



REVIEW

A unifying model that explains the origins of human inverted copy number variants

Bonita J. Brewer ^{*}, Maitreya J. Dunham, M. K. Raghuraman 

Department of Genome Sciences, University of Washington, Seattle, Washington, United States of America

^{*} bbrewer@uw.edu

Abstract

With the release of the telomere-to-telomere human genome sequence and the availability of both long-read sequencing and optical genome mapping techniques, the identification of copy number variants (CNVs) and other structural variants is providing new insights into human genetic disease. Different mechanisms have been proposed to account for the novel junctions in these complex architectures, including aberrant forms of DNA replication, non-allelic homologous recombination, and various pathways that repair DNA breaks. Here, we have focused on a set of structural variants that include an inverted segment and propose that they share a common initiating event: an inverted triplication with long, unstable palindromic junctions. The secondary rearrangement of these palindromes gives rise to the various forms of inverted structural variants. We postulate that this same mechanism (ODIRA: origin-dependent inverted-repeat amplification) that creates the inverted CNVs in inherited syndromes also generates the palindromes found in cancers.

 OPEN ACCESS

Citation: Brewer BJ, Dunham MJ, Raghuraman MK (2024) A unifying model that explains the origins of human inverted copy number variants. *PLoS Genet* 20(1): e1011091. <https://doi.org/10.1371/journal.pgen.1011091>

Editor: Michael Lichten, National Cancer Institute, UNITED STATES

Published: January 4, 2024

Copyright: © 2024 Brewer et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the National Institutes of Health (<https://www.nih.gov>) grants R01GM018926 and R35 GM122497 to BJB and R01GM147040 to MJD. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: aCGH, array comparative genome hybridization; AOH, absence of heterozygosity; BFB, breakage-fusion-bridge; BIR, break-induced replication; CNV, copy number variant; EJ, end joining; FISH, fluorescent in situ hybridization; FoSTeS, fork-stalling and template switching; HR,

Introduction

Human structural variants (SVs) are DNA rearrangements that involve large segments of the human genome and can have profound effects on phenotype. They are classified as resulting in changes in copy number (copy number variants, CNVs), in chromosome location (translocations, local, or dispersed duplications) and/or in orientation (direct versus inverted). This analysis focuses specifically on locally inverted CNVs—that is, the change in copy number and orientation is confined to the vicinity of the initial region. Local inverted CNVs may include duplications and triplications, may be interspersed with copy number neutral regions, or may be associated with neighboring deletions or regions of homozygosity (ROH; also referred to as absence of heterozygosity, AOH), but have at least 1 repeated segment that is inverted with respect to its native orientation.

Many CNVs have been found through analysis of patients with clinical phenotypes. Identifying the genes responsible for these phenotypes is the primary focus of the clinical reports, but the sequencing data can also provide insights into the mechanism(s) responsible for their generation. The novel junction sequences that flank the CNVs have been interpreted as evidence for mechanisms of their formation and are as varied as the events themselves (reviewed in [1]). These mechanisms include breakage-fusion-bridge cycles (BFB [2]), double-stranded break repair (non-homologous end joining—NHEJ [3] and microhomology-mediated end joining—MMEJ [4]), non-allelic homologous recombination (NAHR [5]), and altered DNA

homologous recombination; ICTS, intrachromosomal template switching; LCR, low copy repeat; MMBIR, microhomology-mediated break-induced replication; MMEJ, microhomology-mediated end joining; NAHR, non-allelic homologous recombination; NHEJ, non-homologous end joining; ODIRA, origin-dependent inverted-repeat amplification; ROH, regions of homozygosity; SV, structural variant.

replication that result from either a collapsed replication fork (microhomology-mediated break-induced replication—MMBIR [6]) or a stalled replication fork (fork-stalling and template switching—FoSTeS [7]). Sequencing of the parents' genomes can determine whether the CNV was preexisting or de novo, but does not address the responsible mechanisms. Except for FoSTeS, all of these potential mechanisms have been identified, genetically dissected, and verified in eukaryotic model organisms [2,3,5,8]. Nevertheless, perhaps because FoSTeS is the most flexible of the models, it is frequently invoked as explaining a variety of different CNVs (for example, [9–11]). We believe that an alternative model to FoSTeS that we have experimentally validated in yeast [12–14] can explain the genesis of all locally inverted CNVs.

In this analysis, we have culled from the human literature 45 representative CNV events with localized inverted DNA segments that have well-characterized junction sequences and we have explored the possibility that there is a unifying explanation for their formation. Our hypothesis is that there is a shared, unstable precursor that gives rise to the broad range of events through secondary rearrangements. That unstable structure is the palindrome: long palindromes (with perfect inverted arms), quasi-palindromes (with mismatches between the 2 inverted arms), and interrupted palindromes (with a short spacer between the 2 inverted arms) are unstable in all organisms where they have been introduced and are rare in natural genomes [15,16] or are associated with disease [17]. Our proposal that palindromes are the predisposing structure for inverted CNVs in humans is inspired by the inverted amplification of the *SUL1* locus in the yeast *Saccharomyces cerevisiae* that invariably arises during continuous growth of laboratory strains in limiting sulfate medium. The *SUL1* amplicons have a palindromic structure with junctions that arise at closely spaced short inverted repeats. Such repeats occur at high frequency throughout the yeast genome and serve as the sites for template switching of the leading strand to the lagging strand template. Because this replication error depends on both an origin of replication and the inverted repeats, we have called this model ODIRA (origin-dependent inverted-repeat amplification; [12–14]). Similar amplification events, consistent with an ODIRA mechanism, have also been reported at the yeast *GAP1* and *DUR3* loci [18].

While these palindromic junctions persist in yeast under selective conditions, they are unstable and can undergo secondary rearrangements through end joining (EJ; including NHEJ or MMEJ), break-induced replication (BIR; including MMBIR), or homologous recombination (HR; including NAHR) (Fig 1), creating new junctions and increasing the length of the spacer between the 2 palindromic arms. Every class of localized, inverted human CNV that we have analyzed can be explained by this same type of secondary rearrangement of an inverted triplication. Among these broad classes are inverted triplications located between direct repeats (DUP-TRP/INV-DUP), inverted duplications associated with deletions (INV-DUP-DEL), duplications/triplications flanking copy-neutral segments, AOH distal to the SVs, and telomere deletions distal to inverted SVs [19].

Genome-wide palindrome analysis (using a snap-back assay and sequencing after nuclease S1 treatment) reveals a rise in palindromic sequences in certain cancers [20–25]. “Cancer cells exhibit massive genome rearrangements, which include gene amplifications, translocations, and deletions, and these rearrangements are often associated with the presence of a palindrome, suggesting a possible correlation between the palindrome and the gene rearrangements.” In this National Cancer Institute interview Allison Rattray went on to say, “DNA palindromes are unstable and can lead to genome rearrangements by themselves, further suggesting palindromes could arise not only by sister chromatid fusion, but also by other mechanisms, such as replication errors.” (Platinum Highlight article, NCI, July 29, 2015, by Nancy Parrish, interview with Allison Rattray; <https://ncifrederick.cancer.gov/about/theposter/content/novel-method-developed-further-understanding-dna-palindromes>). We believe that

