



**HAL**  
open science

# Homologous Recombination and the Formation of Complex Genomic Rearrangements

Aurèle Piazza, Wolf-Dietrich Heyer

► **To cite this version:**

Aurèle Piazza, Wolf-Dietrich Heyer. Homologous Recombination and the Formation of Complex Genomic Rearrangements. Trends in Cell Biology, 2019, 29 (2), pp.135-149. 10.1016/j.tcb.2018.10.006 . hal-02999807

**HAL Id: hal-02999807**

**<https://hal.science/hal-02999807>**

Submitted on 11 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1

2

# 3 **Homologous recombination and the formation of complex genomic** 4 **rearrangements**

5

6

7 Aurèle Piazza<sup>1,3</sup> and Wolf-Dietrich Heyer<sup>1,2</sup>

8

9 <sup>1</sup>Department of Microbiology and Molecular Genetics, <sup>2</sup>Department of Molecular and Cellular  
10 Biology, One Shields Avenue, University of California, Davis, CA 95616, USA, <sup>3</sup> Spatial  
11 Regulation of Genomes, Department of Genomes and Genetics, CNRS UMR3525, Institut  
12 Pasteur, 28 Rue du Docteur Roux, 75015 Paris, France

13 \*Corresponding author. E-mail: wdheyer@ucdavis.edu Phone: +1 (530) 752-3001

14

15 **Keywords:** chromothripsis, copy number variation, multi-invasion, non-allelic homologous  
16 recombination, structure-selective endonuclease, structural variant

17

18 Abbreviations: BIR, break-induced replication; CNV, copy number variant; dsDNA: double-  
19 stranded DNA; DSB, double-strand break; EJ, endjoining; hDNA, heteroduplex DNA; HR,  
20 homologous recombination; MI, multi-invasion; MIR, MI-induced rearrangement; MMEJ,  
21 microhomology-mediated end-joining; NE, nuclear envelope; NHEJ: non-homologous end-  
22 joining; ssDNA, single-strand DNA; SSE, structure-selective endonuclease; SV, structural  
23 variant.

1 **Abstract**

2 The maintenance of genome integrity involves multiple independent DNA damage avoidance  
3 and repair mechanisms. Yet, the origin and pathways of the focal chromosomal reshuffling  
4 phenomena collectively referred to as chromothripsis remain mechanistically obscure. Here ,we  
5 discuss the role, mechanisms, and regulation of HR in the formation of simple and complex  
6 chromosomal rearrangements. We emphasize features of the recently characterized Multi-  
7 invasions Induced Rearrangement (MIR) pathway, which uniquely amplifies the initial DNA  
8 damage. HR intermediates and cellular contexts at risk for genomic stability are discussed along  
9 with the emerging roles of various classes of nucleases in the formation of genome  
10 rearrangements. Long-read sequencing and improved mapping of repeats should enable better  
11 appreciation of the significance of recombination in generating genomic rearrangements.

12

## 1 **Complex chromosomal rearrangements and chromothripsis**

2 Chromosomal rearrangements encompass any structural variation (SV) of the genome,  
3 regardless of its association with copy number variation (CNV). The advent of high throughput  
4 sequencing technologies revealed massive and complex clustered structural variations [1],  
5 which have been proposed to constitute a novel genomic instability phenomenon found in  
6 cancer genomes, congenital diseases, as well as in asymptomatic individuals [1-4]. Although  
7 the terminology varies with the precise nature of the alterations and possibly its etiology, this  
8 phenomenon is widely referred to as *chromothripsis*, and we will use this umbrella term here  
9 **(Box: Mutational Phenomena)**. Formal criteria based on the CNV pattern and the physically  
10 confined nature of the rearrangement junctions have been proposed to define chromothripsis  
11 ([1, 5] but see also [6] for a different perspective). The key underlying feature of chromothripsis  
12 is the abrupt acquisition of the associated rearrangements. The number of junctions in localized  
13 rearrangements suggests a continuum of complexity, with chromothripsis being potentially an  
14 extreme expression of mechanism(s) also responsible for simpler SVs [7].

## 15 **Pathways for complex chromosomal rearrangements as studied in *S. cerevisiae***

16 A number of experimental systems have been developed in tractable model organisms to  
17 decipher the origin of both simple and complex SVs as well as the pathways promoting and  
18 preventing their occurrence in various sequence contexts. Here, we will focus on work  
19 performed in *Saccharomyces cerevisiae*, in which the conserved double-strand break (DSB)  
20 repair mechanisms and the consequences of their defect have been best understood thanks to  
21 the exquisite genetic and molecular tools in that organism [8].

22 Most strikingly, and despite their relatively low abundance in the *S. cerevisiae* genome,  
23 repeated DNA elements were the predominant mediators of SVs [9-12]. Beyond obvious  
24 pathological consequences, these repeat-mediated CNVs also contributed to rapid adaptation  
25 upon artificial gene dosage imbalances [13, 14] or nutrient-limiting conditions [15, 16].

26 Genetic screens revealed the complex networks of proteins involved in genome maintenance in  
27 yeast with implications for human cancer [17]. Kolodner and co-workers identified an  
28 astounding 182 genes (3%) playing a primary role and 438 genes (7%) playing a supporting  
29 role in suppressing genomic instability in the absence of exogenous DNA damage [17]. These  
30 genes participate in two broad functions whose simultaneous inactivation synergizes to  
31 destabilize the genome: 1) the prevention or removal of structural DNA damage or aberrant  
32 structures, and 2) the promotion of accurate repair of the damage, with certain functions

1 intersecting both categories (*e.g.* mismatch repair). Structural DNA lesions at risk for genomic  
2 stability are varied in nature and origin. In unchallenged cells they are believed to mainly  
3 originate from replication errors in the form of a persistent ssDNA gaps or broken replication  
4 fork, *i.e.* a single-ended **DSB** (see **Glossary**) [8, 18, 19]. These two types of lesion are  
5 recombinogenic substrates. For simplicity we will focus here on DSBs, which have been best  
6 studied in their double-ended form upon site-specific induction [20].

## 7 **Homologous recombination pathways and their associated risks to genomic stability**

8 The DSB repair strategies can be broadly separated based on their homology requirements and  
9 sub-categorized based on their genetic dependencies, as (i) a homology-independent end-  
10 joining (NHEJ) mechanism, (ii) single strand annealing mechanisms relying either on micro-  
11 homology (Alt-EJ/MMEJ) or extensive homology (SSA), and (iii) homology-dependent DNA  
12 strand invasion mechanisms collectively referred to as HR (**Fig. 1**). While SSA is a homology-  
13 dependent process, it does not involve DNA strand invasion and may be responsible for  
14 homology-directed repair independent of Rad51 [21]. We briefly review here the risks to  
15 genomic stability inherent to the HR pathway, and refer the reader interested in the role of EJ  
16 and annealing mechanisms to other recent reviews [22-24].

17 HR templates DSB repair by locating and copying an extensive identical (homologous) or near-  
18 identical (homeologous) sequence present in intact duplex DNA (**Fig. 1**). Hence, HR uniquely  
19 entails homology search and DNA strand invasion of the broken molecule to identify and invade  
20 a homologous template. The DNA strand invasion reaction results in a **D-loop** (**Fig. 1**)  
21 containing heteroduplex DNA (hDNA). These reactions are catalyzed by a helical filament of  
22 Rad51 (RecA in bacteria) and associated proteins assembled on the resected ssDNA flanking  
23 the DSB [25]. Upon pairing of the 3' extremity of the broken molecule, DNA synthesis restores  
24 the sequence information disrupted by the DSB. The HR sub-pathways branch based on the  
25 differential processing of this extended D-loop intermediate (**Fig. 1**). Two features of this  
26 pathway have important consequences for genomic stability: the unstable nature of the DNA  
27 synthesis occurring in the context of the D-loop and the potential of various types of DNA joint  
28 molecules generated throughout HR to be aberrantly processed by structure-selective  
29 endonuclease (SSE) (see below). Additionally, HR-mediated rearrangements can lead to  
30 dicentric chromosome formation and subsequent chronic instability by **Breakage-Fusion-**  
31 **Bridge** cycles [26, 27].

