Review

CellPress

Human structural variation: mechanisms of chromosome rearrangements

Brooke Weckselblatt and M. Katharine Rudd

Department of Human Genetics, Emory University School of Medicine, Atlanta, GA 30322, USA

Chromosome structural variation (SV) is a normal part of variation in the human genome, but some classes of SV can cause neurodevelopmental disorders. Analysis of the DNA sequence at SV breakpoints can reveal mutational mechanisms and risk factors for chromosome rearrangement. Large-scale SV breakpoint studies have become possible recently owing to advances in next-generation sequencing (NGS) including whole-genome sequencing (WGS). These findings have shed light on complex forms of SV such as triplications, inverted duplications, insertional translocations, and chromothripsis. Sequence-level breakpoint data resolve SV structure and determine how genes are disrupted, fused, and/or misregulated by breakpoints. Recent improvements in breakpoint sequencing have also revealed non-allelic homologous recombination (NAHR) between paralogous long interspersed nuclear element (LINE) or human endogenous retrovirus (HERV) repeats as a cause of deletions, duplications, and translocations. This review covers the genomic organization of simple and complex constitutional SVs, as well as the molecular mechanisms of their formation.

Introduction

Genomic SV refers to abnormalities in chromosome structure. The first human chromosome rearrangements were observed down the microscope in cells from tumors (neoplastic SV) or blood (constitutional SV). Today, standard SV detection methods include chromosome banding, fluorescence *in situ* hybridization (FISH), and array comparative genome hybridization (CGH). Whereas array CGH detects copy-number variation (CNV) in the form of deletions and duplications, chromosome banding and FISH can also detect copy-neutral SV-like inversions and balanced translocations that do not result in changes in copy number. Recently, targeted NGS and WGS technology has been applied to detect CNV and copy-neutral SV using sequence read-depth and paired reads that span breakpoints. Although these techniques detect submicroscopic CNVs missed by traditional cytogenetics, all SV detection methods have some limitations (Box 1).

Although chromosome rearrangements are important for human health because they contribute to genetic diversity and evolution [1–4], they can also drive disease [5–7]. Approximately 15–20% of those with intellectual disability and autism spectrum disorders have a clinically relevant SV [6,8,9]. Exome sequencing studies estimate that single nucleotide variation (SNV) is responsible for another ~25% of neurodevelopmental disorders [10,11]. Constitutional SV arises in premeiotic, meiotic, or postzygotic cells, and in most cases the timing of SV formation is not known. Analysis of SV identifies genes related to disease, breakage hotspots, parent-of-origin biases, and common mutational mechanisms. Together, these data point to risk factors for SV formation and key genes responsible for genetic syndromes. DNA sequence at SV breakpoint junctions reveals signatures of diverse DNA repair mechanisms that shape human chromosome rearrangements (see Glossary). Long stretches of homologous sequence shared between breakpoints indicate NAHR, whereas the absence of sequence homology points to repair by nonhomologous end-joining (NHEJ) (Figure 1) [12,13]. Inserted or inverted sequences at breakpoints suggest DNA replication-based mechanisms, such as microhomology-mediated break-induced replication (MMBIR) or fork stalling and template switching (FoSTeS) (Figure 1) [14,15]. Sequencing breakpoints also

Glossary

Alu: a family of short interspersed nuclear elements; *Alu* elements are approximately 300 bp in length and are the most abundant class of repeats in the human genome.

Fork stalling and template switching (FoSTeS): at a stalled replication fork, the lagging strand disengages and invades a nearby replication fork, then reinitiates DNA synthesis.

Human endogenous retroviral element (HERV): derived from ancient retroviruses, HERV sequences are flanked by long terminal repeats.

Long interspersed nuclear element (LINE): retrotransposons that are ~6 kb when full-length.

Microhomology-mediated break-induced replication (MMBIR): a broken DNA strand at a collapsed replication fork uses microhomology to invade a nearby replication fork.

Non-allelic homologous recombination (NAHR): recombination between regions with high sequence similarity but different genomic positions.

Non-homologous end joining (NHEJ): following a double-strand break (DSB), broken DNA ends ligate together without a homologous template.

Segmental duplication (SD): also known as low-copy repeats, these genomic segments are at least 1 kb in length and share >90% sequence identity.

Corresponding author: Rudd, M.K. (katie.rudd@emory.edu).

Keywords: structural variation; copy-number variation; inverted duplication; translocation; triplication; chromothripsis.

0168-9525/

 http://dx.doi.org/10.1016/j.tig.2015.05.010

Box 1. Methods for SV detection

Chromosome banding: chromosomes are prepared from dividing cells, stained, and viewed with a microscope. Large deletions, duplications, and translocations are detected if the banding pattern or chromosome structure is altered. However, smaller microdeletions and microduplications are not observed.

Fluorescence *in situ* hybridization (FISH): fluorescent-labeled DNA probes hybridize to metaphase or interphase cells to visualize a locus on a chromosome and determine copy number. FISH can determine the location of chromosomal segments identified by microarray, NGS, and WGS.

Microarray: array CGH detects copy-number differences between abnormal and reference genomes. SNP arrays detect changes in copy-number and allelic ratios. CNV location and SV organization are not determined by microarray methods.

Whole-genome sequencing (WGS): sequencing the whole genome provides the most comprehensive SV analysis. Breakpoints of CNV and copy-neutral SV are detectable by paired-end reads that have discordant mappings to the reference genome. Complex genomic structures identified by WGS may require FISH or chromosome banding to place rearranged segments.

Mate-pair sequencing: whereas standard NGS methods sequence the ends of 300–500 bp DNA fragments, mate-pair or jumping libraries sequence the ends of DNA fragments that are several kb in length. These large-fragment mate-pair libraries increase the likelihood of detecting SV that has breakpoints within interspersed repeats.

has the potential to uncover more-complex genomic structures that are missed by low-resolution methods. Recently, WGS studies of pathogenic SV have revealed many genomic breakpoints in complex rearrangements that form as one catastrophic event. We review here the mechanisms and consequences of simple and complex constitutional SV.

Simple intrachromosomal SV

Simple intrachromosomal deletions, duplications, and inversions involve only one chromosome and are the product of one or two double-strand breaks (DSBs) (Figure 2A). Deletions and duplications are easily detected by array-based methods that measure differences in copy number between subject and reference genomes (Box 1). These CNVs may also be detected by measuring NGS read depth because, relative to the rest of the genome, a region with half of the coverage is inferred to be a deletion, and a region with ~50% more read depth is inferred to be a duplication [16–18].

Because inversions are copy-neutral, they escape detection by microarray and read depth methods. Recent use of mate-pair sequencing (Box 1) and fosmid/bacterial artificial chromosome (BAC) end sequencing enabled the identification of hundreds of inversion polymorphisms in the human genome [19–22]. Although most inversions are not associated with an abnormal phenotype, some alter the orientation of repetitive DNA in a way that predisposes the chromosome to rearrangement in the future. Recurrent deletions and duplications of chromosomes 5q35, 8p23.1, 16p12.1, and 17q21.31 occur via NAHR and only arise in parents with an inversion of these chromosomes. Thus, inversion carriers have an increased risk for offspring with genomic disorders [23–26].