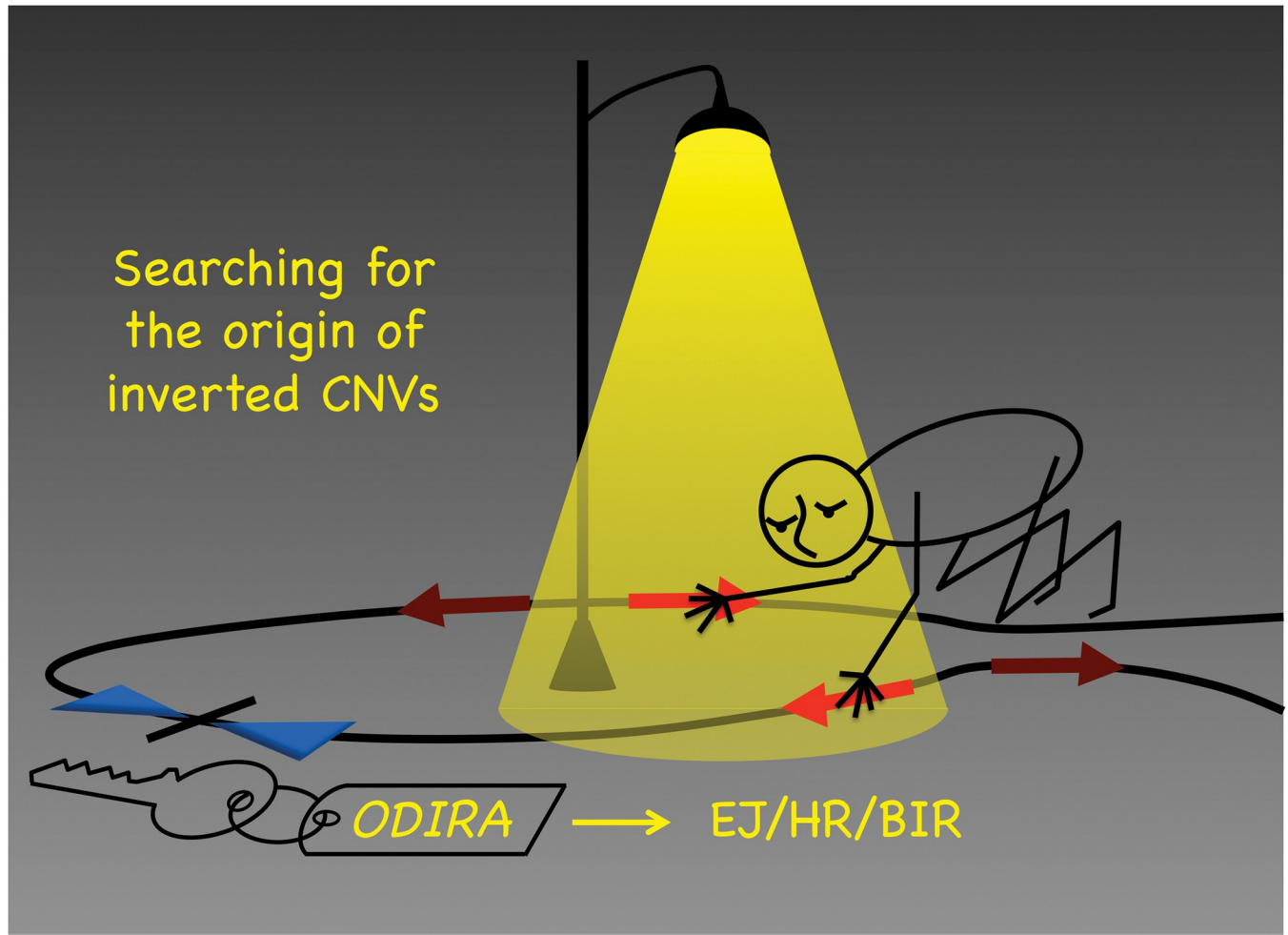


Fig 1. Searching for the origin of inverted CNVs. Noam Chomsky on how science operates: “Science is a bit like the joke about the drunk who is looking under a lamppost for a key that he has lost on the other side of the street, because that’s where the light is. It has no other choice.” [48] Sequencing the DNA of inverted CNVs is like looking under the lamppost. It reveals the eventual rearranged junctions (through forms of EJ, HR, and BIR), but may not capture the “key” initiating event (ODIRA) that remains hidden in the shadows. BIR, break-induced replication; CNV, copy number variant; EJ, end joining; HR, homologous recombination; ODIRA, origin-dependent inverted-repeat amplification.

<https://doi.org/10.1371/journal.pgen.1011091.g001>

we have identified Allison Rattray’s replication error that is responsible for many local, inverted CNVs and may be responsible for a particular subset of amplification events in cancer as well as inherited and de novo inverted CNVs.

Results

CNVs have been routinely discovered by such techniques as array comparative genome hybridization (aCGH) or read-depth analysis of short-read sequencing data [26,27]. However, these techniques do not reveal the genomic location of extra copies or of their orientation with respect to neighboring sequence. Discordant- or split-reads can provide information on the possible location and orientation, but are hard to definitively map in the human genome with its high density of repetitive sequences. Fluorescent in situ hybridization (FISH) has been invaluable for identifying and/or confirming location and orientation of duplicated or triplicated sequences; however, it does not reveal the sequences at the junctions. PCR with appropriately oriented primers can also confirm orientation at junctions. More recently, long-read

platforms such as PacBio and Oxford Nanopore sequencing technologies [26,27], Bionano optical genome mapping [28], and DNA combing [29] have begun to provide much-needed tools in the identification and characterization of these local CNVs.

Inverted triplications

Inverted triplications with long, nearly perfect palindromic junctions (Fig 2A), such as those found at the *SUL1* locus in yeast, are difficult to identify experimentally. They can be inferred through a combination of genome-wide copy number and allele frequency measurements (Fig 2B), but verification of the orientation of the amplified segments is through sequencing of the junction fragments (centromere- and telomere-proximal junctions, CJ and TJ, respectively; Fig 2). The most parsimonious structure that accommodates the number of additional copies and the junction sequences is a triplication where the center copy is inverted (TRP/INV; Fig 2A and 2C). Reports of TRP/INV CNVs in humans are rare, in part because older technologies using PCR and short-read sequencing are biased against recovering palindromic junctions. They are also a challenge for Nanopore sequencing [30], but as long-read sequencing technologies improve, and optical mapping and DNA combing become more widely used, we anticipate that the frequency of inverted CNV discovery is likely to increase.

Inverted triplications are inherently unstable as the direct duplications that flank the inverted segment can recombine with one another and restore euploidy. However, the junctions are also unstable due to their palindromic nature [24] and secondary rearrangements that delete one of the arms of the palindromes increases the distance between inverted segments and improves their stability [31–33]. One contributing factor to palindrome instability is that repetitive elements (such as LINES and SINEs) that were in opposite orientations before the amplification event will have copies that are in direct orientation after the inverted duplication (Fig 3A). NAHR between 2 of the direct repeats would delete the segment between them (Fig 3A). This recombination event results in loss of one of the copies of DNA between the 2 repetitive elements and is marked by a decrease from 4 to 3 in copy number measurements and alters the allele frequencies (Fig 3B). If a similar event occurs at the other junction, then the triplication is flanked by duplications on both margins.

For clarity, we have illustrated NAHR occurring at both junctions, but most interstitial triplications with this DUP-TRP/INV-DUP structure (Fig 3C and Table 1, examples 1–21) do not have junctions that map to LINES, SINEs, or other low copy repeats (LCRs) and are likely created through NHEJ, MMEJ, or MMBIR at sites of little to no microhomology. Some notable examples of recurrent inverted CNVs have 1 junction between closely spaced repeats (with the alternate allele composition illustrated in Fig 3C or in Fig 3D) and the other junction created through microhomology or blunt ended ligation [34,35]. The sizes of the DUP segments in the cases we examined ranged from as little as a few kb to thousands of kb. The triplicated segments also had a wide range of sizes. We found no examples in our limited survey where the presumptive palindromic arms were unambiguously intact. This finding is in contrast to what we recover for the yeast *SUL1* locus where only 3 of 92 sequenced junctions were consistent with secondary rearrangements [14], possibly reflecting differences in the number of cell divisions, selective pressures, and/or availability of different DNA repair pathways in the yeast experiments compared to human development.

Some DUP-TRP/INV-DUP events are flanked by a region that has become homozygous (AOH). These cases have allele compositions within and adjacent to the duplicated region, often extending through the adjacent telomere, that are consistent with a mitotic recombination event between the 2 homologues followed by segregation in mitosis (Fig 3E–3G and Table 1, examples 22–23). In these cases, authors invoke FoSTeS as the mechanism for this

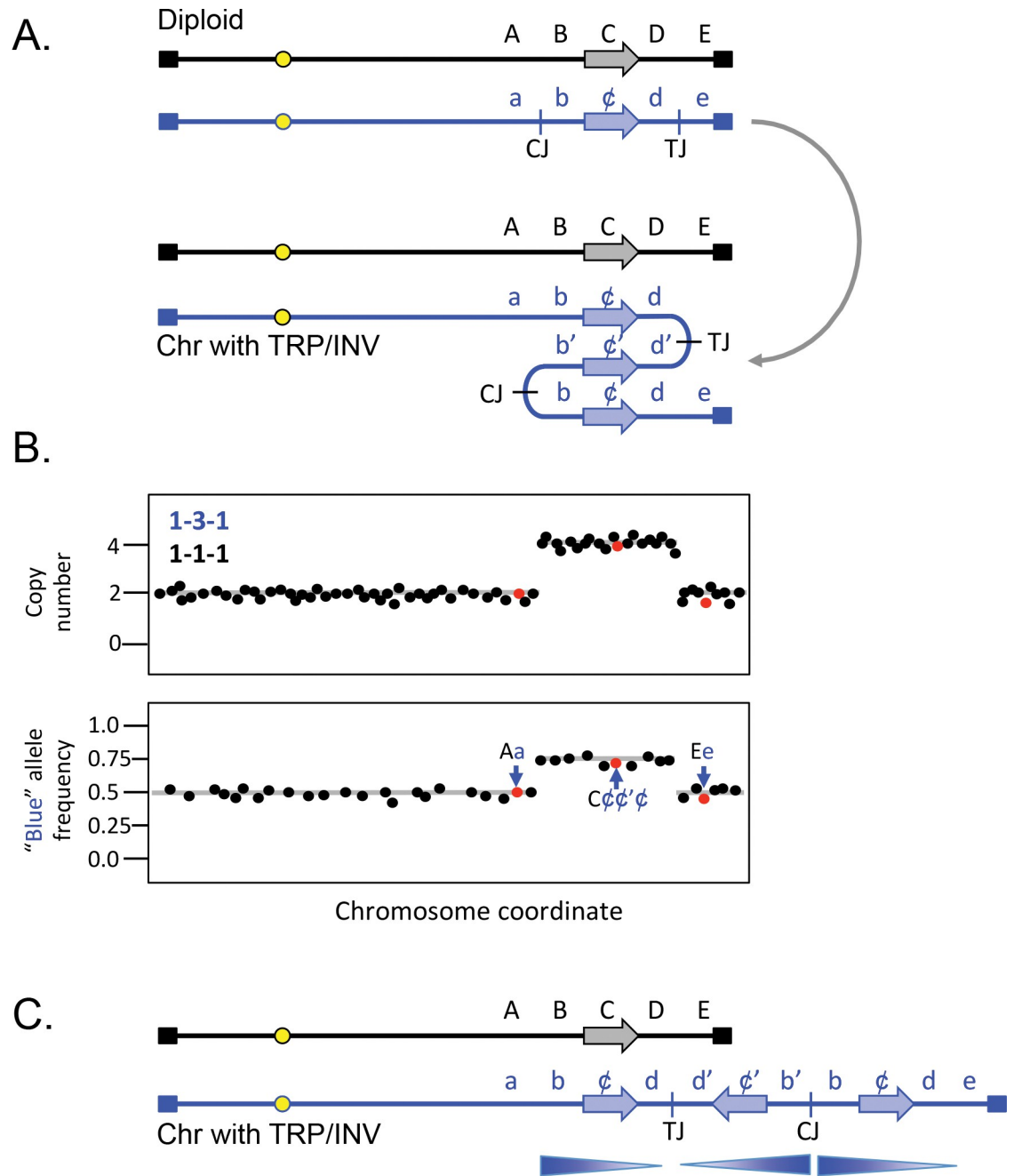


Fig 2. Inverted triplication. (A) A generic example of an inverted triplication in a diploid, affecting the blue chromosome, with SNPs indicated in upper and lower case letters (lower case "c" being depicted as ϕ for clarity). The horizontal arrow represents a potential coding sequence. CJ and TJ refer to the potential centers of the inversion junctions (centromere-proximal and telomere-proximal junctions) identified after the inversion and triplication of the segment containing the b, c, and d SNPs. The derived chromosome is shown folded back on itself to emphasize the triplication and the inverted center copy. (B) Top; expected copy number results (using either aCGH or read depth) of the diploid after triplication of the b–d region. Bottom; allele frequencies for SNPs unique to the blue chromosome. (C) Linear representation of the 2 homologues after triplication affecting the blue chromosome. Arrows indicate the orientation of the 3 segments involved in the triplication. Notice that the right end of the chromosome remains intact. aCGH, array comparative genome hybridization.