1 Disruption of the extended D-loop funnels the pathway towards Synthesis-Dependent Strand  
2 Annealing (SDSA), which leads to a **non-crossover** outcome with minimal associated gene  
3 **conversion**. As such, it is the most conservative sub-pathway of HR, even when using a donor  
4 at an ectopic locus.

5 Alternatively, the displaced strand in the extended D-loop can anneal to the second end of the  
6 DSB leading to the formation of a double Holliday Junction (dHJ), as part of the pathway  
7 historically coined Double-Strand Break Repair (DSBR) [28]. This step may depend on more  
8 extensive DNA synthesis than SDSA [29]. This covalently linked intermediate, detected  
9 physically both in somatic and meiotic cells [30, 31], can either be topologically dissolved [32]  
10 or resolved endonucleolytically [33]. While dissolution always yields a non-crossover outcome  
11 with minimal gene conversion, resolution can lead to a **crossover** outcome. If occurring at an  
12 allelic locus, crossovers will cause a loss of heterozygosity at the next cell division in half of  
13 the cases. If occurring between non-allelic loci, a crossover will lead to a reciprocal  
14 translocation (**Fig. 1**).

15 Alternatively to this second end annealing scenario, a defect in extended D-loop disruption or  
16 the absence of a second end to anneal to (*e.g.* in the case of a single-ended DSB) leads to a long-  
17 range displacement DNA synthesis mechanism known as Break-Induced Replication (BIR)  
18 until stabilization of the broken chromosome by capture of a telomeric sequence or the merging  
19 with a convergent replication fork [34, 35]. BIR is a physiological pathway of broken fork  
20 recovery also prone to generate various types of rearrangements (reviewed in ref. [36]).  
21 Replication forks can break through direct cleavage by structure-selective endonucleases, such  
22 as MUS81-EME1, or by replicating through an existing single-stranded nick in the template  
23 [19]. Indeed, the massive gene conversion resulting from its conservative nature [37, 38] will  
24 lead either to loss of heterozygosity upon repair templated by the homologous chromosome or  
25 an unbalanced SV if initiated at an ectopic site. Ectopic BIR is also associated with copy number  
26 (CN) gain of the template molecule. The overall mutagenic potential of BIR depends on the  
27 extent of synthesis it can achieve, which can be several hundreds of kilobases [37]. Finally, the  
28 long-lived single-stranded DNA associated with BIR [38, 39] is at risk for increased  
29 mutagenesis and *kataegis* [40].

30 A consequence of the unstable displacement DNA synthesis during HR is the occasional  
31 occurrence of template switches at regions of extensive homo- and homeologies over rounds of  
32 extended D-loop dissociation and re-invasion, both as part of BIR and SDSA [41-45]. Template  
33 switches can also occur at nearby micro-homologies when D-loop extension during BIR is

1 impaired in a *PIF1* mutant [46]. These switches thus provide opportunities for multiple CN  
2 gain-associated SVs from a single DSB repair event. This propensity to switch decreases at a  
3 distance from the invasion point, as the migrating synthesis bubble is converted by **structure-**  
4 **selective endonucleases (SSEs)** to a stable replication fork [35, 44, 47, 48]. Hence, BIR-  
5 mediated SVs are predicted to be found clustered at one end of the gained region, where the  
6 initial invasion and synthesis have occurred. Moreover, several HR intermediates are substrates  
7 for SSEs, and a combination of strand cleavage and ligation can convert a D-loop to a **half-**  
8 **crossover (Figs. 1, 2A)** [44, 49-51]. It is predicted to transfer a co-oriented 3'-protruding single-  
9 ended DSB on the donor, thus maintaining the initial amount and orientation of DSB extremities  
10 (**Fig. 2A**). Such DSB transfer on the donor have been physically detected in a related  
11 mechanism (MIR, see below) and shown to be dependent on both Rad51 (*i.e.* DNA strand  
12 invasion) and SSEs [52].

### 13 **Multi-invasions-Induced Rearrangement (MIR): an HR-based mechanism for the** 14 **formation of complex rearrangements**

15 Biochemical work revealed that a single presynaptic filament can pair and invade (*i.e.* form  
16 hDNA with) multiple dsDNA donors at once, resulting in a multi-invasion intermediate (MI;  
17 **Fig. 2B**) [52, 53]. MI joint molecules are readily formed by long presynaptic filaments made  
18 with bacterial RecA, yeast Rad51-Rad54, and human RAD51-RAD54, and are stimulated by  
19 increasing homology length ([52, 53] and Wright & Heyer, unpublished). These results suggest  
20 that MIs are byproducts of basic activities of the presynaptic filament: the inter-segmental  
21 homology sampling process [54] and hDNA formation.

22 It is noteworthy that terminal homologies are not required for MI formation [52, 53]. It suggests  
23 that ssDNA regions, such as replication-associated gaps, distant from or devoid of a freely  
24 rotating end could form MI joint molecules, despite the topological constraints for strand  
25 intertwining [55]. This possibility greatly expands the pathological and physiological contexts  
26 conducive to MI beyond the DSB-induced and -proximal situation studied so far [52] (see below  
27 and **Fig. 3**).

28 MI joint molecules form in *S. cerevisiae* cells and induce the formation of chromosomal  
29 rearrangements (**Fig. 2C**) [52]. Specifically, the endonucleolytic processing of MI by SSEs  
30 leads to a translocation of the donors that inserts the intervening sequence of the invading  
31 molecule between the two invaded regions, in a mechanism referred to as MI-Induced  
32 Rearrangement (MIR; Fig. 2B) [52]. In addition to the translocation, MIR generates additional

1 single-ended DSBs on each donor (**Fig. 2C**; mechanisms detailed in ref. [56]). Moreover, the  
2 initiating DSB is not repaired during the MIR process (**Fig. 2C**). Hence, while D-loop cleavage  
3 leading to half-crossover keeps the total amount of DSB ends constant, MIR generates two  
4 additional single-ended DSBs (compare **Figs. 2A** and **2C**). In both cases, the propensity to  
5 generate additional and sometimes complex rearrangements upon attempted repair of the  
6 remaining DSBs (initial and newly formed) depends on their respective sequences [52, 56].  
7 Hence, MIR is a HR pathway that uniquely amplifies the initial damage, and is thus at risk of  
8 runaway cascades of rearrangements.

### 9 **Reversibility of HR intermediates by HR regulators guards against MIR and repeat-** 10 **mediated genomic instability**

11 HR accuracy relies in part on kinetic proof-reading, which is enforced by the reversal of several  
12 non-covalent intermediates of the pathway as a safeguard against HR-mediated rearrangements  
13 [57]. Distinct but possibly overlapping reversal activities in yeast are supported by a diverse set  
14 of HR regulators including the Srs2 helicase (putative functional human homologs are FBH1,  
15 RTEL, PARI and RECQ5), the Mph1 helicase (human FANCM), the Sgs1-Top3-Rmi1  
16 helicase-topoisomerase complex (STR; human BLM-TOPO3 $\alpha$ -RMI1/2), and mismatch repair  
17 (MMR) factors (reviewed in [57]). These regulators operate at multiple possible steps during  
18 the HR pathway [57]. Srs2 dissociates Rad51 filaments [58, 59] likely formed on ssDNA gaps  
19 generated during replication that are otherwise toxic [60]. Nascent and extended D-loops are  
20 nodes for two radically different decisions: anti-recombination and anti-crossover, respectively  
21 (**Fig. 1**). Srs2, Mph1, and STR disrupt D-loops in reconstituted biochemical reactions with  
22 Rad51, Rad54 and RPA [61-63]. In cells, they inhibit homeologous recombination in  
23 coordination with MMR [9], promote and/or bias the HR repair outcome towards non-crossover  
24 [61, 64-66], and in the case of Mph1 and Srs2 promote template-switch during BIR [43].  
25 Additionally, the STR complex uniquely catalyzes dHJ dissolution [32, 67]. Hence, Srs2, Mph1  
26 and STR are believed to operate throughout the HR pathway at the presynaptic (resection and  
27 Rad51 filament), synaptic (nascent D-loop formation and stability) and post-synaptic (extended  
28 D-loop stability and dHJ processing) steps, which results in conservative repair products and  
29 mitigates the risk of HR-induced genomic instability.