Deletions can lie either within a chromosome arm (interstitial) or truncate the end of a chromosome (terminal) (Figure 2B) [27]. Terminal deletions have been described on almost every human chromosome end, and in some cases these CNVs result in a recognizable genomic disorder. For example, Wolf–Hirshhorn [28], Cri-du-chat [29], Kleefstra [30], Jacobsen [31], and Phelan–McDermid [32]

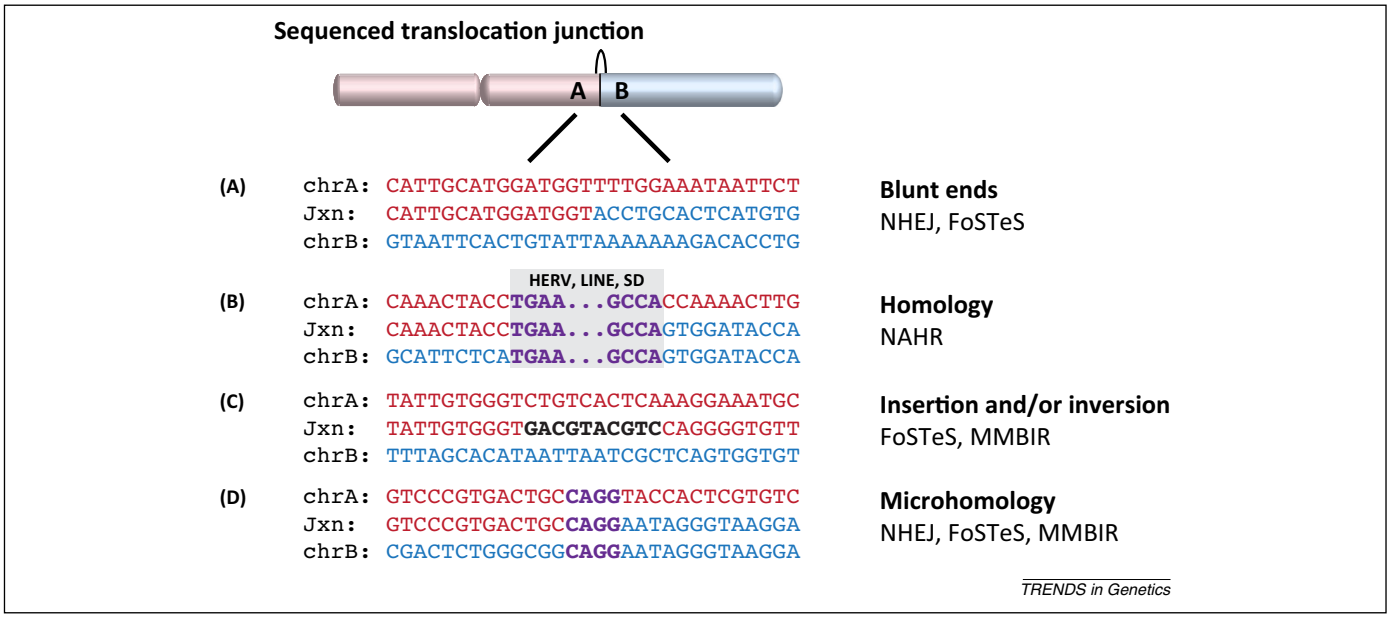

Figure 1. Signatures of mutational mechanisms. DNA sequence that spans the translocation breakpoint junction is aligned to the pink reference chromosome A (chrA) and blue reference chromosome B (chrB). The breakpoints are located where the junction (Jxn) sequence transitions from chrA to chrB. **(A)** Jxn with blunt ends at the breakpoints points to repair by non-homologous end-joining (NHEJ) or fork stalling and template switching (FoSTeS). **(B)** Homology, shown in purple, >1 kb in length and shared between chrA and chrB breakpoints suggests non-allelic homologous recombination (NAHR) between paralogous human endogenous retroviruses (HERVs), long interspersed nuclear elements (LINEs), or segmental duplications (SDs). **(C)** The presence of inverted and/or inserted sequences (shown in black) at the breakpoints are signatures of replicative mechanisms such as FoSTeS and microhomology-mediated break-induced replication (MMBIR). **(D)** 1–15 bp of microhomology between chromosome breakpoints is common and may be due to NHEJ, FoSTeS, or MMBIR.

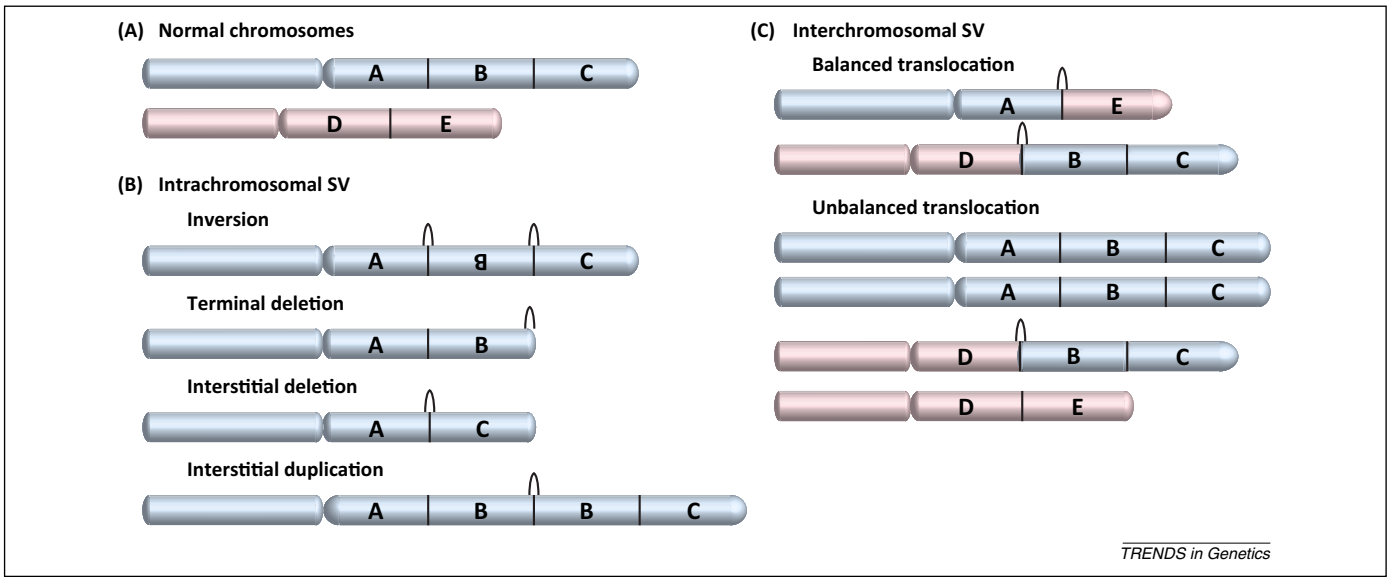

Figure 2. Simple chromosome rearrangements. **(A)** Two nonhomologous chromosomes shown in blue and pink. Segments are labeled with letters A–E. Black arches indicate structural variation (SV) breakpoint junctions. **(B)** Intrachromosomal rearrangements include inversions, interstitial and terminal deletions, and interstitial duplications. **(C)** Simple translocations between two different chromosome ends. Balanced translocations do not result in copy-number variation (CNV), but unbalanced translocations have partial monosomy (segment E) and partial trisomy (segments B,C).

syndromes are caused by terminal deletions of chromosomes 4p, 5p, 9q, 11q, and 22q, respectively. Sequence analysis of terminal deletions revealed guanine-rich motifs that are over-represented at breakpoints. This suggests that either G-rich sequences are risk factors for chromosome breakage or that, once a DSB occurs, G-rich DNA is an ideal substrate for *de novo* telomere synthesis and terminal deletion formation [27,33].

Interstitial deletions and duplications may be caused by NAHR, NHEJ, or MMBIR. The genomic organization of interstitial deletions is relatively simple, and haploinsufficiency for genes within the deleted segment can lead to abnormal outcomes. The phenotypic significance of interstitial duplications is more difficult to interpret because genes at breakpoints may or may not be disrupted depending on the orientation of the duplicated segment. Sequence analysis of a diverse collection of interstitial duplications revealed that they are almost always tandem and in direct orientation relative to the original locus (Figure 2B) [34].