<https://doi.org/10.1371/journal.pgen.1011091.g002>

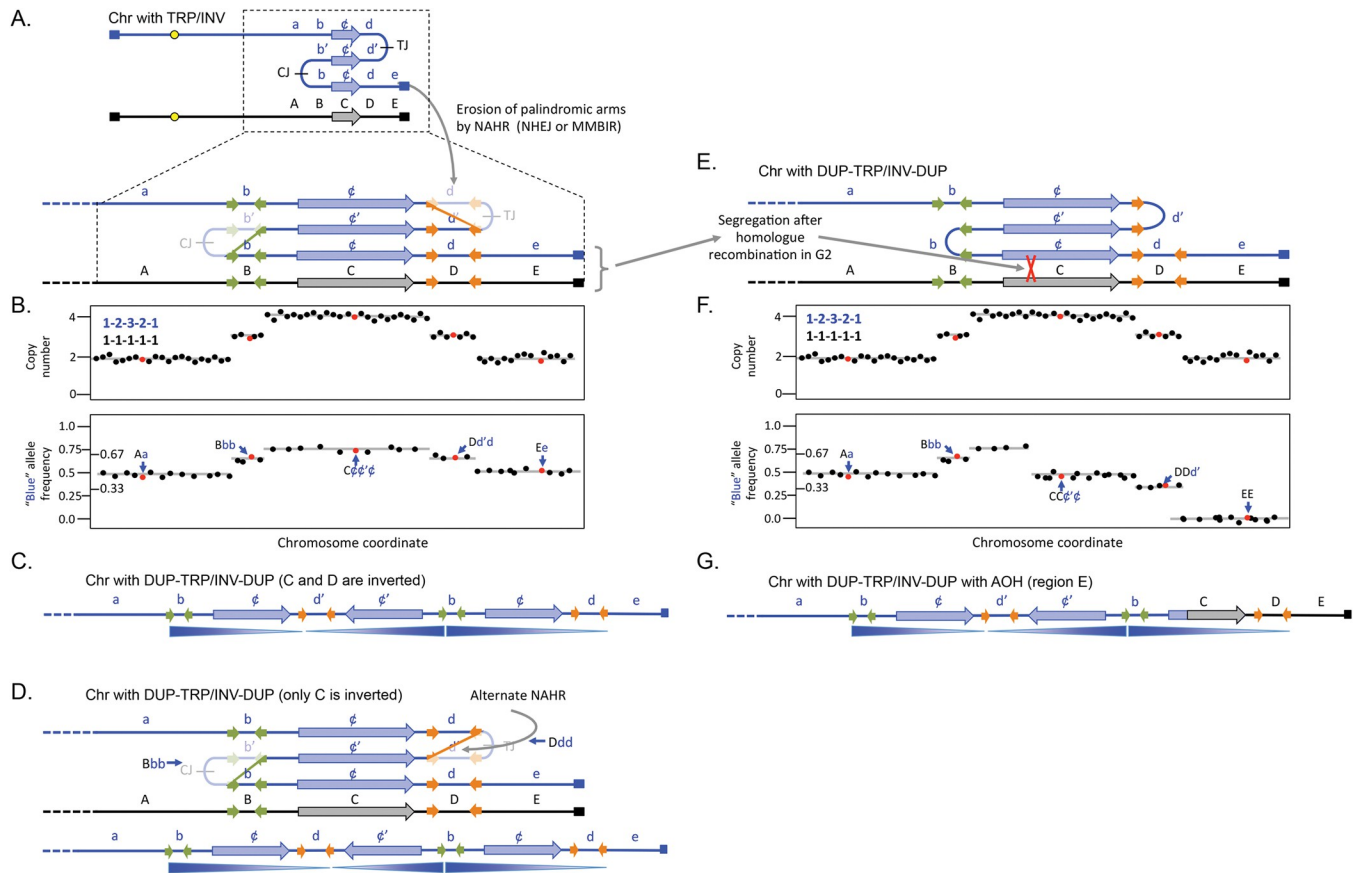


Fig 3. DUP-TRP/INV-DUP with and without adjacent AOH. (A) The same chromosome illustrated in Fig 2 is expanded to show potential short regions of inverted homology such as SINES or Alu sequences (green and orange horizontal arrows). After the triplication, pairs of orange and green repeats are now found in direct orientation and serve as the sites for non-allelic homologous recombination or other forms of rearrangement. The original inverted junctions (CJ and TJ) have been lost and the region between the recombined repeats is reduced in copy number. (B) Top; expected copy number results (using either aCGH or read depth). Bottom; allele frequencies for SNPs unique to the blue chromosome. Notice that the d region is now at 3 copies total with one copy of the c and d regions in inverted orientation. (C) Linear representation of the blue homologue after triplication. (D) An alternate recombination event at the orange repeats produces a chromosome with the same copy number and allele frequency profiles, but in this case, only the c region is inverted. (E and F) After the erosion of palindromes shown in (A), a secondary event of mitotic homologous recombination and subsequent segregation produces the same pattern of copy number estimates, but the telomeric region has become homozygous for the black chromosome E allele and other c and d SNPs have been reduced. (G) Linear representation of the blue homologue after homologous mitotic recombination that replaced the end of the rearranged chromosome with alleles from the black homologue. See S1 Fig for alternate illustrations of the rearrangement events shown in (A) and (D). aCGH, array comparative genome hybridization; AOH, absence of heterozygosity.

<https://doi.org/10.1371/journal.pgen.1011091.g003>

type of event, suggesting that the 3' end of the lagging strand from the stalled fork visits the oppositely oriented repeat to continue synthesis before jumping to the homologue to complete replication of the chromosome (e.g., [36]), all occurring within in a single division cycle. In contrast, we are suggesting that homologue exchange could be an outcome of the reduction of one of the palindromic arms or could be an independent event that is executed at a subsequent division cycle.

Inverted duplications associated with deletions

The dense distribution of repetitive elements in the human genome can precipitate situations where inverted junctions at a distance can lead to other types of secondary outcomes. When the repeats are more widely spaced (left-most orange arrow in Fig 4A, some distance from the pair of closely spaced inverted repeats diagrammed in Fig 3), then other potential

Table 1. Examples from the human literature of categories of local inverted SVs.

Example	Category of CNV	Homology at junction(s)	Chromo-somal location	Gene/syndrome	Reference	Case ID	Copy number on the affected chromosome
1	DUP-TRP/INV-DUP	4 bp; RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	BAB1612/P374	1-2-3-2-1
2	DUP-TRP/INV-DUP	47 bp (AluSx); RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	BAB3698	1-2-3-2-1
3	DUP-TRP/INV-DUP	0 bp; RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	P558	1-2-3-2-1
4	DUP-TRP/INV-DUP	2 bp; RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	P642	1-2-3-2-1
5	DUP-TRP/INV-DUP	1 bp; RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	P674	1-2-3-2-1
6	DUP-TRP/INV-DUP	3 bp; RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	P820	1-2-3-2-1
7	DUP-TRP/INV-DUP	0 bp; RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	P842	1-2-3-2-1
8	DUP-TRP/INV-DUP	2 bp; RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	P1150	1-2-3-2-1
9	DUP-TRP/INV-DUP	1 bp; RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	P1389	1-2-3-2-1
10	DUP-TRP/INV-DUP	2 bp; RE or LCR	Xq22.2	PMB/Pelizaeus-Merzbacher disease	[34]	P1407	1-2-3-2-1
11	DUP-TRP/INV-DUP	2/4 bp; IRs	Xq28	MECP2/ID and DD	[35]	BAB2769	1-2-3-2-1
12	DUP-TRP/INV-DUP	3 bp; IRs	Xq28	MECP2/ID and DD	[35]	BAB2772	1-2-3-2-1
13	DUP-TRP/INV-DUP	2 bp; IRs	Xq28	MECP2/ID and DD	[35]	BAB2796	1-2-3-2-1
14	DUP-TRP/INV-DUP	2 bp; IRs	Xq28	MECP2/ID and DD	[35]	BAB2980	1-2-3-2-1
15	DUP-TRP/INV-DUP	1 bp; IRs	Xq28	MECP2/ID and DD	[35]	BAB2797	1-2-3-2-1
16	DUP-TRP/INV-DUP	? bp; IRs	Xq28	MECP2/ID and DD	[35]	BAB2801	1-2-3-2-1
17	DUP-TRP/INV-DUP	0 bp; IRs	Xq28	MECP2/ID and DD	[35]	BAB2805	1-2-3-2-1
18	DUP-TRP/INV-DUP	? bp; IRs	Xq28	MECP2/ID and DD	[35]	BAB3114	1-2-3-2-1
19	DUP-TRP/INV-DUP	2/3 bp; IRs	Xq28	MECP2/ID and DD	[49]	BAB2727	1-2-3-2-1
20	DUP-TRP/INV-DUP	LCR; 0 bp	Xq22.2	PLP1/Pelizaeus-Merzbacher disease	[50]	9 YO (M)	1-2-3-2-1
21	DUP-TRP/INV-DUP	not reported	Xq28	MECP2/ID and ASD	[51]	Patient 1 (M)	1-2-3-2-1
22	DUP-TRP/INV-DUP with AOH	4 bp; 2 bp	14q23	Temple Syndrome	[36]	BAB7004 (M)	1-2-3-2-1
23	DUP-TRP/INV-DUP with AOH	4 bp; 2 bp	1q42	Tetralogy of Fallot and ID	[52]	25 YO (M)	1-2-3-2-1
24	DUP/INV-NML-DEL	LCR; 2 bp	Xq28	Factor 8 /Haemophilia A	[37]	Proband 3 (F)	1-2-1-0-1
25	DUP/INV-NML-DEL	1 bp; LTRs	Xq28	Factor 8 /Haemophilia A	[37]	Proband 4 (F)	1-0-1-2-1
26	DUP-TRP/INV-DUP with TEL DEL	5 bp; 4 bp	13q	Microphthalmia, anophthalmia and coloboma	[53]	newborn	1-2-3-2-0
27	DUP/INV-NML with TEL DEL	3 bp; de novo telomere addition	10qter	DD, ASD, multiple phenotypes	[38]	Patient 5 (F)	1-2-1-0
28	DUP/INV-NML with TEL DEL	6 bp	13q34	ID and DD	[54]	Patient 1	1-2-1-0
29	DUP/INV-NML with TEL DEL	4 bp	8p23.2	ID and DD	[54]	Patient 4	0-1-2-1
30	DUP/INV-NML with TEL DEL	0 bp	11q25	IGSF9B/ID and DD	[54]	Patient 9	1-2-1-0