30 Consistent with their proposed roles in D-loop disruption, Srs2, Mph1 and STR also suppress  
31 MIR [52]. Additionally, the Rad1-Rad10 endonuclease (human XPF-ERCC1) suppresses MIR  
32 by clipping the 3'-flap of the internal invasion, thus preempting formation of the MI



1 intermediate (**Fig. 2D**) [52]. Analysis of combinations between these MIR suppressive  
2 pathways suggested that these inhibitory activities are exerted at different steps or substrates  
3 [52]. Conversely, physical analysis of BIR intermediates suggested that Srs2 inhibits MI  
4 formation presumably by preventing Rad51 filament assembly on and/or by disrupting internal  
5 D-loops formed by the trailing ssDNA upstream of the extending D-loop [39]. Hence, multiple  
6 independent activities prevent the formation and/or accumulation of HR byproducts at risk for  
7 genomic instability. Their precise substrate(s), mechanisms and interactions have not yet been  
8 deciphered.

### 9 **Long single-stranded DNAs and nucleases are instrumental in MIR**

10 Various endo- and exonucleases are instrumental for both the early and late steps of MIR, by  
11 generating the recombinogenic substrates and by processing the subsequent MI joint molecules,  
12 respectively (**Fig. 3**). Long ssDNAs are key precursors of MIR: the amount of MI species in  
13 reconstituted D-loop reaction *in vitro* and the frequency of MIR *in vivo* are greatly stimulated  
14 by increasing ssDNA and homology length, respectively [52, 53]. By exposing long ssDNAs,  
15 extensive resection presumably enable the concomitant encounter of independent donors by the  
16 Rad51-ssDNA homology search engine [52-54] and exposes repeated sequences located at a  
17 distance from the DSB site, thus enabling SSA [12, 68]. Consequently, various cellular  
18 processes that generate long ssDNA are potential instigators of repeat-mediated genomic  
19 instability by MIR (**Fig. 3**). DSB resection is a regulated process that involves several endo-  
20 and/or exonucleases [69], and which has recently been shown to be reversible [70]. Briefly, the  
21 Sae2-MRX (human CtIP-MRN) endo- and exonuclease complex initiates DSB resection at  
22 short range (~1 kb), while the Exo1 exonuclease (human EXO1) and Sgs1/Dna2 (human  
23 BLM/DNA2) redundantly generate ssDNA of up to several tens of kilobases [69]. Single-strand  
24 breaks as well as gaps on the lagging strand of replication forks can also be exonucleolytically  
25 processed in a similar fashion [71, 72] (**Fig. 3**). Long ssDNAs can also be generated in a  
26 nuclease-independent fashion at the replication fork, such as upon uncoupling between the  
27 leading strand DNA polymerase and the replicative helicase [73-75] or when DNA primase  
28 becomes limiting [74, 76]. The endonucleolytic processing of replication forks by SSEs is also  
29 expected to yield single-ended DSBs exhibiting varying degrees of ssDNA, especially upon  
30 uncoupling (**Fig. 3**). Broken forks can be rescued by BIR, which generates persistent, kilobases-  
31 long newly synthesized ssDNA trailing behind the extending D-loop [38], which was physically  
32 shown to invade *in cis* upstream of the extending D-loop [39]. Previously observed BIR-

1 induced recombination events between donor chromosomes may result from the  
2 endonucleolytic processing of such MIs [43].

3 As discussed earlier, DNA strand invasion intermediates generated during HR exhibit a variety  
4 of branch points that can be recognized and cleaved by SSEs with varying efficiency [33].  
5 Consistently, three SSEs in *S. cerevisiae* (Mus81-Mms4, Slx1-Slx4, and Yen1) are required in  
6 a redundant fashion for MIR [52]. Functional redundancy between these SSEs in cleaving  
7 other DNA structures had been previously reported [44]. Importantly, while moving the  
8 pathway towards the completion of the donor translocation, MI cleavage by SSEs propagates  
9 additional single-ended DSBs onto the donors (**Fig. 2C**).

10 In conclusion, nucleases contribute to HR-mediated genomic stability and specifically play a  
11 two-fold role in MIR by: 1) generating the initiating recombinogenic substrate and 2)  
12 processing DNA joint molecules into both genomic rearrangements and additional  
13 recombinogenic damage [44, 52].

#### 14 **Chromothripsis as a consequence of a spatial and temporal failure to separate nucleases** 15 **from their potential substrate?**

16 Recent inducible experimental systems of chromothripsis in mammalian cells started to shed  
17 light on cellular contexts prone to generate massive and localized structural DNA damage and  
18 their commonalities, and the repair mechanisms involved in the formation of the rearrangement.

#### 19 Nuclear compartmentalization defects are associated with DNA damage formation and 20 chromothripsis

21 Cellular compartmentalization and cell cycle regulation ensure that mutually exclusive  
22 metabolic activities are separated spatially and/or temporally [77]. For instance, failure to  
23 temporally separate DNA replication and mitotic entry results in mild chromosomal fragility  
24 up to full chromosome shattering [78, 79]. The breakage results from replication fork collapse  
25 triggered by the activation of SSEs [80-82], which normally act at mitosis as a last resort  
26 mechanism for disentangling various types of DNA joint molecules that dissolution and  
27 decatenation have failed to process [83, 84]. From a spatial point of view, one role of the nuclear  
28 envelope (NE) is to isolate the genomic DNA from the nucleases that patrol the cytosol as part  
29 of the cellular immune system (**Fig. 4**) [85]. In fact, NE disruption and/or programmed entry of  
30 endo- and exonucleases (endoG, TREX1 and NM23-H1) into the nucleus are integral to the  
31 genomic DNA elimination program of apoptosis induced during viral infection or in tumor cells

1 [86-88]. The cytosol also contains Yen1/GEN1, an SSE involved in the elimination of various  
2 persistent replication and HR intermediates upon NE breakdown at mitosis [89]. Its activity is  
3 subjected to a dual inhibitory control in human: a spatial exclusion from its potential substrates  
4 and a cell-cycle regulated inactivation of the enzyme by phosphorylation [90, 91].  
5 Consequently, NE rupture will accidentally expose genomic DNA to various nucleases, thus  
6 fulfilling one of the prerequisite for chromothripsis: the formation of massive DNA damage  
7 (**Fig. 4**).

8 Two independent chromothripsis-inducing contexts that both entail nuclear  
9 compartmentalization defects have recently been investigated (**Fig. 5**) [92-94]. In two studies,  
10 chromothripsis arose from the isolation of whole or pieces of chromosome within **micronuclei**  
11 consecutive to their mis-segregation at the previous mitosis (**Fig. 5A**) [92, 93]. A consequence  
12 of this disrupted nuclear homeostasis in micronuclei is the massive formation of DNA damage  
13 containing, or coinciding with, long ssDNA (as revealed by colocalized phosphorylated RPA  
14 and  $\gamma$ H2AX staining) specifically during or following S-phase, but not in G1 (**Fig. 5A**) [95].  
15 Their S-phase-dependent nature suggested that defective replication itself or the exposure of  
16 DNA structures generated during replication to cytoplasmic components such as SSEs leads to  
17 DNA damage (**Fig. 4**).