Most deletion and duplication CNVs have non-recurrent breakpoints, with blunt ends or microhomology at breakpoint junctions [27,35,36]. Although this microhomology may seem coincidental, many CNV sequencing studies have revealed greater microhomology than would be expected by chance [2,34,36]. Recurrent deletions and duplications make up ~20% of pathogenic intrachromosomal rearrangements [27,37]. Genomic disorders caused by these recurrent CNVs are ideal for genotype–phenotype correlations because the same contiguous genes are deleted or duplicated in unrelated individuals [23]. The earliest recurrent deletions and duplications discovered turned out to be mediated by NAHR between segmental duplications (SDs) hundreds of kb in length on the same chromosome [12]. More recent studies have used genomic approaches to predict intrachromosomal CNVs mediated by long (>10 kb) SDs with high sequence identity (>95%) [38–40]. NAHR frequency is positively correlated with SD length, proximity, and sequence identity, and the most common CNVs are therefore flanked by long stretches of near-perfect homology [41].

Shorter paralogous repeats can also mediate NAHR, albeit less frequently than long SDs. Sequencing across interspersed repeats is challenging, and until recently many of these breakpoints were missed by CNV sequencing studies. This year, recombination between LINE pairs was discovered at the breakpoints of 44 pathogenic CNVs. High sequence identity appears to be a requirement for LINE–LINE rearrangements because the minimum identity between recombining LINEs was 96%, and most pairs were greater than 97% identical [42]. Some LINEs had less than 1 kb of homology, suggesting that even fragmented LINEs can participate in NAHR. Recombination between HERV elements can also give rise to recurrent CNVs. Deletions and duplications mediated by HERV–HERV recombination at three intrachromosomal loci were sequenced in a recent study [43]. Similarly to other HERV-mediated chromosome rearrangements [44–46], all the CNVs are flanked by HERV-H elements that are at least 3 kb in length and 93–96% identical. The longer length and significant sequence identity of intact HERV-H elements may make them particularly recombinogenic.

By contrast, *Alu* repeats are only ~300 bp in length, and *Alu* pairs that flank deletions and duplications are 75–91% identical [27,34–36,47–49]. Sequencing the breakpoints of 54 CNVs at the *Alu*-rich *SPAST* locus revealed 38 that spanned hybrid *Alu* elements [47]. Lower sequence identity between *Alu* pairs suggests that these CNVs may not be the product of NAHR, but rather are the result of homeologous, or near-homologous, recombination that occurs between more-divergent sequences [50]. Compared to deletions and duplications mediated by LINE–LINE and HERV–HERV events (30 kb–5.5 Mb; median 523 kb),

those flanked by *Alu* elements tend to be smaller (1.9 kb–4.2 Mb; median 65.4 kb) [27,34–36,47–49].

Simple interchromosomal SV

Translocation is the exchange of genomic material between two different chromosomes (Figure 2C). The initial event that gives rise to translocations is usually reciprocal, producing two derivative chromosomes that are balanced. However, derivative translocation chromosomes may segregate in a balanced or an unbalanced manner. Balanced translocations are copy-neutral and do not cause a phenotype unless they disrupt developmentally important gene(s) at breakpoints. By contrast, unbalanced translocations result in trisomy and monosomy of chromosome ends and are usually found in individuals with developmental delay, intellectual disability, and/or birth defects, depending on the genes affected by the CNVs. Unbalanced translocations are easily detected by several methods, whereas detecting balanced translocations requires techniques that capture breakpoints, such as WGS or targeted NGS (Box 1).

Like intrachromosomal rearrangements, most constitutional translocations are non-recurrent, and microhomology is the most common feature at breakpoint junctions. Two recent studies of unbalanced translocations reported different types of DNA repair at junctions [45,46]. Sequencing the junctions of 37 unbalanced translocations revealed that 34 lacked extensive sequence homology [46]. Three unbalanced translocations had breakpoints consistent with NAHR between pairs of LINEs, HERVs, or short SDs; however, most breakpoint junctions had blunt ends, microhomology, inserted sequence, or inversions, indicating that most unbalanced translocations arise by NHEJ or MMBIR. These breakpoint signatures are similar to those from sequenced balanced translocations [51–53]. By contrast, sequencing another cohort of nine unbalanced translocations revealed that six were mediated by NAHR between 6 kb LINE, 3 kb HERV, or 1.7 kb SD pairs that are each >90% identical [45]. Although in this group NAHR between paralogous repeats appeared to be the 'driver' of unbalanced translocations, this is unlikely to be the major mechanism of translocation formation. In both unbalanced translocation studies, LINE and HERV elements were capable of NAHR, whereas no *Alu–Alu* events were detected [45,46]. Indeed, *Alu–Alu* recombination has been reported in only three translocations [53–55]. This trend suggests that, for NAHR-mediated rearrangements, those that are interchromosomal may require longer stretches of homology and greater sequence identity than those that are intrachromosomal.

Recurrent translocations are caused by NAHR between homologous sequences on different chromosomes, or by breakage hotspots in palindromic AT-rich repeats (PATRRs). The same SDs responsible for reciprocal deletions and duplications of the short arm of chromosome 8 also underlie recurrent translocations between chromosomes 4, 8, and 12 [56–58]. A recurrent translocation between chromosomes 4 and 18 is also caused by NAHR between 92% identical HERV-H repeats [44]. PATRRs on chromosomes 3, 8, 11, 17, and 22 give rise to recurrent translocations, the best known of which is the der(22)t(11:22), which causes Emanuel syndrome [59].

Complex chromosome rearrangement

Complex chromosome rearrangements have three or more breakpoints and may lead to a balanced or an unbalanced copy-number state [60,61]. Recent NGS breakpoint studies have paved the way to understanding the mutational mechanisms and defining the genomic structure of these rearrangements. We describe here insights into the major classes of complex chromosome rearrangements.

Inverted duplication adjacent to terminal deletion

Inverted duplication next to terminal deletion is a common type of rearrangement that has been recognized in cancer and constitutional genomes [62,63]. Several models have been put forth to explain these CNVs, and all include a dicentric chromosome that goes through a breakage-fusion-bridge cycle. Analysis of 34 sequenced breakpoints revealed spacers with normal copy number (median size 3 kb) between the inverted duplications (Figure 3A) and short inverted homology at the edges of the inverted segments. These molecular features support a model wherein the initial DSB leads to a terminal deletion, followed by fold-back of the truncated chromosome, formation of a dicentric chromosome, and a second DSB between the two centromeres that is repaired by addition of a new telomere [63]. The disomic spacers between inverted duplications correspond to the fold-back portion of the chromosome, and their discovery provided important insight in the formation of these complex chromosome rearrangements. Spacers are too small to detect by array-based methods, and sequencing breakpoint junctions was therefore a major advance in understanding this rearrangement mechanism.

Inverted duplications adjacent to deletions have also been described in ring chromosomes [64], an interstitial chromosome rearrangement [65], and unbalanced translocations [63]. These rearrangements are also are formed through a dicentric chromosome step, but instead of resolving as a terminal deletion the second DSB is repaired by an internal site on the same chromosome or by capture of a nonhomologous chromosome (Figure 3E).

Duplication-normal-duplication (DUP-NML-DUP)

Adjacent duplications with a normal copy-number region between them have a characteristic 'DUP-NML-DUP' pattern by array CGH (Figure 3B). Sequencing DUP-NML-DUP junctions revealed that most are interconnected with duplications in direct or inverted orientation [34,66–68]. These interstitial duplications are derived from regions of the same chromosome arm that are hundreds of kb to Mb apart. DUP-NML-DUPs are not associated with a particular syndrome because they are derived from diverse genomic loci and involve different genes. Depending on the spacing of probes, some DUP-NML-DUPs may appear as a single duplication by array CGH, and their prevalence is therefore likely to be underestimated. DUP-NML-DUPs have the potential to duplicate, fuse, and/or disrupt genes at breakpoints; therefore, determining their genomic structure is essential to identify genes involved in disease. For example, a DUP-NML-DUP of chromosome 14 fuses the *KCNH5* (potassium channel, voltage-gated eag-related subfamily H, member 5) and *FUT8* (fucosyltransferase 8) genes at