(Continued)

Table 1. (Continued)

Example	Category of CNV	Homology at junction(s)	Chromo-somal location	Gene/syndrome	Reference	Case ID	Copy number on the affected chromosome
31	DUP/INV-NML with TEL DEL	3 bp	4p	not reported	[54]	case 1	1-2-1-0
32	DUP/INV-NML with TEL DEL	2 bp	4q	not reported	[54]	case 2	1-2-1-0
33	DUP/INV-NML with TEL DEL	0 bp	9p	not reported	[54]	case 4	0-1-2-1
34	DUP/INV-NML with TEL DEL	2 bp	10q	not reported	[54]	case 5	1-2-1-0
35	DUP/INV-NML with TEL DEL	3 bp	11q	not reported	[54]	case 6	1-2-1-0
36	DUP/INV-NML with TEL DEL	3 bp	18p	not reported	[54]	case 7	0-1-2-1
37	DUP/INV-NML with TEL DEL	Olfactory receptor gene cluster	8p	not reported	[54]	case 9	0-1-2-1
38	DUP/INV-NML with TEL DEL	Olfactory receptor gene cluster	8p	not reported	[44]	case 10	0-1-2-1
39	DUP/INV-NML-DUP	LCR; 5 bp	Xq28	Factor 8 /Haemophilia A	[37]	Proband 2 (M)	1-2-1-2-1
40	DUP/INV-NML-DUP with TEL DEL	6 bp; 2 bp	13qter	Tetralogy of Fallot; DD	[38]	Patient 1 (F)	1-2-1-2-1-0
41	DUP/INV-NML-DUP with TEL DEL	2 bp; 0 bp	4pter	Wolf-Hirschhorn syndrome and DD	[38]	Patient 2 (M)	0-1-2-1-2-1
42	DUP/INV-NML-DUP with TEL DEL	14 bp; 0 bp	4pter	generalized hypotonia	[38]	Patient 3 (F)	0-1-2-1-2-1
43	DUP/INV-NML-DUP with TEL DEL	3 bp; 1 bp	4pter	Developmental delay	[38]	Patient 4 (F)	0-1-2-1-2-1
44	DUP/INV-NML-DEL-NML with TEL DEL	4 bp; 3 bp	1q43	AKT3; CNST; KMO/ID and DD	[54]	Patient 11	1-2-1-0-1-0
45	DEL-NML-DUP/INV-NML with TEL-DEL	4 bp; 4 bp	3q29	ID and DD	[54]	Patient 7	1-0-1-2-1-0

AOH, absence of heterozygosity; ASD, autism spectrum disorder; DD, developmental delay; DEL, deletion; DUP, duplication; DUP/INV, inverted duplication; ID, intellectual disability; IR, inverted repeat; LCR, low copy repeat; LTR, long terminal repeat; NML, normal copy number; RE, repeated element; TEL DEL, telomere deletion; TRP/INV, inverted triplication.

<https://doi.org/10.1371/journal.pgen.1011091.t001>

recombination partners can be involved in the rearrangement of the palindrome. These duplications are illustrated using NAHR at repetitive elements, but similar structures could be generated at non-repetitive sequences through repair of DNA breaks by NHEJ, MMEJ, or MMBIR. This particular outcome has recently been reported at the Factor 8 locus where NAHR at long, highly homologous inverted repeats generates one of the junctions and the second junction occurs at regions of little to no homology [37]. It produces the pattern of a deletion and inverted duplication separated by a stretch of copy-neutral DNA in between (1-0-1-2-1; Figs 4A and 5 and Table 1, examples 24, 25, 39, 44, 45).

The final example, also illustrated as occurring on the X chromosome, is a common form of human SV that occurs in telomeric regions. The telomere-proximal palindrome is resolved by NAHR, NHEJ, or MMEJ, but the centromere-proximal palindromic junction is lost when a break is capped by a new telomere or captures a telomere from another chromosome end (Fig 4B and Table 1, examples 26–38, 40–45). Shimojima Yamamoto and colleagues [38] described an example of an inverted duplication on chromosome 10 that was accompanied by the loss of one of the palindromic arms, the presumptive telomere-proximal junction and all distal sequences.

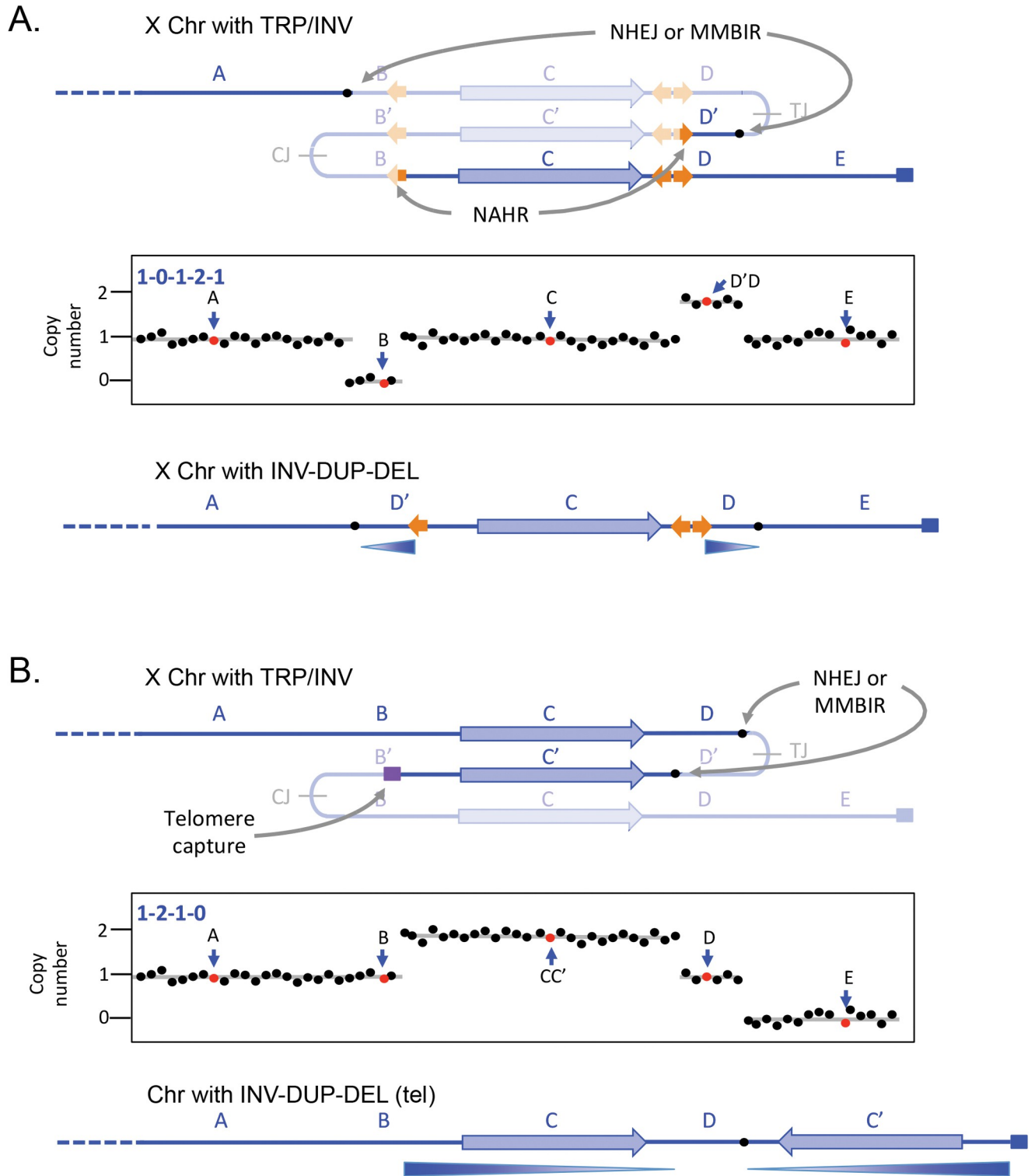


Fig 4. Direct and inverted duplications. (A) In this example, modeled after the Factor 8 locus on the X chromosome in humans, additional repeats provide other opportunities for rearrangements of the triplicated locus that remove the centromere-proximal junction. The telomeric-proximal junction is removed by NHEJ or MMBIR at short regions of microhomology. This pattern of palindrome erosion results in a deletion and an inverted duplication separated by a copy-neutral segment of chromosome. (B) A failed recombination/MMBIR attempt to erode the centromere-proximal junction leaves a dsDNA break that acquires or captures a new telomere. It results in the complete loss of sequences from the point of the inverted duplication to the end of the chromosome. See S1 Fig for alternate illustrations of the 2 rearrangement events. MMBIR, microhomology-mediated break-induced replication; NHEJ, non-homologous end joining.