18 In another study, chromothripsis was initiated by the attempted segregation of a dicentric  
19 chromosome (**Fig. 5B**) [94]. The force exerted by the spindle at metaphase ( $\approx 1$  nN) [96] is  
20 expected to cause structural transitions and partial melting of protein-free dsDNA [97, 98], but  
21 remains insufficient to break the molecules, as rupture of covalent bonds occurs only above 2  
22 nN [99]. Instead, resolution of various types of DNA bridges involved the action of nucleases,  
23 either at the mid-body during cytokinesis or upon breakdown or rupture of the NE [83, 84, 94].  
24 In the case of chromatin bridges, the NE reformed around the segregated chromosomes as well  
25 as the bridge that persisted between the two daughter cells, which remains subjected to robust  
26 antagonistic forces (**Fig. 5B**) [94]. These forces ultimately lead to localized NE rupture at the  
27 base of the bridge, exposing the intervening chromatin to cytoplasmic components. The TREX1  
28 exonuclease, either by exploiting pre-existing nicks or aided by an uncharacterized  
29 endonuclease, resects the stretched bridge until its resolution, causing the snapback of massive  
30 amounts of RPA-coated ssDNA into the daughter cells (**Fig. 5B**). This damaged DNA prompted  
31 the formation of 53BP1,  $\gamma$ H2AX and Mre11 foci in the daughter cells, where it induced  
32 chromothriptic rearrangements [94]. Notably, these micronuclei-independent chromothripsis

1 events were uniquely found associated with kataegis, consistent with long-lived ssDNA as an  
2 initiating substrate for chromothripsis (see below).

3 Hence, the commonality of these two experimentally distinct chromothripsis-inducing contexts  
4 is a defect in nuclear compartmentalization, and in one case the demonstrated involvement of a  
5 cytoplasmic nuclease and long ssDNAs.

### 6 Repair mechanisms generating chromothriptic rearrangements

7 Initial examination of the genomic characteristics of chromothriptic rearrangements, such as  
8 the SV junction sequences and the limited and oscillating copy number variation, suggested  
9 that chromothripsis resulted either from the EJ-mediated stitching of numerous DSBs and/or  
10 of microhomology-mediated template-switching during replicative or repair processes [1, 4].  
11 Experimental evidence indicated that the DNA damage generated in both micronuclei during  
12 S-phase and upon bridge resolution in G1 or S phase stains with RPA antibodies, and certain  
13 chromothripsis events were found associated with kataegis over kilobases-long regions [94,  
14 100, 101]. These observations suggested the involvement of long ssDNAs in chromothripsis, a  
15 recombinogenic substrate that is not readily processed by EJ mechanisms. Furthermore, HR-  
16 proficient breast cancer cells exhibit clustered rearrangements of various types while HR-  
17 deficient cells do not [102]. Moreover, chromothripsis was found mutually exclusive with  
18 biallelic BRCA2 mutations in metastatic prostate cancers [103]. Finally, p53-deficient mice  
19 inactivated for EJ (XRCC4 or LIG4) or HR factors (BRCA2) still exhibited complex  
20 rearrangements consistent with chromothripsis or *chromoanaysynthesis*, suggesting the  
21 involvement of both pathways in chromothripsis etiology [104].

22 Ly and colleagues experimentally addressed whether DNA damage, in addition to be generated  
23 in micronuclei, could also be repaired there [92]. Depletion of canonical NHEJ factors (LIG4  
24 and DNA-PKcs), but not HR (BRCA2 and RAD51) or MMEJ factors (LIG3 and PARP1), led  
25 to an increase of the fragmentation of the micronuclei-contained Y chromosome observed at  
26 the subsequent mitosis. It suggests that a significant fraction of the micronuclei-induced  
27 damages are substrate for the NHEJ machinery, and that this type of repair is active in the  
28 context of the micronuclei while HR repair is not. The remaining damages are subsequently  
29 repaired upon reincorporation of the micronuclear DNA into the main nucleus at mitosis (**Fig.**  
30 **5A**) [93], the nature of which remains unknown. Although not yet demonstrated by sequencing,  
31 this NHEJ-dependent minimization of micronuclear DNA fragmentation suggests that the  
32 mechanisms generating chromothriptic rearrangements could occur in the micronuclei in

1 addition to the main nucleus. The involvement of NHEJ is consistent with the analysis of  
2 junction sequences in chromothriptic genomes [1].

3 In contrast to the Y-chromosome system, bridge-induced chromothripsis originated from  
4 extensive ssDNA and did not undergo partial repair in micronuclei [94]. How could a single  
5 long ssDNA or multiple ssDNA gaps that are not readily substrate for EJ repair trigger  
6 chromothripsis in the daughter cell? Since long ssDNAs are suitable for Rad51 filament  
7 assembly, bridge-induced chromothripsis potentially resulted from attempted HR repair. This  
8 repair is expected to be independent of factors required for resection initiation and extension,  
9 thus bypassing an important (although not unique [105]) G1-specific block to HR, at the  
10 resection initiation level (reviewed in ref. [69]). Defects in p53 in the strain used may also have  
11 potentiated HR [106] by failing to repress *RAD51* expression [107]. This situation is  
12 nonetheless relevant as p53 is inactivated in the majority of human cancers [108]. Furthermore,  
13 the prevalence of chromothripsis is higher in p53-deficient tumors [109]. Given the concomitant  
14 NE integrity defect, we suspect that HR intermediates generated along this ssDNA may lead to  
15 damage transfer and amplification similar to what is observed during half-crossover and MIR,  
16 respectively (**Fig. 2**).

17 Hence, the etiology of chromothripsis is complex and likely influenced by the damage-inducing  
18 context, with different DSB and/or ssDNA gap repair pathways involved in the formation of  
19 the rearrangements.

20 **Why is there no extensive homology at the chromothriptic SV junctions? The need for**  
21 **long-read genome assemblies.**

22 It is striking that prior to the advent of high-throughput DNA sequencing the finely-mapped SV  
23 junctions observed in human genomes involved repeated elements [110, 111], in agreement  
24 with yeast studies [8]. Then how is it that the vast majority of the SV junctions documented  
25 since then exhibit significant micro-homology (1-10 nt) but rarely longer homologies [112],  
26 despite the massive repeat content of the human genome? Instead of a biological reality, we  
27 suspect that this bias originates from technical limitations of the dominant paired-end  
28 sequencing technology and analysis pipelines. First, reads ambiguously mapped (*i.e.* in repeats)  
29 are usually discarded. Second, paired-end sequencing of <400 bp fragments used to detect SV  
30 poses an absolute upper threshold to the repeat size that can be detected at a SV junction.  
31 Consistently, germline and somatic cancer rearrangements in humans identified by such  
32 methods are found enriched near or at short repeats (SINEs) but strongly depleted at long

1 repeats (LINEs) [113]. Accordingly, the exhaustive establishment of mutational profiles  
2 (particularly SVs) in *S. cerevisiae* required the use of additional molecular techniques in  
3 addition to paired-end sequencing [114]. We anticipate that the improved mappability of long  
4 repeated regions thanks to long-read technologies, optical mapping or Hi-C approaches will  
5 reveal more SV formed at repeats [115-118]. The recent application of long-read sequencing  
6 and across-platform comparisons reported a dramatic increase in the identification of SVs [115,  
7 119-121], and improved SV calling algorithms could detect significant enrichment of junctions  
8 within LINEs [112]. These technical advances will produce a more comprehensive picture of  
9 the patterns of SVs in genomic rearrangements and will enable to test the involvement of HR  
10 and especially MIR in chromothripsis.