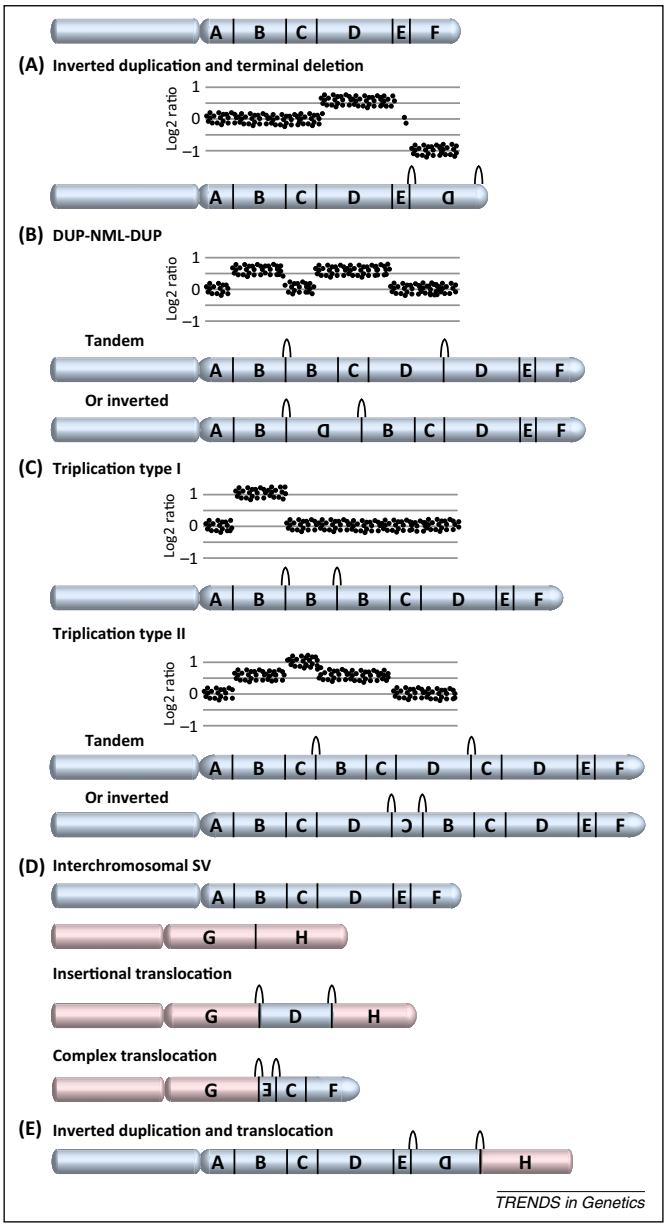

Figure 3. Complex chromosome rearrangements. Complex rearrangements and their array comparative genome hybridization (CGH) signatures are shown relative to the blue reference chromosome (top) divided into segments A–F. **(A)** Inverted duplications adjacent to terminal deletions have a short disomic spacer region (segment E) between inverted duplications. **(B)** A duplication-normal-duplication (DUP-NML-DUP) appears by array CGH as two copy-number gains (segments B and D). The duplications may be in direct orientation, or one duplicated segment (D) may be inverted between two copies of the other (B). **(C)** Triplication type I has three direct copies of B. In triplication type II the triplication (C) is embedded within a duplicated region (B–D). The triplicated segment may be in direct or inverted orientation. **(D)** Complex interchromosomal rearrangements occur between the blue and pink chromosomes. An insertional translocation involves the interstitial insertion of one chromosome segment (D) into another chromosome. Some complex translocations have multiple chromosome segments and/or inversion at the breakpoint junction. **(E)** An inverted duplication with terminal deletion may end with the translocated end of a nonhomologous chromosome.

an inverted junction that is predicted to produce an in-frame fusion transcript [26].

Triplication

Triplications are often recognized by array or NGS as segments with increased copy number within a duplicated segment. Type I triplications are oriented head-to-tail, without flanking duplications, and are formed via NAHR between SDs [39] (Figure 3C). Type II triplications lie within larger duplications and may or may not involve SDs at breakpoints (Figure 3C). In most type II CNVs, the triplicated segment is inverted relative to the duplications, a structure known as DUP-TRP/INV-DUP (Figure 4) [69–73].

DUP-TRP/INV-DUP of the *PLP1* (proteolipid protein 1) gene on the X chromosome causes Pelizaeus–Merzbacher disease, and these complex triplications lead to a more severe clinical phenotype than *PLP1* duplications that also cause the disease [69]. Triplication breakpoints cluster at inverted SDs distal of *PLP1*, and sequence analysis of 17 *PLP1* DUP-TRP/INV-DUPs revealed that a recurrent breakpoint junction lies within these inverted repeats [69]. Such DUP-TRP/INV-DUPs are proposed to form via a two-step process involving replication fork collapse and strand invasion between inverted repeats, followed by MMBIR or NHEJ (Figure 4) [67,69]. DUP-TRP/INV-DUPs of *MECP2* (methyl CpG binding protein 2) also have recurrent breakpoints within inverted repeats and cause a more severe form of *MECP2* duplication syndrome [71].

Triplications in the same orientation as flanking duplications have been described at other loci [34]. Whereas inverted triplications tend to have inverted repeats at junctions, direct triplications lack inverted repeats [34,67,69]. Recently, terminal regions of absence of heterozygosity were detected distal of some triplications. Extended absence of heterozygosity adjacent to triplications is likely due to MMBIR template switching between homologous chromosomes, which leads to regional uniparental disomy at end of the chromosome [70].

Insertional translocation

In common with other complex chromosome rearrangements, insertional translocations have more than two breakpoints. As opposed to more common translocations of chromosome ends, these translocated segments are inserted interstitially into a nonhomologous chromosome (Figure 3D). Insertional translocations often appear to be simple interstitial duplications by copy-number studies; however, FISH and breakpoint analyses revealed that ~2% of genomic gains detected by array CGH are inserted in another chromosome [34,74–76]. Unbalanced insertional translocations may be inherited from parents with the balanced form of the rearrangement [75,76], and in some cases the insertional translocation includes multiple segments in direct or inverted orientation [34,46,52].

Chromoanagenesis

The most severe forms of genomic reorganization are described as 'chromoanagenesis,' or chromosome rebirth, so named because the chromosomes are rearranged beyond recognition [77]. Chromosome shattering, or 'chromothripsis' [78], and chromosome reconstitution, or 'chromoanasynthesis' [68], are two types of chromoanagenesis, and their underlying mechanisms are only now beginning to be understood.

Chromothripsis

Chromothripsis was originally detected in chronic lymphocytic leukemia, where dozens of breakpoints were clustered