<https://doi.org/10.1371/journal.pgen.1011091.g004>

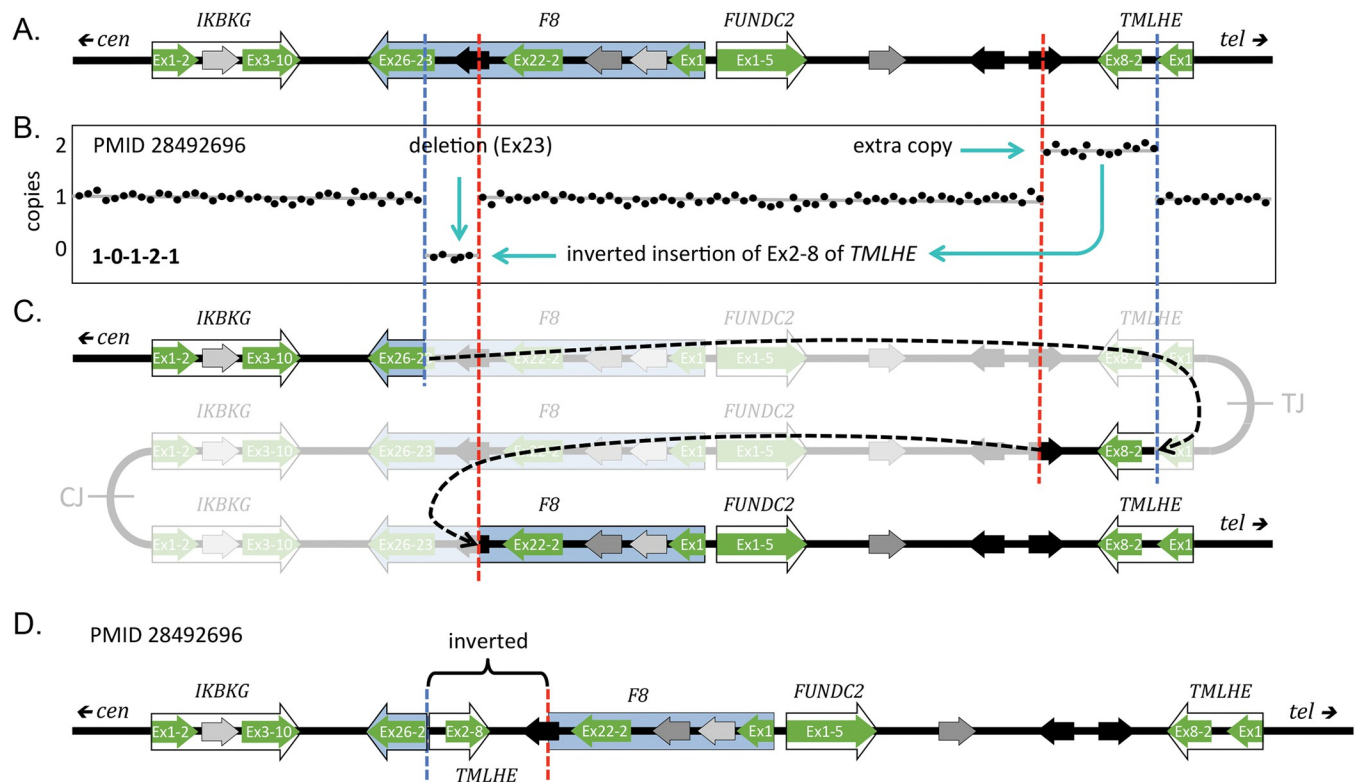


Fig 5. Rearrangement of an inverted triplication gives rise to a Factor 8 mutation. (A) The original sequence near the Xq telomere; F8 is the Factor 8 gene, highlighted in blue. Diagram is adapted from [37]. Exons of 3 adjacent genes (green arrows) and 3 different LCRs (black and gray arrows) are highlighted. (B) Stylized aCGH data for the male patient with Factor 8 deficiency (reported in [37] as PMID 28492696). Cyan arrows indicate the deleted and inverted duplicate regions. (C) Hypothesized initial inverted triplication with CJ and TJ palindromic junctions. Red and blue dashed lines indicate rearrangement junctions produced through NHEJ or MMEJ and NAHR at the 2 oppositely oriented black LCRs. Black dashed arrows indicate the joining events; the grayed-out regions indicate the regions lost during the secondary rearrangements. (D) The final structure of the F8 region of the patient in PMID 28492696 with a deletion of exon 23 and an inverted duplication of exons 2–8 of *TMMLHE* and flanking regions. aCGH, array comparative genome hybridization; LCR, low copy repeat; MMEJ, microhomology-mediated end joining; NAHR, non-allelic homologous recombination; NHEJ, non-homologous end joining.

<https://doi.org/10.1371/journal.pgen.1011091.g005>

Source of the initiating TRP/INV structures

We propose that all of the above CNVs share the same starting point: a chromosome with a segment present as an inverted triplication where the center copy is inverted between 2 directly repeated segments with palindromes at the junctions (Figs 3 and 4). All of the different inverted CNVs can be created through rearrangements of the TRP/INV CNVs by well-characterized pathways, but what is the mechanism for forming the initial, inverted triplications and what is the nature of the initial palindrome?

Clues into the mechanism that generates TRP/INV CNVs has come from our studies in yeast. We have been investigating how yeast adapts to growth in continuous culture under conditions of sulfate limitation [12,14,39]. Over the course of 50 to 200 generations, variants arise that have an increased fitness due to an interstitial inverted triplication of *SUL1*, the gene that codes for the primary sulfate transporter. Analysis of nearly 100 events revealed that the amplicons always contain the adjacent origin of replication (*ARS228*) and centromere- and telomere-proximal palindromic junctions that map to preexisting short (5 to 6 bp) inverted repeats in the genome that are interrupted by 40 to 80 bp of non-palindromic DNA [13]. The model that we proposed and experimentally tested involves a replication error and not a double-stranded break as the initiating event [13]. At stalled replication forks the nascent leading strands at divergent replication forks switch to the lagging strand templates and become

continuous with the nascent lagging strands (Fig 6A and 6B). The aberrant structure is expelled from the chromosome by a fork from an adjacent replicon and, after replication in the next cell cycle, re-integrates into the undamaged chromosome by homologous recombination (Fig 6C–6E). Because the model is dependent on the presence of an origin of replication in the amplified segment and on short inverted repeats, we named this mechanism ODIRA [13].

We propose that this same aberrant replication pathway occurs in the human genome where it generates the TRP/INV substrates that give rise to all manner of local inverted CNVs through secondary rearrangements (Fig 1). Unlike the human genome that is composed of roughly 50% repeated sequences [40], the region of the yeast genome that contains *SUL1* has no significant repeated sequences (such as Tys, deltas, or tRNA genes) so the options for secondary rearrangements of the palindromic junctions are limited. However, in roughly 5% to 10% of cases of yeast inverted triplications, we recovered new junctions produced through regions of microhomology [11,13] that increase the spacing between the inverted arms (from 40 to 80 bp to 1.1 to 34.9 kb). While we originally proposed that the 2 junctions are created by simultaneous errors at the 2 diverging replication forks flanking *SUL1*, we have new evidence that the 2 template switches can occur in different cell cycles, but, after recombination with the chromosome, produce identical inverted triplication products [14].

In Figs 3, 4, and 6, the extrachromosomal inverted duplicated circular intermediate that arose through ODIRA is shown as re-integrating into the chromosome from which it arose. However, it is also possible for the intermediate to recombine with the homologue, producing a triplication with a 2:1 ratio of SNPs from the 2 homologues. When this event occurs in the germ line, it is possible to see the contribution from both homologues of that parent in the CNV. While most studies do not go into this level of analysis, some reports of human inverted triplications have this 2:1 ratio of SNPs from the contributing parent (for example, [41–43]).

Discussion

Inverted duplications that end in terminal deletions are easily and economically explained by Barbara McClintock's BFB model [2,44] but are also compatible with replication-based mechanisms [45]. However, interstitial inverted SVs where the end of the chromosome remains intact require another explanation as the sister chromatid fusion that occurs during BFB results in the loss of sequences distal to the point of fusion. All inverted CNVs presented in this work (including inverted duplications with terminal deletion) can be explained by secondary rearrangements of inverted triplications produced through a strand switching mechanism between the 2 strands at a replication fork. It is the very nature of the palindromic structure of the inverted triplications that makes them prone to rearrangement [24]: in the course of subsequent replication cycles, processes such as NHEJ, MMEJ, NAHR, or MMBIR can repair breaks arising from the unstable palindromes.