### 11 **Concluding remarks**

12 Far from an error-free pathway, HR can generate a variety of rearrangements due to the  
13 repetitive nature of the genome. The HR pathway entails the formation of DNA joint molecules  
14 at risk to induce rearrangements when cleaved inappropriately. Consequently, dysregulation of  
15 SSEs and other nucleases puts the cell at risk of a HR-mediated runaway cascade of damage  
16 and rearrangement, as demonstrated with the MIR pathway. Given that the integrity of the NE  
17 is compromised in experimental models for chromothripsis, we suspect that a key aspect of  
18 both SV and damage formation of chromothripsis reside in the accidental exposure of  
19 replication and HR intermediates to various nucleases. Finally, we highlight the need for  
20 genomes assembled from long-read DNA sequencing techniques to allow evaluating the  
21 contributions of HR between repeated DNA in the formation of genomic rearrangements and  
22 chromothripsis.

### 23 **Acknowledgements**

24 We are grateful to present and former members of the Heyer, Hunter and Kowalczykowski  
25 laboratories for stimulating discussions, Steven Gore, Thomas Guérin and Stéphane Marcand  
26 for helpful insights and comments, and Jim Haber, John McPherson, Richard Kolodner and  
27 Markus Löbrich for their critical reading of the manuscript. We apologize to all authors whose  
28 primary work could not be cited due to tight constraints in the allowed number of references.  
29 AP was supported by fellowships from the Fondation ARC pour la Recherche sur le Cancer  
30 (PDF20171206726), the EMBO (ALTF-238-2013), the Framework Project 7 of the European  
31 Union (Marie Curie International Outgoing Fellowship 628355), and received financial support

1 from the Philippe Foundation. Research in the WDH laboratory is supported by NIH grants  
2 GM58015 and CA92276 and the France-Berkeley Fund.

3

#### 4 **References**

- 5 1. Stephens, P.J. et al. (2011) Massive genomic rearrangement acquired in a single catastrophic event  
6 during cancer development. *Cell* 144 (1), 27-40.
- 7 2. Cortes-Ciriano, I. et al. (2018) Comprehensive analysis of chromothripsis in 2,658 human cancers  
8 using whole-genome sequencing. *bioRxiv*.
- 9 3. Kloosterman, W.P. and Cuppen, E. (2013) Chromothripsis in congenital disorders and cancer:  
10 similarities and differences. *Curr Opin Cell Biol* 25 (3), 341-8.
- 11 4. Liu, P. et al. (2011) Chromosome catastrophes involve replication mechanisms generating complex  
12 genomic rearrangements. *Cell* 146 (6), 889-903.
- 13 5. Korbel, J.O. and Campbell, P.J. (2013) Criteria for inference of chromothripsis in cancer genomes.  
14 *Cell* 152 (6), 1226-36.
- 15 6. Kinsella, M. et al. (2014) The elusive evidence for chromothripsis. *Nucleic Acids Res* 42 (13),  
16 8231-42.
- 17 7. Storchova, Z. and Kloosterman, W.P. (2016) The genomic characteristics and cellular origin of  
18 chromothripsis. *Curr Opin Cell Biol* 40, 106-13.
- 19 8. Putnam, C.D. and Kolodner, R.D. (2017) Pathways and mechanisms that prevent genome instability  
20 in *Saccharomyces cerevisiae*. *Genetics* 206 (3), 1187-1225.
- 21 9. Putnam, C.D. et al. (2009) Specific pathways prevent duplication-mediated genome rearrangements.  
22 *Nature* 460 (7258), 984-989.
- 23 10. Chan, J.E. and Kolodner, R.D. (2011) A genetic and structural study of genome rearrangements  
24 mediated by high copy repeat Ty1 elements. *PLoS Genet* 7 (5).
- 25 11. Argueso, J.L. et al. (2008) Double-strand breaks associated with repetitive DNA can reshape the  
26 genome. *Proc Natl Acad Sci U S A* 105 (33), 11845-50.
- 27 12. Hoang, M.L. et al. (2010) Competitive repair by naturally dispersed repetitive DNA during non-  
28 allelic homologous recombination. *PLoS Genet* 6 (12).
- 29 13. Koszul, R. et al. (2004) Eucaryotic genome evolution through the spontaneous duplication of large  
30 chromosomal segments. *EMBO J* 23 (1), 234-43.
- 31 14. Payen, C. et al. (2008) Segmental duplications arise from Pol32-dependent repair of broken forks  
32 through two alternative replication-based mechanisms. *PLoS Genet* 4 (9), e1000175.
- 33 15. Brown, C.J. et al. (1998) Multiple duplications of yeast hexose transport genes in response to  
34 selection in a glucose-limited environment. *Mol Biol Evol* 15 (8), 931-42.
- 35 16. Dunham, M.J. et al. (2002) Characteristic genome rearrangements in experimental evolution of  
36 *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 99 (25), 16144-16149.
- 37 17. Putnam, C.D. et al. (2016) A genetic network that suppresses genome rearrangements in  
38 *Saccharomyces cerevisiae* and contains defects in cancers. *Nat Commun* 7, 11256.
- 39 18. Lisby, M. et al. (2001) Rad52 forms DNA repair and recombination centers during S phase. *Proc*  
40 *Natl Acad Sci U S A* 98 (15), 8276-82.
- 41 19. Branzei, D. and Foiani, M. (2010) Maintaining genome stability at the replication fork. *Nat Rev*  
42 *Mol Cell Biol* 11 (3), 208-219.
- 43 20. Haber, J.E. (2012) Mating-type genes and *MAT* switching in *Saccharomyces cerevisiae*. *Genetics*  
44 191 (1), 33-64.
- 45 21. Verma, P. and Greenberg, R.A. (2016) Noncanonical views of homology-directed DNA repair.  
46 *Genes Dev.* 30 (10), 1138-1154.
- 47 22. Pannunzio, N.R. et al. (2018) Nonhomologous DNA end-joining for repair of DNA double-strand  
48 breaks. *J Biol Chem* 293 (27), 10512-10523.
- 49 23. Sallmyr, A. and Tomkinson, A.E. (2018) Repair of DNA double-strand breaks by mammalian  
50 alternative end-joining pathways. *J Biol Chem* 293 (27), 10536-10546.