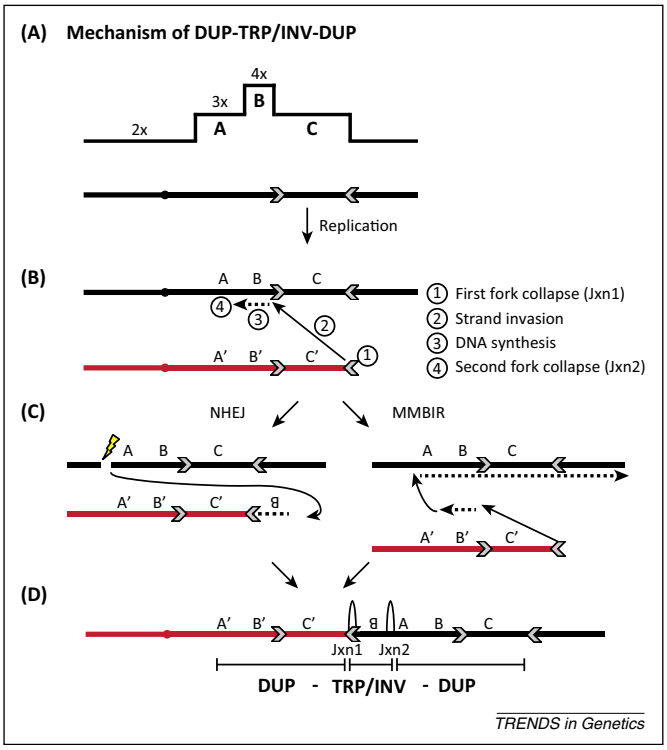

Figure 4. DUP-TRP/INV-DUP (duplication-inverted triplication-duplication) formation. **(A)** Copy-number changes are detected relative to the black reference chromosome. 2× indicates normal disomic copy number, while 3× genomic copies of A and C are duplications, and 4× total copies of segment B represents a triplication. Inverted repeats (grey arrows) are present at the edges of segment C. **(B)** At a collapsed replication fork, sequence homology drives strand invasion from one inverted repeat into one from the opposite strand. DNA synthesis is reinitiated until the occurrence of a second collapsed replication fork. **(C)** This second junction may arise from a non-homologous end-joining (NHEJ) or microhomology-mediated break-induced replication (MMBIR) mechanism. In NHEJ, a double-strand break (DSB) occurs on the original DNA strand and is repaired by joining the end of the replicated strand. In MMBIR, the lagging strand disengages, invades upstream sequence, and synthesizes DNA along the rest of the chromosome. **(D)** The resulting structure is a duplication, inverted triplication, and duplication. Orientation of the triplicated 'B' is confirmed by sequencing across junctions Jxn1 and Jxn2. Figure adapted with permission from [71].

on a single chromosome arm [78]. Chromothripsis is present in ~2% of cancer genomes [79] and has been reported at similar frequencies in constitutional chromosome rearrangements.

Chromothripsis involving up to five different chromosomes has been described in children with neurodevelopmental disorders (Figure 5A). Long stretches of homology are absent from the breakpoint junctions, and thus DNA repair likely occurs via NHEJ [46,52,80–86]. Despite tens of breakpoints per genome, constitutional chromothripsis is largely copy-neutral. Retention of essentially normal copy number in chromothriptic genomes could be mechanistically important, or could simply reflect selective pressure in liveborn individuals [86]. Some breakpoints have adjacent deletions, and many are inverted, but duplications are rare [46,81,85]. WGS is ideal to capture tens of breakpoints in one experiment, including balanced translocations and inversions in chromothriptic genomes that go unnoticed by other methods [46,52,81,84,85]. However, visualization of chromosomes is still necessary to localize rearranged segments and determine the contiguous structure of chromosomes scrambled by chromothripsis [46,82,83]. Breakpoint analysis of a growing number of complex rearrangements has revealed that translocations involving three or more different chromosomes are likely formed via chromothripsis [46,81,85].

Most constitutional chromothripsis events occur *de novo*, and those investigated thus far have all turned out to be paternal in origin [46,81,85]. This raises the possibility that chromosome shattering occurs in the male germline. By contrast, mitotic errors in the early embryo [87] or pulverization of micronuclei [88] could be responsible for numerous DNA breaks. Chromothriptic chromosomes have also been transmitted from mothers with a balanced form of the rearrangement [46,89]. Although children who inherit unbalanced chromothriptic chromosomes have fewer breakpoints than their parent with the balanced form, CNV due to unbalanced inheritance of the derivative chromosomes leads to a phenotype [46,89].

In addition to cancer and constitutional situations, chromothripsis has been observed upon integration of a transgene [52] and in a hematopoietic stem cell lineage [90]. Somatic chromothripsis was recently described in a woman with WHIM (warts, hypogammaglobulinemia, infections, and myelokathexis) syndrome, a rare immunodeficiency disorder resulting from a mutated copy of the *CXCR4* (chemokine C-X-C motif receptor 4) gene. In this case, chromothriptic deletion of her dominant *CXCR4* mutation led to reversion of the disease [90].

Chromoanasynthesis

Chromosome reconstitution confined to a single chromosome or locus has been termed chromoanasynthesis [68]. Whereas chromothripsis is limited to two copy-number states, has features of NHEJ at breakpoints, and may involve multiple chromosomes, chromoanasynthesis leads

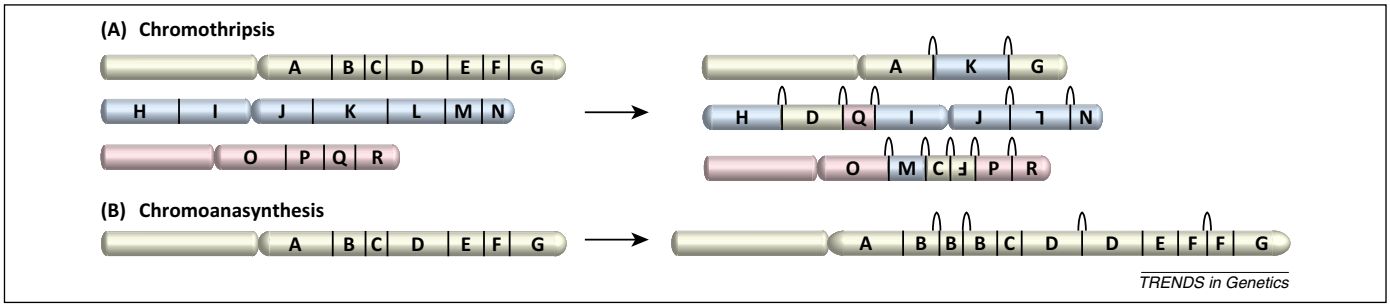

Figure 5. Massive genomic reorganization. **(A)** Chromothripsis shatters three nonhomologous chromosomes. The only copy-number variations (CNVs) are deletions of B and D, but translocating segments and inversions have shuffled the contents of the three chromosomes. The 12 breakpoint junctions have blunt ends or short microhomology. **(B)** Chromoanagenesis leads to triplication (B) and duplications (D and F) across one chromosome. These breakpoint junctions contain microhomology and insertions that suggest a DNA replication-based mechanism of repair.

to deletions, duplications, and triplications along a single chromosome (Figure 5B) [68,91,92]. Constitutional chromoanasynthesis has been recognized in rearrangements that involve eight to 33 breakpoints [68,92], and sequenced junctions bear signatures of FoSTeS and MMBIR [68]. Going forward, as WGS is more widely applied to SV, we expect to better define the features and origins of these highly complex chromosome rearrangements.

SV hotspots

De novo chromosome breakpoints at the same position suggest that the underlying DNA sequence and/or chromatin are risk factors for breakage. One of the most common SV hotspots lies between exons 8 and 9 of the *SHANK3* (SH3 and multiple ankyrin repeat domains 3) gene. Breakage in G-rich tandem repeats at this locus most often resolves as a terminal deletion of chromosome 22, but can also lead to other types of SV [27,93]. Similarly to PATRRs, these repetitive sequences may form aberrant secondary structures that are particularly susceptible to rearrangement.

Some *Alu*, HERV, and LINE repeats are also recognized as SV hotspots. As described above, some paralogous repeats give rise to recurrent SVs mediated by NAHR [44,47]. Others are not substrates for NAHR but underlie common breakpoints from independent rearrangements. Two translocation breakpoints in chromosome 12 lie in the same HERV-H, although the homologous HERV-H partner for each translocation is located on a different chromosome [45,46]. A L1PA4 element on chromosome 9 is another reused translocation breakpoint. One translocation is mediated by LINE–LINE recombination [45], and the other translocation occurred via NHEJ [46]. Thus, even at the same breakpoint, diverse DNA repair mechanisms give rise to different SVs.

In three separate chromothriptic genomes, *PTPRD* (protein tyrosine phosphatase receptor type D) on chromosome 9 is disrupted by several breakpoints [46,82,89]. The breakpoints on chromosome 9 are not the same, but this raises the possibility that the *PTPRD* locus is a chromothripsis hotspot. Deletions and duplications within *PTPRD* are among the most common SVs in chromothriptic neuroblastoma genomes [94,95].