We propose that the initial formation of inverted triplications occurs by a replication error in which the leading strand at a replication fork switches to the lagging strand template at very short, interrupted inverted repeats [13] which are found at very high density in all genomes. The size of the interruptions at ODIRA junctions in yeast is consistent with the length of the single-stranded gaps between Okazaki fragments on the lagging strand [14]. The presence of a single-stranded gap on the lagging strand provides the opportunity for the leading strand to switch to the lagging strand template. As Okazaki fragment size is the same across eukaryotes, we would expect to find similar junctions in human inverted triplications. All of the TRP/INV junctions we found in the literature were large enough to include multiple replication origins but had considerably larger interruptions in their palindromic arms, similar to the secondary rearrangements we find in approximately 5% of yeast junctions. These results suggest that in

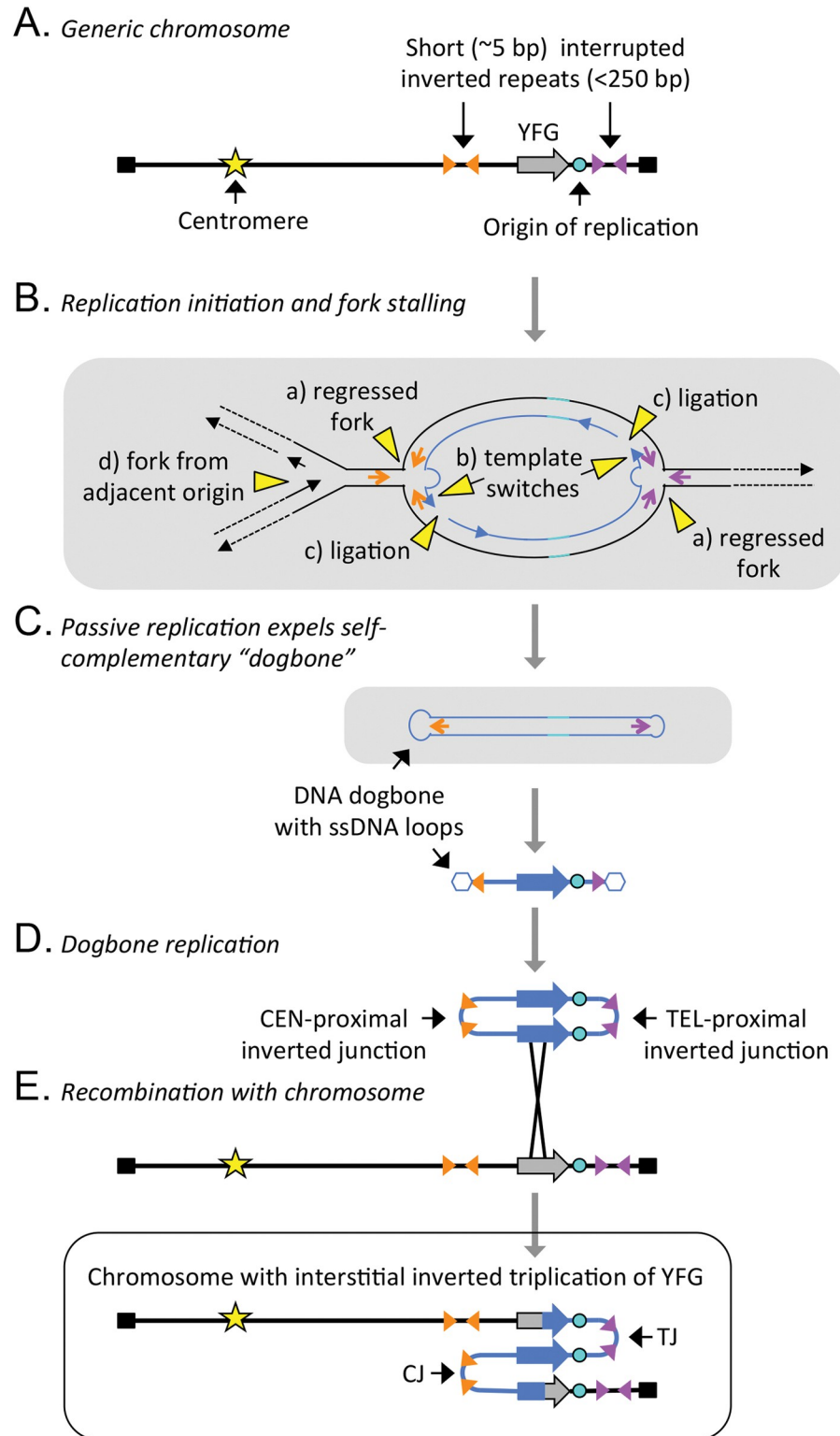


Fig 6. Mechanism that generates an inverted triplication (ODIRA)—modified from Fig 1 in Martin and colleagues [14]. (A) A chromosome with a segment that will be amplified, containing “your favorite gene” (YFG), an origin of replication, and adjacent short (~5 nt) inverted repeats separated by 40–80 nt of sequence. These interrupted inverted repeats mark the sites of the centromere- and telomere-proximal junctions (CJ and TJ, respectively) after inverted triplication. (B) Process of template switching between the leading strand and its migration to complementary

sequences on the lagging strand template. After extension of the 3' end of the leading strand, it becomes ligated to the adjacent Okazaki fragment. If both forks undergo the same event, it results in a closed loop of self-complementary DNA. (C) Displacement of the self-complementary loop can be achieved by branch migration ahead of an incoming replication fork. (D) The resulting expelled molecule (dogbone) can replicate in the next cell cycle to produce a dimeric circular molecule with 2 copies of the expelled DNA in inverted orientation with junctions formed from the 2 inverted repeats (C) and (TJ). (E) In a subsequent cell cycle, recombination of the dimeric circle into the original location on the chromosome leads to the inverted triplication. Note that the junctions CJ and TJ retain the sequence of the 2 short inverted repeats where the template switching occurred. ODIRA, origin-dependent inverted-repeat amplification.

<https://doi.org/10.1371/journal.pgen.1011091.g006>

the human examples, the palindromic junctions had already undergone secondary rearrangements to increase the size of the interruption to the point that the palindromic arms were no longer inducing unstable secondary structures.

The FoSTeS model [7] has also been suggested as a mechanism for various forms of SVs. While FoSTeS and ODIRA both propose template switching of one of the nascent strands at a replication fork, they differ in which strand is switching templates and the location of the new template. FoSTeS was inspired by work in *Escherichia coli* where Lac⁺ direct amplicons arise under conditions of stress. Mutations implicate the flap endonuclease function of DNA polymerase I and the lagging strand template [46] in their formation. The model, devised to account for the microhomology at the novel junctions, proposes that the 3' end of an Okazaki fragment makes the jump between the *E. coli* chromosome and the F' conjugating plasmid and back again to the chromosome. But in eukaryotes, because Okazaki fragments are only approximately 150 bp (>10× times shorter than is found in *E. coli*), it is perhaps surprising that the 3' end of such a short strand can sequentially invade a different replication fork, dissociate from it after synthesizing long stretches of DNA and then return to the original fork before completing the nascent lagging strand, all within a single S phase. In the ODIRA model, the lagging strand template is in close proximity to the 3' end of the leading strand, making template switching of the leading strand within a single fork a more energetically feasible mechanism compared to FoSTeS. Despite FoSTeS being widely cited as a likely explanation for various inverted segmental variants in humans, we have not been able to find any published accounts with corroborating experimental evidence in model eukaryotes for FoSTeS. One report of intrachromosomal template switching (ICTS) in yeast by Tsaponina and Haber [47] shares some similarities with FoSTeS but differs in that the length of sequences synthesized after the jump to a new chromosomal site is very short. In addition, they did not investigate which strand at the replication fork was involved in the template switch.

In addition to providing a unifying model for a variety of congenital inverted SVs in humans, we propose that ODIRA may also be a significant pathway to forming SVs in cancer cells and may provide an experimental system to dissect this mechanism. Reports of genome-wide palindrome analysis reveal a rise in palindromic sequences in a variety of cancers [20–25] but the length of spacers and their association with amplified segments remains unexplored. Are these palindromes evidence of inverted triplications? Are they a cause or consequence of some other process that is permissive for genome rearrangements? If they are formed through the same replication error we are proposing for yeast *SUL1*, then what microenvironment/stress conditions cause the aberrant template switching and are there interventions that can block these replication products or their processing? Long-read sequencing data on tumors in early stages of their development, as well as genetic dissection in models such as yeast, will be invaluable for answering these important questions.

Supporting information

S1 Fig. Alternate depictions for secondary rearrangements of inverted triplications. 3init, 3A and 3D refer to images from Fig 3. Init = inverted triplication before rearrangement. 4init,

4A and 4B refer to images from Fig 4. Init = inverted triplication before rearrangement. Brackets indicate rearrangement junctions created by NAHR, NHEJ, or MMBIR. The grayed-out regions indicate the regions of the inverted triplication that are deleted during the secondary rearrangements.

(PDF)

Acknowledgments

We are grateful for the discussions with and thoughtful comments of our colleagues Elizabeth Kwan, Rebecca Martin, Gina Alvino, Amy Moore, Joe Armstrong, Evan Eichler, and Xavi Guittart. We apologize in advance for overlooking anyone's work that is relevant to the discussion here and encourage you to contact BJB with your thoughts and comments. There were many more examples that could have been included in our analysis but we limited our analysis to those that had well-characterized junctions. Again, we apologize and look forward to hearing more about these other patients.

Author Contributions

Conceptualization: Bonita J. Brewer, Maitreya J. Dunham, M. K. Raghuraman.

Data curation: Bonita J. Brewer, Maitreya J. Dunham.

Formal analysis: Bonita J. Brewer, Maitreya J. Dunham, M. K. Raghuraman.

Funding acquisition: Bonita J. Brewer, Maitreya J. Dunham.

Investigation: Bonita J. Brewer, Maitreya J. Dunham, M. K. Raghuraman.

Methodology: Bonita J. Brewer, Maitreya J. Dunham, M. K. Raghuraman.

Project administration: Bonita J. Brewer, Maitreya J. Dunham, M. K. Raghuraman.

Resources: Bonita J. Brewer, Maitreya J. Dunham, M. K. Raghuraman.

Supervision: Bonita J. Brewer, Maitreya J. Dunham, M. K. Raghuraman.

Visualization: Bonita J. Brewer, M. K. Raghuraman.

Writing – original draft: Bonita J. Brewer, Maitreya J. Dunham, M. K. Raghuraman.

Writing – review & editing: Bonita J. Brewer, Maitreya J. Dunham, M. K. Raghuraman.