- 1 24. Bhargava, R. et al. (2016) Regulation of Single-Strand Annealing and its Role in Genome  
2 Maintenance. *Trends Genet* 32 (9), 566-575.
- 3 25. Bell, J.C. and Kowalczykowski, S.C. (2016) RecA: regulation and mechanism of a molecular  
4 search engine. *Trends Biochem Sci* 41 (6), 491-507.
- 5 26. Mizuno, K. et al. (2013) Recombination-restarted replication makes inverted chromosome fusions  
6 at inverted repeats. *Nature* 493 (7431), 246-249.
- 7 27. Lambert, S. et al. (2010) Homologous recombination restarts blocked replication forks at the  
8 expense of genome rearrangements by template exchange. *Mol Cell* 39 (3), 346-59.
- 9 28. Szostak, J.W. et al. (1983) The double-strand-break repair model for recombination. *Cell* 33, 25-  
10 35.
- 11 29. Juhasz, S. et al. (2018) ATRX Promotes DNA Repair Synthesis and Sister Chromatid Exchange  
12 during Homologous Recombination. *Molecular Cell* 71 (1), 11-24.
- 13 30. Bzymek, M. et al. (2010) Double Holliday junctions are intermediates of DNA break repair.  
14 *Nature* 464, 937-941.
- 15 31. Schwacha, A. and Kleckner, N. (1995) Identification of double Holliday junctions as intermediates  
16 in meiotic recombination. *Cell* 83, 783-791.
- 17 32. Wu, L.J. and Hickson, I.D. (2003) The Bloom's syndrome helicase suppresses crossing-over  
18 during homologous recombination. *Nature* 426, 870-874.
- 19 33. Schwartz, E.K. and Heyer, W.D. (2011) Processing of joint molecule intermediates by structure-  
20 selective endonucleases during homologous recombination in eukaryotes. *Chromosoma* 120, 109-127.
- 21 34. Malkova, A. et al. (1996) Double-strand break repair in the absence of *RAD51* in yeast: a possible  
22 role for break-induced DNA replication. *Proc Natl Acad Sci U S A* 93 (14), 7131-7136.
- 23 35. Mayle, R. et al. (2015) Mus81 and converging forks limit the mutagenicity of replication fork  
24 breakage. *Science* 349 (6249), 742-747.
- 25 36. Anand, R.P. et al. (2013) Break-induced DNA replication. *Cold Spring Harbor Persp Biol* 5 (12),  
26 a010397.
- 27 37. Donnianni, R.A. and Symington, L.S. (2013) Break-induced replication occurs by conservative  
28 DNA synthesis. *Proc Natl Acad Sci U S A* 110 (33), 13475-13480.
- 29 38. Saini, N. et al. (2013) Migrating bubble during break-induced replication drives conservative DNA  
30 synthesis. *Nature* 502 (7471), 389-392.
- 31 39. Elango, R. et al. (2017) Break-induced replication promotes formation of lethal joint molecules  
32 dissolved by Srs2. *Nat Commun* 8 (1), 1790.
- 33 40. Sakofsky, C.J. et al. (2014) Break-induced replication is a source of mutation clusters underlying  
34 kataegis. *Cell Rep* 7 (5), 1640-8.
- 35 41. Smith, C.E. et al. (2007) Template switching during break-induced replication. *Nature* 447 (7140),  
36 102-105.
- 37 42. Anand, R.P. et al. (2014) Chromosome rearrangements via template switching between diverged  
38 repeated sequences. *Genes Dev.* 28 (21), 2394-2406.
- 39 43. Stafa, A. et al. (2014) Template switching during break-induced replication is promoted by the  
40 Mph1 helicase in *Saccharomyces cerevisiae*. *Genetics* 196 (4), 1017-28.
- 41 44. Pardo, B. and Aguilera, A. (2012) Complex chromosomal rearrangements mediated by break-  
42 induced replication involve structure-selective endonucleases. *PLoS Genet.* 8 (9).
- 43 45. Hicks, W.M. et al. (2010) Increased mutagenesis and unique mutation signature associated with  
44 mitotic gene conversion. *Science* 329 (5987), 82-85.
- 45 46. Wilson, M.A. et al. (2013) Pif1 helicase and Poldelta promote recombination-coupled DNA  
46 synthesis via bubble migration. *Nature* 502 (7471), 393-6.
- 47 47. Lydeard, J.R. et al. (2007) Break-induced replication and telomerase-independent telomere  
48 maintenance require Pol32. *Nature* 448 (7155), 820-823.
- 49 48. Lydeard, J.R. et al. (2010) Break-induced replication requires all essential DNA replication factors  
50 except those specific for pre-RC assembly. *Genes Dev.* 24 (11), 1133-1144.
- 51 49. Deem, A. et al. (2008) Defective break-induced replication leads to half-crossovers in  
52 *Saccharomyces cerevisiae*. *Genetics* 179 (4), 1845-1860.
- 53 50. Smith, C.E. et al. (2009) aberrant double-strand break repair resulting in half crossovers in mutants  
54 defective for Rad51 or the DNA polymerase delta complex. *Mol Cell Biol* 29 (6), 1432-1441.



- 1 51. Hum, Y.F. and Jinks-Robertson, S. (2018) DNA strand-exchange patterns associated with double-  
2 strand break-induced and spontaneous mitotic crossovers in *Saccharomyces cerevisiae*. PLoS Genet  
3 14 (3), e1007302.
- 4 52. Piazza, A. et al. (2017) Multi-invasions are recombination byproducts that induce chromosomal  
5 rearrangements. Cell 170 (4), 760-773 e15.
- 6 53. Wright, W.D. and Heyer, W.D. (2014) Rad54 functions as a heteroduplex DNA pump modulated  
7 by its DNA substrates and Rad51 during D-loop formation. Mol Cell 53 (3), 420-432.
- 8 54. Forget, A.L. and Kowalczykowski, S.C. (2012) Single-molecule imaging of DNA pairing by RecA  
9 reveals a three-dimensional homology search. Nature 482 (7385), 423-427.
- 10 55. Wright, W.D. et al. (2018) Homologous recombination and the repair of DNA double-strand  
11 breaks. J Biol Chem 293 (27), 10524-10535.
- 12 56. Piazza, A. and Heyer, W.D. (2018) Multi-invasion-induced rearrangements as a pathway for  
13 physiological and pathological recombination. Bioessays 40 (5), e1700249.
- 14 57. Heyer, W.D. (2015) Regulation of recombination and genomic maintenance. Cold Spring Harb  
15 Perspect Biol 7 (8), a016501.
- 16 58. Veaute, X. et al. (2003) The Srs2 helicase prevents recombination by disrupting Rad51  
17 nucleoprotein filaments. Nature 423, 309-312.
- 18 59. Kaniecki, K. et al. (2017) Dissociation of Rad51 presynaptic complexes and heteroduplex DNA  
19 joints by tandem assemblies of Srs2. Cell Rep 21 (11), 3166-3177.
- 20 60. Fabre, F. et al. (2002) Alternate pathways involving Sgs1/Top3, Mus81/Mms4, and Srs2 prevent  
21 formation of toxic recombination intermediates from single-stranded gaps created by DNA replication.  
22 Proc Natl Acad Sci U S A 99, 16887-16892.
- 23 61. Prakash, R. et al. (2009) Yeast Mph1 helicase dissociates Rad51-made D-loops: implications for  
24 crossover control in mitotic recombination. Genes Dev 23 (1), 67-79.
- 25 62. Liu, J. et al. (2017) Srs2 promotes synthesis-dependent strand annealing by disrupting DNA  
26 polymerase delta-extending D-loops. Elife 6.
- 27 63. Fasching, C.L. et al. (2015) Top3-Rmi1 dissolve Rad51-mediated D-loops by a topoisomerase-  
28 based mechanism. Mol Cell 57, 595-606.
- 29 64. Ira, G. et al. (2003) Srs2 and Sgs1-Top3 suppress crossovers during double-strand break repair in  
30 yeast. Cell 115 (4), 401-411.
- 31 65. Mazon, G. and Symington, L.S. (2013) Mph1 and Mus81-Mms4 prevent aberrant processing of  
32 mitotic recombination intermediates. Mol Cell 52 (1), 63-74.
- 33 66. Mitchel, K. et al. (2013) Heteroduplex DNA position defines the roles of the Sgs1, Srs2, and Mph1  
34 helicases in promoting distinct recombination outcomes. PLoS Genet. 9 (3).
- 35 67. Cejka, P. et al. (2010) Rmi1 stimulates decatenation of double Holliday junctions during  
36 dissolution by Sgs1-Top3. Nature Struct. Mol. Biol. 17, 1377-1382.
- 37 68. Mimitou, E.P. and Symington, L.S. (2008) Sae2, Exo1 and Sgs1 collaborate in DNA double-strand  
38 break processing. Nature 455, 770-774.
- 39 69. Symington, L.S. (2016) Mechanism and regulation of DNA end resection in eukaryotes. Crit Rev  
40 Biochem Mol Biol 51 (3), 195-212.
- 41 70. Mirman, Z. et al. (2018) 53BP1-RIF1-shieldin counteracts DSB resection through CST- and  
42 Polalpha-dependent fill-in. Nature 560 (7716), 112-116.
- 43 71. Hashimoto, Y. et al. (2010) Rad51 protects nascent DNA from Mre11-dependent degradation and  
44 promotes continuous DNA synthesis. Nat Struct Mol Biol 17 (11), 1305-11.
- 45 72. Schlacher, K. et al. (2011) Double-strand break repair-independent role for BRCA2 in blocking  
46 stalled replication fork degradation by MRE11. Cell 145 (4), 529-42.
- 47 73. Sogo, J.M. et al. (2002) Fork reversal and ssDNA accumulation at stalled replication forks owing  
48 to checkpoint defects. Science 297, 599-602.
- 49 74. Graham, J.E. et al. (2017) Independent and stochastic action of DNA polymerases in the  
50 replisome. Cell 169 (7), 1201-1213.
- 51 75. Yeeles, J.T. and Marians, K.J. (2013) Dynamics of leading-strand lesion skipping by the  
52 replisome. Mol Cell 52 (6), 855-65.
- 53 76. Lemoine, F.J. et al. (2005) Chromosomal translocations in yeast induced by low levels of DNA  
54 polymerase a model for chromosome fragile sites. Cell 120 (5), 587-98.