Consequences of SV

Fusion genes

In many cases, SV breakpoints intersect open reading frames of genes. Although the transcriptional consequences of most SVs have not been investigated, breakpoints that disrupt or fuse genes have the potential to wreak havoc on normal development. Fusion genes are predicted at the breakpoints of constitutional deletions [47,96], duplications [34,96], balanced translocations [97,98], unbalanced translocations [46], an insertional translocation [34], an inverted DUP-NML-DUP [34], and chromothriptic rearrangements [84]. Genes disrupted or fused at the breakpoints of balanced rearrangements are excellent candidates for neurodevelopmental disorders because the rest of the genome is intact. However, fusion genes in unbalanced rearrangements also have the potential to acquire new functions related to phenotypic outcomes.

Mutations adjacent to breakpoint junctions

Although DNA sequence at breakpoints is known to be altered by resection, insertion, and inversion, recent studies suggest that regions further from junctions are also mutated. Complex duplications of the *MECP2* locus have SNV within 50 bp of breakpoint junctions that arose at the same time as the *de novo* duplications [67]. Similar 'micromutations' have been detected adjacent to pathogenic deletions of five different chromosomes [99]. It remains to be determined whether these mutations occur at other SV, but this phenomenon may be similar to the mutations induced by APOBEC (apolipoprotein B mRNA editing enzyme, catalytic polypeptide) cytosine deaminase that are associated with somatic mutations in cancer [100].

Position effect

SVs may also exert position effects that alter the expression of intact genes near breakpoints. Position effects have been noted at the *FOXL2* (forkhead box L2) [101–103], *PLP1* [104], *SHOX* (short stature homeobox) [105], and *SOX9* (sex-determining region box 8) [106,107] genes, among others. In the recurrent translocation between chromosomes 11 and 22, aberrant nuclear positioning of translocated regions results in differential expression of many genes on different chromosomes [108]. Future studies of *cis* and *trans* position effects related to SV may inform phenotypes when even thorough breakpoint analysis by WGS fails to pinpoint genes related to disease [109].

Technical caveats

Although NGS and associated data analysis methods continue to improve, there are still some major challenges in breakpoint sequencing. Filtering discordant sequence reads to identify paired ends that span SV breakpoints is one successful strategy to pinpoint breakpoints; however, it is difficult to uniquely map short reads to repetitive DNA. One way to increase the likelihood of capturing breakpoints in repeats is by creating large-insert 'jumping' libraries, where mate pairs span both sides of interspersed repeats [110,111]. This is especially useful for inversions and balanced translocations that are not detected by copy-number methods. As opposed to more common paired-end sequencing of short DNA fragments, single-molecule real-time (SMRT) technology generates long reads (mean mapped length 5.8 kb) that cross repetitive elements. Although SMRT has a higher error rate per read, it works well for GC-rich regions and tandem repeats, which is essential to resolve complex SV structure [112].

WGS is ideal for finding breakpoints in chromothripsis, which has many copy-neutral segments and few breakpoints within repeats; however, it can be difficult to intuit complex contigs with tens of breakpoint junctions across multiple chromosomes. Existing algorithms call SV based on discordant reads from paired-end data; however, there is the potential for significant false discovery [4,17,113]. This can be overcome with laborious rounds of FISH to place genomic segments on the correct derivative chromosome [46], but more high-throughout methods are needed for accurate placement of complex SV. Outstanding questions are listed in Box 2.

Box 2. Outstanding questions

- The consequences of SV include fusion genes, SNV adjacent to breakpoints, and position effects acting on nearby genes. To what extent do these SV byproducts contribute to clinical phenotypes?
- The ability to sequence across and accurately map repetitive regions prohibits comprehensive analysis of SV structure. What breakpoints are we missing?
- Chromoanagenesis, including chromothripsis and chromoanasynthesis, describes ultracomplex rearrangements. How prevalent is chromoanagenesis, and what are the molecular mechanisms that distinguish chromothripsis and chromoanasynthesis?

Concluding remarks

SV breakpoint analyses reveal multiple pathways to recurrent and non-recurrent chromosome rearrangements. In addition to SDs, LINE–LINE and HERV–HERV NAHR is responsible for some recurrent rearrangements. *Alu–Alu* recombination may also lead to recurrent intrachromosomal CNVs, but via a homeologous mechanism. Analysis of thousands of breakpoints revealed that most human SVs are non-recurrent and are formed via NHEJ or MMBIR. The increased use of high-resolution SV detection methods has teased apart complex SV architectures, including chromothripsis and chromoanasynthesis. With the capability to sustain fusion genes, tens of breakpoints that cause massive genomic reorganization, and breakage hotspots, some constitutional rearrangements may be as striking as those found in cancer. Functional analyses of chromosome breakage in SV hotspots and position effects acting on genes near breakpoints are important next steps that will tell us more about the causes and effects of human SV.

References

1 Redon, R. *et al.* (2006) Global variation in copy number in the human genome. *Nature* 444, 444–454
2 Conrad, D.F. *et al.* (2010) Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. *Nat. Genet.* 42, 385–391
3 Kidd, J.M. *et al.* (2010) A human genome structural variation sequencing resource reveals insights into mutational mechanisms. *Cell* 143, 837–847
4 Mills, R.E. *et al.* (2011) Mapping copy number variation by population-scale genome sequencing. *Nature* 470, 59–65
5 Stankiewicz, P. and Lupski, J.R. (2010) Structural variation in the human genome and its role in disease. *Annu. Rev. Med.* 61, 437–455
6 Cooper, G.M. *et al.* (2011) A copy number variation morbidity map of developmental delay. *Nat. Genet.* 43, 838–846
7 Coe, B.P. *et al.* (2014) Refining analyses of copy number variation identifies specific genes associated with developmental delay. *Nat. Genet.* 46, 1063–1071
8 Kaminsky, E.B. *et al.* (2011) An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. *Genet. Med.* 13, 777–784
9 Miller, D.T. *et al.* (2010) Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am. J. Hum. Genet.* 86, 749–764
10 Yang, Y. *et al.* (2014) Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 312, 1870–1879
11 Lee, H. *et al.* (2014) Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 312, 1880–1887
12 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
13 Lieber, M.R. (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* 79, 181–211
14 Hastings, P.J. *et al.* (2009) Mechanisms of change in gene copy number. *Nat. Rev. Genet.* 10, 551–564
15 Zhang, F. *et al.* (2009) The DNA replication FoSTeS/MMBIR mechanism can generate genomic, genic and exonic complex rearrangements in humans. *Nat. Genet.* 41, 849–853
16 Abyzov, A. *et al.* (2011) CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* 21, 974–984
17 Chiang, D.Y. *et al.* (2009) High-resolution mapping of copy-number alterations with massively parallel sequencing. *Nat. Methods* 6, 99–103
18 Haraksingh, R.R. *et al.* (2011) Genome-wide mapping of copy number variation in humans: comparative analysis of high resolution array platforms. *PLoS ONE* 6, e27859
19 Kidd, J.M. *et al.* (2008) Mapping and sequencing of structural variation from eight human genomes. *Nature* 453, 56–64
20 Rasekh, M.E. *et al.* (2015) Discovery of large genomic inversions using pooled clone sequencing. *bioRxiv* Published online February 11, 2015. http://dx.doi.org/10.1101/015156
21 Williams, L.J. *et al.* (2012) Paired-end sequencing of fosmid libraries by Illumina. *Genome Res.* 22, 2241–2249
22 Tuzun, E. *et al.* (2005) Fine-scale structural variation of the human genome. *Nat. Genet.* 37, 727–732
23 Watson, C.T. *et al.* (2014) The genetics of microdeletion and microduplication syndromes: an update. *Annu. Rev. Genomics Hum. Genet.* 15, 215–244
24 Giglio, S. *et al.* (2001) Olfactory receptor-gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. *Am. J. Hum. Genet.* 68, 874–883
25 Antonacci, F. *et al.* (2010) A large and complex structural polymorphism at 16p12.1 underlies microdeletion disease risk. *Nat. Genet.* 42, 745–750
26 Antonacci, F. *et al.* (2009) Characterization of six human disease-associated inversion polymorphisms. *Hum. Mol. Genet.* 18, 2555–2566
27 Luo, Y. *et al.* (2011) Diverse mutational mechanisms cause pathogenic subtelomeric rearrangements. *Hum. Mol. Genet.* 20, 3769–3778
28 Battaglia, A. *et al.* (1999) Natural history of Wolf–Hirschhorn syndrome: experience with 15 cases. *Pediatrics* 103, 830–836
29 Zhang, X. *et al.* (2005) High-resolution mapping of genotype–phenotype relationships in cri du chat syndrome using array comparative genomic hybridization. *Am. J. Hum. Genet.* 76, 312–326
30 Kleefstra, T. *et al.* (2009) Further clinical and molecular delineation of the 9q subtelomeric deletion syndrome supports a major contribution of EHMT1 haploinsufficiency to the core phenotype. *J. Med. Genet.* 46, 598–606
31 Grossfeld, P.D. *et al.* (2004) The 11q terminal deletion disorder: a prospective study of 110 cases. *Am. J. Med. Genet. A* 129A, 51–61
32 Durand, C.M. *et al.* (2007) Mutations in the gene encoding the synaptic scaffolding protein SHANK3 are associated with autism spectrum disorders. *Nat. Genet.* 39, 25–27
33 Bose, P. *et al.* (2014) Tandem repeats and G-rich sequences are enriched at human CNV breakpoints. *PLoS ONE* 9, e101607
34 Newman, S. *et al.* (2015) Next-generation sequencing of duplication CNVs reveals that most are tandem and some create fusion genes at breakpoints. *Am. J. Hum. Genet.* 96, 208–220
35 Verdin, H. *et al.* (2013) Microhomology-mediated mechanisms underlie non-recurrent disease-causing microdeletions of the FOXL2 gene or its regulatory domain. *PLoS Genet.* 9, e1003358
36 Vissers, L.E. *et al.* (2009) Rare pathogenic microdeletions and tandem duplications are microhomology-mediated and stimulated by local genomic architecture. *Hum. Mol. Genet.* 18, 3579–3593
37 Itsara, A. *et al.* (2012) Resolving the breakpoints of the 17q21.31 microdeletion syndrome with next-generation sequencing. *Am. J. Hum. Genet.* 90, 599–613
38 Sharp, A.J. *et al.* (2006) Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat. Genet.* 38, 1038–1042
39 Liu, P. *et al.* (2012) Mechanisms for recurrent and complex human genomic rearrangements. *Curr. Opin. Genet. Dev.* 22, 211–220
40 Dittwald, P. *et al.* (2013) NAHR-mediated copy-number variants in a clinical population: mechanistic insights into both genomic disorders and Mendelizing traits. *Genome Res.* 23, 1395–1409