References

1. Burssed B, Zamariolli M, Bellucco FT, Melaragno MI. Mechanisms of structural chromosomal rearrangement formation. *Mol Cytogenet.* 2022; 15(1):23. Epub 20220614. <https://doi.org/10.1186/s13039-022-00600-6> PMID: 35701783; PubMed Central PMCID: PMC9199198.
2. McClintock B. The Stability of Broken Ends of Chromosomes in Zea Mays. *Genetics.* 1941; 26(2):234–282. Epub 1941/03/01. <https://doi.org/10.1093/genetics/26.2.234> PMID: 17247004; PubMed Central PMCID: PMC1209127.
3. Moore JK, Haber JE. Cell cycle and genetic requirements of two pathways of nonhomologous end-joining repair of double-strand breaks in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 1996; 16(5):2164–2173. <https://doi.org/10.1128/MCB.16.5.2164> PMID: 8628283; PubMed Central PMCID: PMC231204.
4. Ma JL, Kim EM, Haber JE, Lee SE. Yeast Mre11 and Rad1 proteins define a Ku-independent mechanism to repair double-strand breaks lacking overlapping end sequences. *Mol Cell Biol.* 2003; 23(23):8820–8828. <https://doi.org/10.1128/MCB.23.23.8820-8828.2003> PMID: 14612421; PubMed Central PMCID: PMC262689.
5. Sasaki M, Lange J, Keeney S. Genome destabilization by homologous recombination in the germ line. *Nat Rev Mol Cell Biol.* 2010; 11(3):182–195. Epub 20100218. <https://doi.org/10.1038/nrm2849> PMID: 20164840; PubMed Central PMCID: PMC3073813.

6. Hastings PJ, Ira G, Lupski JR. A microhomology-mediated break-induced replication model for the origin of human copy number variation. *PLoS Genet.* 2009; 5(1):e1000327. Epub 2009/01/31. <https://doi.org/10.1371/journal.pgen.1000327> PMID: 19180184; PubMed Central PMCID: PMC2621351.
7. Lee JA, Carvalho CM, Lupski JR. A DNA replication mechanism for generating nonrecurrent rearrangements associated with genomic disorders. *Cell.* 2007; 131(7):1235–1247. Epub 2007/12/28. <https://doi.org/10.1016/j.cell.2007.11.037> [pii] PMID: 18160035.
8. Sakofsky CJ, Ayyar S, Deem AK, Chung WH, Ira G, Malkova A. Translesion Polymerases Drive Microhomology-Mediated Break-Induced Replication Leading to Complex Chromosomal Rearrangements. *Mol Cell.* 2015; 60(6):860–872. Epub 20151206. <https://doi.org/10.1016/j.molcel.2015.10.041> PMID: 26669261; PubMed Central PMCID: PMC4688117.
9. Jourdy Y, Bardel C, Fretigny M, Diguët F, Rollat-Farnier PA, Mathieu ML, et al. Complete characterisation of two new large Xq28 duplications involving F8 using whole genome sequencing in patients without haemophilia A. *Haemophilia.* 2022; 28(1):117–124. Epub 20210904. <https://doi.org/10.1111/hae.14402> PMID: 34480810.
10. Tokoro M, Tamura S, Suzuki N, Kakiyama M, Hattori Y, Odaira K, et al. Aberrant X chromosomal rearrangement through multi-step template switching during sister chromatid formation in a patient with severe hemophilia A. *Mol Genet Genomic Med.* 2020; 8(9):e1390. Epub 20200705. <https://doi.org/10.1002/mgg3.1390> PMID: 32627361; PubMed Central PMCID: PMC7507428.
11. Dery T, Chatron N, Alqahtani A, Pugeat M, Till M, Edery P, et al. Follow-up of two adult brothers with homozygous CEP57 pathogenic variants expands the phenotype of Mosaic Variegated Aneuploidy Syndrome. *Eur J Med Genet.* 2020; 63(11):104044. Epub 20200828. <https://doi.org/10.1016/j.ejmg.2020.104044> PMID: 32861809.
12. Brewer BJ, Payen C, Di Rienzi SC, Higgins MM, Ong G, Dunham MJ, et al. Origin-Dependent Inverted-Repeat Amplification: Tests of a Model for Inverted DNA Amplification. *PLoS Genet.* 2015; 11(12):e1005699. <https://doi.org/10.1371/journal.pgen.1005699> PMID: 26700858; PubMed Central PMCID: PMC4689423.
13. Brewer BJ, Payen C, Raghuraman MK, Dunham MJ. Origin-dependent inverted-repeat amplification: a replication-based model for generating palindromic amplicons. *PLoS Genet.* 2011; 7(3):e1002016. Epub 2011/03/26. <https://doi.org/10.1371/journal.pgen.1002016> PMID: 21437266; PubMed Central PMCID: PMC3060070.
14. Martin R, Espinoza CY, Large CRL, Rosswork J, Van Bruinisse C, Miller AW, et al. (2024) Template switching between the leading and lagging strands at replication forks generates inverted copy number variants through hairpin-capped extrachromosomal DNA. *PLoS Genet* 20(1): e1010850. <https://doi.org/10.1371/journal.pgen.1010850>.
15. Akgun E, Zahn J, Baumes S, Brown G, Liang F, Romanienko PJ, et al. Palindrome resolution and recombination in the mammalian germ line. *Mol Cell Biol.* 1997; 17(9):5559–5570. <https://doi.org/10.1128/MCB.17.9.5559> PMID: 9271431; PubMed Central PMCID: PMC232404.
16. Cunningham LA, Cote AG, Cam-Ozdemir C, Lewis SM. Rapid, stabilizing palindrome rearrangements in somatic cells by the center-break mechanism. *Mol Cell Biol.* 2003; 23(23):8740–8750. <https://doi.org/10.1128/MCB.23.23.8740-8750.2003> PMID: 14612414; PubMed Central PMCID: PMC262683.
17. Ganapathiraju MK, Subramanian S, Chaparala S, Karunakaran KB. A reference catalog of DNA palindromes in the human genome and their variations in 1000 Genomes. *Hum Genome Var.* 2020; 7(1):40. Epub 20201120. <https://doi.org/10.1038/s41439-020-00127-5> PMID: 33298903; PubMed Central PMCID: PMC7680136.
18. Lauer S, AVECILLA G, Spealman P, Sethia G, Brandt N, Levy SF, et al. Single-cell copy number variant detection reveals the dynamics and diversity of adaptation. *PLoS Biol.* 2018; 16(12):e3000069. Epub 20181218. <https://doi.org/10.1371/journal.pbio.3000069> PMID: 30562346; PubMed Central PMCID: PMC6298651.
19. Schuy J, Grochowski CM, Carvalho CMB, Lindstrand A. Complex genomic rearrangements: an underestimated cause of rare diseases. *Trends Genet.* 2022; 38(11):1134–1146. Epub 20220709. <https://doi.org/10.1016/j.tig.2022.06.003> PMID: 35820967; PubMed Central PMCID: PMC9851044.
20. Guenthoer J, Diede SJ, Tanaka H, Chai X, Hsu L, Tapscott SJ, et al. Assessment of palindromes as platforms for DNA amplification in breast cancer. *Genome Res.* 2012; 22(2):232–245. Epub 20110713. <https://doi.org/10.1101/gr.117226.110> PMID: 21752925; PubMed Central PMCID: PMC3266031.
21. Marotta M, Onodera T, Johnson J, Budd GT, Watanabe T, Cui X, et al. Palindromic amplification of the ERBB2 oncogene in primary HER2-positive breast tumors. *Sci Rep.* 2017; 7:41921. Epub 20170217. <https://doi.org/10.1038/srep41921> PMID: 28211519; PubMed Central PMCID: PMC5314454.
22. Murata MM, Giuliano AE, Tanaka H. Genome-Wide Analysis of Palindrome Formation with Next-Generation Sequencing (GAPF-Seq) and a Bioinformatics Pipeline for Assessing De Novo Palindromes in