- 1 77. Hartwell, L.H. and Weinert, T.A. (1989) Checkpoints: controls that ensure order of cell cycle  
2 events. *Science* 246, 629-634.
- 3 78. Kato, H. and Sandberg, A.A. (1968) Chromosome pulverization in human cells with micronuclei. *J*  
4 *Natl Cancer Inst* 40 (1), 165-79.
- 5 79. Letessier, A. et al. (2011) Cell-type-specific replication initiation programs set fragility of the  
6 *FRA3B* fragile site. *Nature* 470 (7332), 120-3.
- 7 80. Ying, S. et al. (2013) MUS81 promotes common fragile site expression. *Nat Cell Biol* 15 (8),  
8 1001-1007.
- 9 81. Naim, V. et al. (2013) ERCC1 and MUS81-EME1 promote sister chromatid separation by  
10 processing late replication intermediates at common fragile sites during mitosis. *Nat Cell Biol* 15 (8),  
11 1008-1015.
- 12 82. Techer, H. et al. (2016) Signaling from Mus81-Eme2-dependent DNA damage elicited by Chk1  
13 deficiency modulates replication fork speed and origin usage. *Cell Rep* 14 (5), 1114-1127.
- 14 83. Szakal, B. and Branzei, D. (2013) Premature Cdk1/Cdc5/Mus81 pathway activation induces  
15 aberrant replication and deleterious crossover. *EMBO J* 32 (8), 1155-1167.
- 16 84. Matos, J. et al. (2013) Cell-cycle kinases coordinate the resolution of recombination intermediates  
17 with chromosome segregation. *Cell Rep* 4 (1), 76-86.
- 18 85. Atianand, M.K. and Fitzgerald, K.A. (2013) Molecular basis of DNA recognition in the immune  
19 system. *J Immunol* 190 (5), 1911-8.
- 20 86. Chowdhury, D. et al. (2006) The exonuclease TREX1 is in the SET complex and acts in concert  
21 with NM23-H1 to degrade DNA during granzyme A-mediated cell death. *Mol Cell* 23 (1), 133-42.
- 22 87. Fan, Z. et al. (2003) Tumor suppressor NM23-H1 is a granzyme A-activated DNase during CTL-  
23 mediated apoptosis, and the nucleosome assembly protein SET is its inhibitor. *Cell* 112 (5), 659-72.
- 24 88. Li, L.Y. et al. (2001) Endonuclease G is an apoptotic DNase when released from mitochondria.  
25 *Nature* 412 (6842), 95-9.
- 26 89. West, S.C. et al. (2015) Resolution of recombination intermediates: Mechanisms and regulation.  
27 *Cold Spring Harb Symp Quant Biol* 80, 103-9.
- 28 90. Blanco, M.G. et al. (2014) Dual control of Yen1 nuclease activity and cellular localization by Cdk  
29 and Cdc14 prevents genome instability. *Mol Cell* 54 (1), 94-106.
- 30 91. Eissler, C.L. et al. (2014) The Cdk/Cdc14 module controls activation of the Yen1 Holliday  
31 junction resolvase to promote genome stability. *Mol Cell* 54 (1), 80-93.
- 32 92. Ly, P. et al. (2017) Selective Y centromere inactivation triggers chromosome shattering in  
33 micronuclei and repair by non-homologous end joining. *Nat Cell Biol* 19 (1), 68-75.
- 34 93. Zhang, C.Z. et al. (2015) Chromothripsis from DNA damage in micronuclei. *Nature* 522 (7555),  
35 179-84.
- 36 94. Maciejowski, J. et al. (2015) Chromothripsis and kataegis induced by telomere crisis. *Cell* 163 (7),  
37 1641-1654.
- 38 95. Crasta, K. et al. (2012) DNA breaks and chromosome pulverization from errors in mitosis. *Nature*  
39 482 (7383), 53-8.
- 40 96. Houchmandzadeh, B. et al. (1997) Elasticity and structure of eukaryote chromosomes studied by  
41 micromanipulation and micropipette aspiration. *J Cell Biol* 139 (1), 1-12.
- 42 97. Cluzel, P. et al. (1996) DNA: an extensible molecule. *Science* 271 (5250), 792-4.
- 43 98. Smith, S.B. et al. (1996) Overstretching B-DNA: the elastic response of individual double-stranded  
44 and single-stranded DNA molecules. *Science* 271 (5250), 795-9.
- 45 99. Grandbois, M. et al. (1999) How strong is a covalent bond? *Science* 283 (5408), 1727-30.
- 46 100. Nik-Zainal, S. et al. (2012) Mutational processes molding the genomes of 21 breast cancers. *Cell*  
47 149 (5), 979-993.
- 48 101. Chan, K. and Gordenin, D.A. (2015) Clusters of multiple mutations: incidence and molecular  
49 mechanisms. *Annu Rev Genet* 49, 243-67.
- 50 102. Nik-Zainal, S. et al. (2016) Landscape of somatic mutations in 560 breast cancer whole-genome  
51 sequences. *Nature* 534 (7605), 47-54.
- 52 103. Quigley, D.A. et al. (2018) Genomic hallmarks and structural variation in metastatic prostate  
53 cancer. *Cell* 174 (3), 758-769 e9.
- 54 104. Ratnaparkhe, M. et al. (2018) Defective DNA damage repair leads to frequent catastrophic  
55 genomic events in murine and human tumors. *bioRxiv*.