41 Liu, P. *et al.* (2011) Frequency of nonallelic homologous recombination is correlated with length of homology: evidence that ectopic synapsis precedes ectopic crossing-over. *Am. J. Hum. Genet.* 89, 580–588
42 Startek, M. *et al.* (2015) Genome-wide analyses of LINE–LINE-mediated nonallelic homologous recombination. *Nucleic Acids Res.* 43, 2188–2198
43 Campbell, I.M. *et al.* (2014) Human endogenous retroviral elements promote genome instability via non-allelic homologous recombination. *BMC Biol.* 12, 74
44 Hermetz, K.E. *et al.* (2012) A recurrent translocation is mediated by homologous recombination between HERV-H elements. *Mol. Cytogenet.* 5, 6
45 Robberecht, C. *et al.* (2013) Nonallelic homologous recombination between retrotransposable elements is a driver of de novo unbalanced translocations. *Genome Res.* 23, 411–418
46 Weckselblatt, B. *et al.* (2015) Unbalanced translocations arise from diverse mutational mechanisms including chromothripsis. *Genome Res.* 25, 937–947
47 Boone, P.M. *et al.* (2014) The Alu-rich genomic architecture of SPAST predisposes to diverse and functionally distinct disease-associated cnv alleles. *Am. J. Hum. Genet.* 95, 143–161
48 Carvalho, C.M. *et al.* (2014) Dosage changes of a segment at 17p13.1 lead to intellectual disability and microcephaly as a result of complex genetic interaction of multiple genes. *Am. J. Hum. Genet.* 95, 565–578
49 Shlien, A. *et al.* (2010) A common molecular mechanism underlies two phenotypically distinct 17p13.1 microdeletion syndromes. *Am. J. Hum. Genet.* 87, 631–642
50 Rossetti, L.C. *et al.* (2004) Homeologous recombination between AluSx-sequences as a cause of hemophilia. *Hum. Mutat.* 24, 440
51 Chen, W. *et al.* (2008) Mapping translocation breakpoints by next-generation sequencing. *Genome Res.* 18, 1143–1149
52 Chiang, C. *et al.* (2012) Complex reorganization and predominant non-homologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. *Nat. Genet.* 44, 390–397 S391
53 Higgins, A.W. *et al.* (2008) Characterization of apparently balanced chromosomal rearrangements from the developmental genome anatomy project. *Am. J. Hum. Genet.* 82, 712–722
54 Rouyer, F. *et al.* (1987) A sex chromosome rearrangement in a human XX male caused by Alu–Alu recombination. *Cell* 51, 417–425
55 Fruhmesser, A. *et al.* (2013) Disruption of EXOC6B in a patient with developmental delay, epilepsy, and a de novo balanced t(2;8) translocation. *Eur. J. Hum. Genet.* 21, 1177–1180
56 Giglio, S. *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
57 Goldlust, I.S. *et al.* (2013) Mouse model implicates GNB3 duplication in a childhood obesity syndrome. *Proc. Natl. Acad. Sci. U.S.A.* 110, 14990–14994
58 Ou, Z. *et al.* (2011) Observation and prediction of recurrent human translocations mediated by NAHR between nonhomologous chromosomes. *Genome Res.* 21, 33–46
59 Kato, T. *et al.* (2012) Chromosomal translocations and palindromic AT-rich repeats. *Curr. Opin. Genet. Dev.* 22, 221–228
60 Quinlan, A.R. and Hall, I.M. (2012) Characterizing complex structural variation in germline and somatic genomes. *Trends Genet.* 28, 43–53
61 Zhang, F. *et al.* (2009) Complex human chromosomal and genomic rearrangements. *Trends Genet.* 25, 298–307
62 Tanaka, H. *et al.* (2007) Intrastrand annealing leads to the formation of a large DNA palindrome and determines the boundaries of genomic amplification in human cancer. *Mol. Cell. Biol.* 27, 1993–2002
63 Hermetz, K.E. *et al.* (2014) Large inverted duplications in the human genome form via a fold-back mechanism. *PLoS Genet.* 10, e1004139
64 Murmann, A.E. *et al.* (2009) Inverted duplications on acentric markers: mechanism of formation. *Hum. Mol. Genet.* 18, 2241–2256
65 Milosevic, J. *et al.* (2014) Inverted duplication with deletion: first interstitial case suggesting a novel undescribed mechanism of formation. *Am. J. Med. Genet. A* 164A, 3180–3186
66 Brand, H. *et al.* (2014) Cryptic and complex chromosomal aberrations in early-onset neuropsychiatric disorders. *Am. J. Hum. Genet.* 95, 454–461
67 Carvalho, C.M. *et al.* (2013) Replicative mechanisms for CNV formation are error prone. *Nat. Genet.* 45, 1319–1326
68 Liu, P. *et al.* (2011) Chromosome catastrophes involve replication mechanisms generating complex genomic rearrangements. *Cell* 146, 889–903
69 Beck, C.R. *et al.* (2015) Complex genomic rearrangements at the PLP1 locus include triplication and quadruplication. *PLoS Genet.* 3, e1005050
70 Carvalho, C.M. *et al.* (2015) Absence of heterozygosity due to template switching during replicative rearrangements. *Am. J. Hum. Genet.* 96, 555–564
71 Carvalho, C.M. *et al.* (2011) Inverted genomic segments and complex triplication rearrangements are mediated by inverted repeats in the human genome. *Nat. Genet.* 43, 1074–1081
72 Ishmukhametova, A. *et al.* (2013) Dissecting the structure and mechanism of a complex duplication-triplication rearrangement in the DMD gene. *Hum. Mutat.* 34, 1080–1084
73 Shimojima, K. *et al.* (2012) Pelizaeus–Merzbacher disease caused by a duplication-inverted triplication-duplication in chromosomal segments including the PLP1 region. *Eur. J. Med. Genet.* 55, 400–403
74 Neill, N.J. *et al.* (2011) Recurrence, submicroscopic complexity, and potential clinical relevance of copy gains detected by array CGH that are shown to be unbalanced insertions by FISH. *Genome Res.* 21, 535–544
75 Kang, S.H. *et al.* (2010) Insertional translocation detected using FISH confirmation of array-comparative genomic hybridization (aCGH) results. *Am. J. Med. Genet. A* 152A, 1111–1126
76 Nowakowska, B.A. *et al.* (2012) Parental insertional balanced translocations are an important cause of apparently de novo CNVs in patients with developmental anomalies. *Eur. J. Hum. Genet.* 20, 166–170
77 Holland, A.J. and Cleveland, D.W. (2012) Chromoanagenesis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements. *Nat. Med.* 18, 1630–1638
78 Stephens, P.J. *et al.* (2011) Massive genomic rearrangement acquired in a single catastrophic event during cancer development. *Cell* 144, 27–40
79 Kim, T.M. *et al.* (2013) Functional genomic analysis of chromosomal aberrations in a compendium of 8000 cancer genomes. *Genome Res.* 23, 217–227
80 Genesio, R. *et al.* (2013) Pure 16q21q22.1 deletion in a complex rearrangement possibly caused by a chromothripsis event. *Mol. Cytogenet.* 6, 29
81 Kloosterman, W.P. *et al.* (2012) Constitutional chromothripsis rearrangements involve clustered double-stranded DNA breaks and nonhomologous repair mechanisms. *Cell Rep.* 1, 648–655
82 Macera, M.J. *et al.* (2014) Prenatal diagnosis of chromothripsis, with nine breaks characterized by karyotyping, FISH, microarray and whole-genome sequencing. *Prenat. Diagn.* 35, 299–301
83 Nazaryan, L. *et al.* (2014) The strength of combined cytogenetic and mate-pair sequencing techniques illustrated by a germline chromothripsis rearrangement involving FOXP2. *Eur. J. Hum. Genet.* 22, 338–343
84 van Heesch, S. *et al.* (2014) Genomic and functional overlap between somatic and germline chromosomal rearrangements. *Cell Rep.* 9, 2001–2010
85 Kloosterman, W.P. *et al.* (2011) Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline. *Hum. Mol. Genet.* 20, 1916–1924
86 Kloosterman, W.P. and Cuppen, E. (2013) Chromothripsis in congenital disorders and cancer: similarities and differences. *Curr. Opin. Cell Biol.* 25, 341–348
87 Vanneste, E. *et al.* (2009) Chromosome instability is common in human cleavage-stage embryos. *Nat. Med.* 15, 577–583
88 Crasta, K. *et al.* (2012) DNA breaks and chromosome pulverization from errors in mitosis. *Nature* 482, 53–58
89 de Pagter, M.S. *et al.* (2015) Chromothripsis in healthy individuals affects multiple protein-coding genes and can result in severe congenital abnormalities in offspring. *Am. J. Hum. Genet.* 96, 651–656
90 McDermott, D.H. *et al.* (2015) Chromothriptic cure of WHIM syndrome. *Cell* 160, 686–699