- Cancer Genomes. *Methods Mol Biol.* 2023; 2660:13–22. https://doi.org/10.1007/978-1-0716-3163-8_2 PMID: 37191787.
23. Neiman PE, Elsaesser K, Loring G, Kimmel R. Myc oncogene-induced genomic instability: DNA palindromes in bursal lymphomagenesis. *PLoS Genet.* 2008; 4(7):e1000132. Epub 20080718. <https://doi.org/10.1371/journal.pgen.1000132> PMID: 18636108; PubMed Central PMCID: PMC2444050.
 24. Svetec Miklenic M, Svetec IK. Palindromes in DNA-A Risk for Genome Stability and Implications in Cancer. *Int J Mol Sci.* 2021; 22(6). Epub 20210311. <https://doi.org/10.3390/ijms22062840> PMID: 33799581; PubMed Central PMCID: PMC7999016.
 25. Tanaka H, Bergstrom DA, Yao MC, Tapscott SJ. Large DNA palindromes as a common form of structural chromosome aberrations in human cancers. *Hum Cell.* 2006; 19(1):17–23. <https://doi.org/10.1111/j.1749-0774.2005.00003.x> PMID: 16643603.
 26. Liu Z, Roberts R, Mercer TR, Xu J, Sedlazeck FJ, Tong W. Towards accurate and reliable resolution of structural variants for clinical diagnosis. *Genome Biol.* 2022; 23(1):68. Epub 20220303. <https://doi.org/10.1186/s13059-022-02636-8> PMID: 35241127; PubMed Central PMCID: PMC8892125.
 27. Mahmoud M, Gobet N, Cruz-Davalos DI, Mounier N, Dessimoz C, Sedlazeck FJ. Structural variant calling: the long and the short of it. *Genome Biol.* 2019; 20(1):246. Epub 20191120. <https://doi.org/10.1186/s13059-019-1828-7> PMID: 31747936; PubMed Central PMCID: PMC6868818.
 28. Mantere T, Neveling K, Pebrel-Richard C, Benoist M, van der Zande G, Kater-Baats E, et al. Optical genome mapping enables constitutional chromosomal aberration detection. *Am J Hum Genet.* 2021; 108(8):1409–1422. Epub 20210707. <https://doi.org/10.1016/j.ajhg.2021.05.012> PMID: 34237280; PubMed Central PMCID: PMC8387289.
 29. Tschernoster N, Erger F, Walsh PR, McNicholas B, Fistrek M, Habbig S, et al. Unraveling Structural Rearrangements of the CFH Gene Cluster in Atypical Hemolytic Uremic Syndrome Patients Using Molecular Combing and Long-Fragment Targeted Sequencing. *J Mol Diagn.* 2022; 24(6):619–631. Epub 20220408. <https://doi.org/10.1016/j.jmoldx.2022.02.006> PMID: 35398599.
 30. Spealman P, Burrell J, Gresham D. Inverted duplicate DNA sequences increase translocation rates through sequencing nanopores resulting in reduced base calling accuracy. *Nucleic Acids Res.* 2020; 48(9):4940–4945. <https://doi.org/10.1093/nar/gkaa206> PMID: 32255181; PubMed Central PMCID: PMC7229812.
 31. Jackson EK, Bellott DW, Cho TJ, Skaletsky H, Hughes JF, Pyntikova T, et al. Large palindromes on the primate X Chromosome are preserved by natural selection. *Genome Res.* 2021; 31(8):1337–1352. Epub 20210721. <https://doi.org/10.1101/gr.275188.120> PMID: 34290043; PubMed Central PMCID: PMC8327919.
 32. Lobachev KS, Shor BM, Tran HT, Taylor W, Keen JD, Resnick MA, et al. Factors affecting inverted repeat stimulation of recombination and deletion in *Saccharomyces cerevisiae*. *Genetics.* 1998; 148(4):1507–1524. <https://doi.org/10.1093/genetics/148.4.1507> PMID: 9560370; PubMed Central PMCID: PMC1460095.
 33. Svetec Miklenic M, Gatalica N, Matanovic A, Zunar B, Stafa A, Lisnic B, et al. Size-dependent antirecombinogenic effect of short spacers on palindrome recombinogenicity. *DNA Repair (Amst).* 2020; 90:102848. Epub 20200503. <https://doi.org/10.1016/j.dnarep.2020.102848> PMID: 32388488.
 34. Beck CR, Carvalho CM, Banser L, Gambin T, Stubbolo D, Yuan B, et al. Complex genomic rearrangements at the PLP1 locus include triplication and quadruplication. *PLoS Genet.* 2015; 11(3):e1005050. Epub 20150306. <https://doi.org/10.1371/journal.pgen.1005050> PMID: 25749076; PubMed Central PMCID: PMC4352052.
 35. Carvalho CM, Ramocki MB, Pehlivan D, Franco LM, Gonzaga-Jauregui C, Fang P, et al. Inverted genomic segments and complex triplication rearrangements are mediated by inverted repeats in the human genome. *Nat Genet.* 2011; 43(11):1074–1081. <https://doi.org/10.1038/ng.944> PMID: 21964572; PubMed Central PMCID: PMC3235474.
 36. Carvalho CMB, Coban-Akdemir Z, Hijazi H, Yuan B, Pendleton M, Harrington E, et al. Interchromosomal template-switching as a novel molecular mechanism for imprinting perturbations associated with Temple syndrome. *Genome Med.* 2019; 11(1):25. Epub 20190423. <https://doi.org/10.1186/s13073-019-0633-y> PMID: 31014393; PubMed Central PMCID: PMC6480824.
 37. Li Y, Ding B, Mao Y, Zhang H, Wang X, Ding Q. Tandem and inverted duplications in haemophilia A: Breakpoint characterisation provides insight into possible rearrangement mechanisms. *Haemophilia.* 2023; 29(4):1121–1134. Epub 20230516. <https://doi.org/10.1111/hae.14799> PMID: 37192522.
 38. Shimojima Yamamoto K, Tamura T, Okamoto N, Nishi E, Noguchi A, Takahashi I, et al. Identification of small-sized intrachromosomal segments at the ends of INV-DUP-DEL patterns. *J Hum Genet.* 2023. Epub 20230710. <https://doi.org/10.1038/s10038-023-01181-x> PMID: 37423943.

39. Miller AW, Befort C, Kerr EO, Dunham MJ. Design and use of multiplexed chemostat arrays. *J Vis Exp*. 2013;(72):e50262. <https://doi.org/10.3791/50262> PMID: 23462663; PubMed Central PMCID: PMC3610398.
40. Liao X, Zhu W, Zhou J, Li H, Xu X, Zhang B, et al. Repetitive DNA sequence detection and its role in the human genome. *Commun Biol*. 2023; 6(1):954. Epub 20230919. <https://doi.org/10.1038/s42003-023-05322-y> PMID: 37726397; PubMed Central PMCID: PMC10509279.
41. Devriendt K, Matthijs G, Holvoet M, Schoenmakers E, Fryns JP. Triplication of distal chromosome 10q. *J Med Genet*. 1999; 36(3):242–245. Epub 1999/04/16. PMID: 10204854; PubMed Central PMCID: PMC1734335.
42. Mercer CL, Browne CE, Barber JC, Maloney VK, Huang S, Thomas NS, et al. A complex medical phenotype in a patient with triplication of 2q12.3 to 2q13 characterized with oligonucleotide array CGH. *Cytogenet Genome Res*. 2009; 124(2):179–186. Epub 2009/05/08. 000207526 [pii] <https://doi.org/10.1159/000207526> PMID: 19420931.
43. Ungaro P, Christian SL, Fantes JA, Mutirangura A, Black S, Reynolds J, et al. Molecular characterisation of four cases of intrachromosomal triplication of chromosome 15q11-q14. *J Med Genet*. 2001; 38(1):26–34. Epub 2001/01/03. <https://doi.org/10.1136/jmg.38.1.26> PMID: 11134237; PubMed Central PMCID: PMC1734721.
44. Kato T, Inagaki H, Miyai S, Suzuki F, Naru Y, Shinkai Y, et al. The involvement of U-type dicentric chromosomes in the formation of terminal deletions with or without adjacent inverted duplications. *Hum Genet*. 2020; 139(11):1417–1427. Epub 20200602. <https://doi.org/10.1007/s000439-020-02186-8> PMID: 32488466.
45. Yatsenko SA, Hixson P, Roney EK, Scott DA, Schaaf CP, Ng YT, et al. Human subtelomeric copy number gains suggest a DNA replication mechanism for formation: beyond breakage-fusion-bridge for telomere stabilization. *Hum Genet*. 2012; 131(12):1895–1910. Epub 20120814. <https://doi.org/10.1007/s00439-012-1216-9> PMID: 22890305; PubMed Central PMCID: PMC3493700.
46. Slack A, Thornton PC, Magner DB, Rosenberg SM, Hastings PJ. On the mechanism of gene amplification induced under stress in *Escherichia coli*. *PLoS Genet*. 2006; 2(4):e48. Epub 20060407. <https://doi.org/10.1371/journal.pgen.0020048> PMID: 16604155; PubMed Central PMCID: PMC1428787.
47. Tsaponina O, Haber JE. Frequent Interchromosomal Template Switches during Gene Conversion in *S. cerevisiae*. *Mol Cell*. 2014; 55(4):615–625. Epub 20140724. <https://doi.org/10.1016/j.molcel.2014.06.025> PMID: 25066232; PubMed Central PMCID: PMC4150392.
48. Barsky RF. Noam Chomsky: a life of dissent. Repr. ed. Cambridge, Mass: MIT Press; 1998.
49. Zhang F, Carvalho CM, Lupski JR. Complex human chromosomal and genomic rearrangements. *Trends Genet*. 2009; 25(7):298–307. Epub 2009/06/30. <https://doi.org/10.1016/j.tig.2009.05.005> [pii] PMID: 19560228.
50. Shimojima K, Mano T, Kashiwagi M, Tanabe T, Sugawara M, Okamoto N, et al. Pelizaeus-Merzbacher disease caused by a duplication-inverted triplication-duplication in chromosomal segments including the PLP1 region. *Eur J Med Genet*. 2012; 55(6–7):400–403. Epub 20120321. <https://doi.org/10.1016/j.ejmg.2012.02.013> PMID: 22490426.
51. Abdala BB, Goncalves AP, Dos Santos JM, Boy R, de Carvalho CMB, Grochowski CM, et al. Molecular and clinical insights into complex genomic rearrangements related to MECP2 duplication syndrome. *Eur J Med Genet*. 2021; 64(12):104367. Epub 20211019. <https://doi.org/10.1016/j.ejmg.2021.104367> PMID: 34678473.
52. Kohmoto T, Okamoto N, Naruto T, Murata C, Ouchi Y, Fujita N, et al. A case with concurrent duplication, triplication, and uniparental isodisomy at 1q42.12-qter supporting microhomology-mediated break-induced replication model for replicative rearrangements. *Mol Cytogenet*. 2017; 10:15. Epub 20170428. <https://doi.org/10.1186/s13039-017-0316-6> PMID: 28465723; PubMed Central PMCID: PMC5410019.
53. Boerkoel PK, Dixon K, Fitzsimons C, Shen Y, Huynh S, Schlade-Bartusiak K, et al. Long-read genome sequencing resolves a complex 13q structural variant associated with syndromic anophthalmia. *Am J Med Genet A*. 2022; 188(5):1589–1594. Epub 20220205. <https://doi.org/10.1002/ajmg.a.62676> PMID: 35122461.
54. Xiao B, Ye X, Wang L, Fan Y, Gu X, Ji X, et al. Whole Genome Low-Coverage Sequencing Concurrently Detecting Copy Number Variations and Their Underlying Complex Chromosomal Rearrangements by Systematic Breakpoint Mapping in Intellectual Deficiency/Developmental Delay Patients. *Front Genet*. 2020; 11:616. Epub 20200706. <https://doi.org/10.3389/fgene.2020.00616> PMID: 32733533; PubMed Central PMCID: PMC7357533.