19293338]

Figure 1.

Frequency of specific chromosomes involved in CCR. The *x*-axis lists specific human chromosomes and the *y*-axis shows the percentage of times a particular chromosome is involved in a complex rearrangement. Data compiled on the basis of 226 CCRs.

Figure 2.

Resolutions of genomic complexities. Resolving complex genomic rearrangements can require different genome analysis tools. **(a)** Complex rearrangements of the short arm of human chromosome 17 revealed by FISH analysis (adapted, with permission, from Ref. [47]). **(b)** Complex rearrangement at 1q revealed by BAC aCGH. A BAC array is shown on the left with the chromosome numbers depicted above. The y-axis shows log ratios of patient versus control BAC aCGH with green dots showing gain and red dots showing loss. To the right is a depiction of the complex rearrangement. **(c)** The additional small duplication of a potential complex rearrangement (duplication–normal–duplication) at 17p11.2 was confirmed by oligonucleotide aCGH. Copy number gains were shown in red. See Ref. [67] for an alternative explanation of these data. **(d)** FoSTeS-mediated complex *PLP1* duplication was revealed by breakpoint sequencing in the apparent ‘simple’ rearrangement cases without detectable complexity by oligonucleotide aCGH (adapted, with permission, from Ref. [59]). Abbreviations: del, deletion; DistRef, distal reference sequence; dup,

duplication; nml, normal; ProxRef, proximal reference sequence; 1264F, forward sequence of Subject 1264.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1

Frequency of breakpoints detected for CCRs in 226 patients

Number of breakpoints	Frequency for CCRs
3	70
4	67
5	36
6	22
7	17
8	10
9	8
10	2
11	0
12	2
13	0
14	1
15	1

Table 2

Examples of complex genomic rearrangements

Chromosome or gene involved	Events ^a	Sequenced breakpoint	Microhomology	Size (bp) of microhomology	SINE or LINE	Refs
1p36.12/ <i>PINK1, DDOST</i>	del + ins	2	100%	3	–	[102]
2p21/ <i>MSH2</i>	dup + inv + del	3	100%	2	–	[103]
3p22.2/ <i>MLH1</i>	del + ins + nml + ins + del	2	50%	1	–	[104]
8q13.3/ <i>EYAI</i>	del + ins	2	0%	–	–	[105]
	dup + inv + del	2	100%	1–8		
15q21/ <i>CYP19</i>	del + inv + del	2	50%	1–3	–	[106]
	del + inv	2	100%			
	del + inv + del + ins	2	50%			
	ins + inv	2	–			
16p13.3/ α -globin cluster	del + inv	2	100%	4–8	–	[107]
16q24.3/ <i>SPG7</i>	del + ins + del	2	50%	10	1 <i>Alu–Alu</i>	[108]
Xp21.2/ <i>DMD</i>	inv + dup + del	2	50%	2	–	[65]
Xq28/ <i>EMD, FLN1</i>	del + dup + inv	1	100%	2	–	[63]
Xq28/ <i>F8, MPPI</i>	inv + del	1	0%	–	–	[64]
Xq28/ <i>MECP2, L1CAM, CLIC2</i>	dup + ins	3	100%	2–39	3 <i>Alu–Alu</i>	[61]
	dup + ins + inv	3				

^a Abbreviations: del, deletion; dup, duplication; ins, insertion; inv, inversion; nml, normal.

Table 3

Complex genomic rearrangements identified by aCGH and/or breakpoint sequencing

Disease	Gene	% complex in non-recurrent rearrangements ^d	Sequenced breakpoints	Microhomology	Size (bp) of microhomology	Refs
PMD [MIM 312080]	<i>PLP1</i>	65%	6	100%	2–5	[59,109]
dup <i>MECP2</i>	<i>MECP2</i>	27%	2	100%	2	[62]
del(9)(q34.3)	<i>EHMT1</i>	7%	1	100%	3	[110]
dup <i>LIS1</i>	<i>PAFAH1B1</i>	29%	4	100%	2–21	[66]
PTLS [MIM 610883]	<i>RAI1</i>	57%	7	86%	2–31	[67]
CMT1A [MIM 118220] or HNPP [MIM 162500]	<i>PMP22</i>	29%	3	100%	3–5	[67]

^dPercentages represent the lower estimates because in the case of each locus specific breakpoint sequencing was not achieved for several of the rearrangements.

Table 4

Several mechanisms potentially involved in complex rearrangement formation

Mechanism	Cause	DSB	Process	Breakpoint feature	Refs
FoSTeS	Stalled fork	Not required	Template switching between forks	Microhomology	[59]
MMBIR	Collapsed or broken fork	One-ended break	Template switching between forks	Microhomology	[79]
BIR	Broken fork	One-ended break	Homologous recombination-based template invasion between forks	Homology	[74–76]
MMIR	Broken fork	One-ended break	Rad52-independent template switching between forks	Microhomology or microsatellite	[78]
SRS	Slipped strand mispairing	Not required	Serial replication slippage within forks	Microhomology	[80,83]
Multiple NHEJ	DSB	> 1 break	Multistep recombination repair of DSB ends	Homology not required	[84,85]