91 Zanardo, E.A. *et al.* (2014) Complex structural rearrangement features suggesting chromoanagenesis mechanism in a case of 1p36 deletion syndrome. *Mol. Genet. Genomics* 289, 1037–1043
92 Plaisancie, J. *et al.* (2014) Constitutional chromoanasynthesis: description of a rare chromosomal event in a patient. *Eur. J. Med. Genet.* 57, 567–570
93 Bonaglia, M.C. *et al.* (2006) Identification of a recurrent breakpoint within the SHANK3 gene in the 22q13.3 deletion syndrome. *J. Med. Genet.* 43, 822–828
94 Molenaar, J.J. *et al.* (2012) Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature* 483, 589–593
95 Boeva, V. *et al.* (2013) Breakpoint features of genomic rearrangements in neuroblastoma with unbalanced translocations and chromothripsis. *PLoS ONE* 8, e72182
96 Rippey, C. *et al.* (2013) Formation of chimeric genes by copy-number variation as a mutational mechanism in schizophrenia. *Am. J. Hum. Genet.* 93, 697–710
97 Backx, L. *et al.* (2011) A balanced translocation t(6;14)(q25.3;q13.2) leading to reciprocal fusion transcripts in a patient with intellectual disability and agenesis of corpus callosum. *Cytogenet. Genome Res.* 132, 135–143
98 Utami, K.H. *et al.* (2014) Detection of chromosomal breakpoints in patients with developmental delay and speech disorders. *PLoS ONE* 9, e90852
99 Wang, Y. *et al.* (2015) Characterization of 26 deletion CNVs reveals the frequent occurrence of micro-mutations within the breakpoint-flanking regions and frequent repair of double-strand breaks by templated insertions derived from remote genomic regions. *Hum. Genet.* 134, 589–603
100 Roberts, S.A. *et al.* (2013) An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. *Nat. Genet.* 45, 970–976
101 Beysen, D. *et al.* (2005) Deletions involving long-range conserved nongenic sequences upstream and downstream of FOXL2 as a novel disease-causing mechanism in blepharophimosis syndrome. *Am. J. Hum. Genet.* 77, 205–218
102 Fantes, J. *et al.* (1995) Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum. Mol. Genet.* 4, 415–422
103 Bhatia, S. *et al.* (2013) Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved PAX6 enhancer causes aniridia. *Am. J. Hum. Genet.* 93, 1126–1134
104 Lee, J.A. *et al.* (2006) Spastic paraplegia type 2 associated with axonal neuropathy and apparent PLP1 position effect. *Ann. Neurol.* 59, 398–403
105 Fukami, M. *et al.* (2006) Transactivation function of an approximately 800-bp evolutionarily conserved sequence at the SHOX 3′ region: implication for the downstream enhancer. *Am. J. Hum. Genet.* 78, 167–170
106 Velagaleti, G.V. *et al.* (2005) Position effects due to chromosome breakpoints that map approximately 900 Kb upstream and approximately 1.3 Mb downstream of SOX9 in two patients with campomelic dysplasia. *Am. J. Hum. Genet.* 76, 652–662
107 Hill-Harfe, K.L. *et al.* (2005) Fine mapping of chromosome 17 translocation breakpoints ≥900 Kb upstream of SOX9 in acampomelic campomelic dysplasia and a mild, familial skeletal dysplasia. *Am. J. Hum. Genet.* 76, 663–671
108 Harewood, L. *et al.* (2010) The effect of translocation-induced nuclear reorganization on gene expression. *Genome Res.* 20, 554–564
109 Gilissen, C. *et al.* (2014) Genome sequencing identifies major causes of severe intellectual disability. *Nature* 511, 344–347
110 Hanscom, C. and Talkowski, M. (2014) Design of large-insert jumping libraries for structural variant detection using illumina sequencing. *Curr. Protoc. Hum. Genet.* 80, 7 22 21–29
111 Korbel, J.O. *et al.* (2007) Paired-end mapping reveals extensive structural variation in the human genome. *Science* 318, 420–426
112 Chaisson, M.J. *et al.* (2015) Resolving the complexity of the human genome using single-molecule sequencing. *Nature* 517, 608–611
113 Korbel, J.O. *et al.* (2009) PEMer: a computational framework with simulation-based error models for inferring genomic structural variants from massive paired-end sequencing data. *Genome Biol.* 10, R23