- 1 105. Orthwein, A. et al. (2015) A mechanism for the suppression of homologous recombination in G1  
2 cells. *Nature* 528 (7582), 422-6.
- 3 106. Mekeel, K.L. et al. (1997) Inactivation of p53 results in high rates of homologous recombination.  
4 *Oncogene* 14 (15), 1847-57.
- 5 107. Arias-Lopez, C. et al. (2006) p53 modulates homologous recombination by transcriptional  
6 regulation of the *RAD51* gene. *EMBO Rep* 7 (2), 219-24.
- 7 108. Hollstein, M. et al. (1991) p53 mutations in human cancers. *Science* 253 (5015), 49-53.
- 8 109. Rausch, T. et al. (2012) Genome sequencing of pediatric medulloblastoma links catastrophic  
9 DNA rearrangements with TP53 mutations. *Cell* 148 (1-2), 59-71.
- 10 110. Carvalho, C.M. and Lupski, J.R. (2016) Mechanisms underlying structural variant formation in  
11 genomic disorders. *Nat Rev Genet* 17 (4), 224-38.
- 12 111. Sasaki, M. et al. (2010) Genome destabilization by homologous recombination in the germ line.  
13 *Nat Rev Mol Cell Biol* 11 (3), 182-195.
- 14 112. Wala, J.A. et al. (2017) Selective and mechanistic sources of recurrent rearrangements across the  
15 cancer genome. *bioRxiv*.
- 16 113. van Heesch, S. et al. (2014) Genomic and functional overlap between somatic and germline  
17 chromosomal rearrangements. *Cell Rep* 9 (6), 2001-10.
- 18 114. Serero, A. et al. (2014) Mutational landscape of yeast mutator strains. *Proc Natl Acad Sci U S A*  
19 111 (5), 1897-1902.
- 20 115. Cretu Stancu, M. et al. (2017) Mapping and phasing of structural variation in patient genomes  
21 using nanopore sequencing. *Nat Commun* 8 (1), 1326.
- 22 116. Harewood, L. et al. (2017) Hi-C as a tool for precise detection and characterisation of  
23 chromosomal rearrangements and copy number variation in human tumours. *Genome Biol* 18 (1), 125.
- 24 117. Monlong, J. et al. (2018) Human copy number variants are enriched in regions of low  
25 mappability. *bioRxiv*.
- 26 118. Dixon, J.R. et al. (2018) Integrative detection and analysis of structural variation in cancer  
27 genomes. *Nat Genet* 50 (10), 1388-1398.
- 28 119. Nattestad, M. et al. (2018) Complex rearrangements and oncogene amplifications revealed by  
29 long-read DNA and RNA sequencing of a breast cancer cell line. *Genome Res* 28 (8), 1126-1135.
- 30 120. Sedlazeck, F.J. et al. (2018) Accurate detection of complex structural variations using single-  
31 molecule sequencing. *Nat Methods* 15 (6), 461-468.
- 32 121. Chaisson, M.J.P. et al. (2018) Multi-platform discovery of haplotype-resolved structural variation  
33 in human genomes. *bioRxiv*.
- 34 122. McClintock, B. (1941) The stability of broken ends of chromosomes in *Zea Mays*. *Genetics* 26,  
35 234-282.
- 36 123. Hatch, E.M. et al. (2013) Catastrophic nuclear envelope collapse in cancer cell micronuclei. *Cell*  
37 154 (1), 47-60.
- 38 124. Liu, S. et al. (2018) Nuclear envelope assembly defects link mitotic errors to chromothripsis.  
39 *bioRxiv*.
- 40 125. Kloosterman, W.P. et al. (2011) Chromothripsis as a mechanism driving complex *de novo*  
41 structural rearrangements in the germline. *Hum Mol Genet* 20 (10), 1916-24.
- 42 126. Liu, P. et al. (2012) Mechanisms for recurrent and complex human genomic rearrangements. *Curr*  
43 *Opin Genet Dev* 22 (3), 211-20.
- 44 127. Berger, M.F. et al. (2011) The genomic complexity of primary human prostate cancer. *Nature*  
45 470 (7333), 214-20.
- 46 128. Wang, J. et al. (2007) Evidence for mutation showers. *Proc Natl Acad Sci U S A* 104 (20), 8403-  
47 8.
- 48 129. Roberts, S.A. et al. (2012) Clustered mutations in yeast and in human cancers can arise from  
49 damaged long single-strand DNA regions. *Mol Cell* 46 (4), 424-435.
- 50 130. Mirman, Z. et al. (2018) 53BP1-RIF1-shieldin counteracts DSB resection through CST- and Pol  
51 alpha-dependent fill-in. *Nature* 560 (7716), 112-6.

52

53

## 1 **Highlights**

- 2 • Homologous recombination generates genome rearrangements involving repeated DNA  
3 elements with identical (homologous) or near identical (homeologous) sequences that  
4 can be located anywhere in the genome.
- 5 • Multi-invasions are recombination byproducts that physically bridge two copies of a  
6 repeated DNA element that can be processed by structure-selective endonucleases into  
7 genome rearrangements and additional DNA double-stranded breaks.
- 8 • Different chromothripsis-inducing contexts in mammalian cells feature defective  
9 isolation of genomic DNA from cytoplasmic components, including various types of  
10 nucleases.
- 11 • Long DNA sequence read assemblies paired with additional approaches and improved  
12 bioinformatic pipelines are needed to fully evaluate the contributions of homologous  
13 recombination between repeated DNA to generate genomic rearrangements.

14

## 15 **Glossary**

16 **Breakage-Fusion-Bridge cycle:** Breakage-Fusion-Bridge cycle denotes a genetic  
17 phenomenon first described by Barbara McClintock of telomere fusions generating a dicentric  
18 chromosome that cannot be properly segregated during anaphase leading to breakage and a new  
19 cycle of fusions [122]. Any genome rearrangements between sister chromatids, homologs or  
20 different chromosomes resulting in a dicentric chromosomes will lead to chronic genomic  
21 instability until stabilization of the broken fragments.

22 **Conversion/Crossover/half-crossover/non-crossover:** Conversion is a non-reciprocal  
23 exchange of genetic information. Crossover is the reciprocal exchange of genetic markers  
24 during HR, where both reciprocal products are recovered. In somatic cells, crossover is actively  
25 avoided by using the SDSA pathway and dissolution of double Holliday junctions, leading to  
26 non-crossovers which often involve conversion (**Fig. 1**). Half-crossover is an event where one  
27 of the expected reciprocal products is lost.

28 **D-loop (Displacement loop):** D-loops are a central HR intermediate, whose processing  
29 determines the HR subpathway (SDSA, DSBR, or BIR) and crossover/non-crossover outcome  
30 (**Fig. 1**). D-loops are formed by Rad51-ssDNA filament through DNA strand invasion. Nascent

1 and extended D-loop denotes D-loops prior and after extension of the 3'-OH of the invading  
2 DNA strand by DNA polymerase, respectively.

3 **Double-strand break (DSB):** A DSB consists in the interruption of the phospho-diester  
4 backbone of two complementary DNA strands at the same or nearby position. DSBs come in  
5 two flavors: either single-ended such as upon replication fork breakage or frank (two-sided)  
6 (**Fig. 1**). Single-ended DSBs can be accurately repaired only by HR.

7 **Micronucleus:** Micronuclei are a classic hallmark of genotoxic stress and form when a  
8 chromosome or chromosome fragment is not incorporated into the two daughter nuclei after  
9 anaphase. Micronuclei are nuclear-like structures that features defective isolation of the  
10 nucleoplasm from the cytoplasm (**Fig. 4**) [93, 95, 123]. Micronuclei are depleted for nuclear  
11 pore complexes and other NE proteins, which is suspected to dysregulate nuclear protein  
12 homeostasis and underlie micronuclei-specific NE fragility (**Fig. 4**) [124].

13 **Structure-selective endonucleases (SSE):** Non-linear DNA structures including intermediates  
14 generated during HR (flaps, D-loops or HJs) and stalled replications forks are substrates for  
15 SSEs. Mostly studied for their role in genome maintenance are Mus81-Mms4, Slx1-Slx4, Rad1-  
16 Rad10 and Yen1 in *S. cerevisiae*, respectively MUS81-EME1/2, SLX1-SLX4, XPF-ERCC1  
17 and GEN1 in humans.

18

## 19 **Outstanding Questions**

- 20 • What is the role of SSEs and cytoplasmic nucleases in generating massive DNA damage  
21 and the complex rearrangements of chromothripsis?
- 22 • What are the DNA repair pathways involved in different damage-inducing contexts?  
23 Especially for the repair of long ssDNAs?
- 24 • Are HR regulators involved in suppressing chromothripsis?
- 25 • Does MIR operates in human cells? In which context? Is there a mutational signature  
26 unique to MIR?
- 27 • What additional types of SVs and junction sequences will be uncovered by long-read  
28 assemblies of genomes with simple and complex rearrangements?

29

## 30 **Box: Mutational Phenomena**

1 **Chromothripsis** was originally described by Campbell and colleagues as massive focal  
2 genome rearrangements in a patient with Chronic Lymphocytic Leukemia [1] and is  
3 associated with a significant and variable proportion of cancer genomes depending on cancer  
4 type [2] . Independently, chromothripsis was identified as a constitutional genetic change in a  
5 child with severe congenital abnormalities [125]. The term derives from Greek (chromos for  
6 chromosome; thripsis, shattering into pieces) describing the interpretation that chromothripsis  
7 results from catastrophic chromosomal breakage into many individual DNA fragments.

8

9 **Chromoanasythesis** is a constitutionally acquired complex genome arrangement  
10 phenomenon that was discovered in individuals with developmental delay and cognitive  
11 anomalies [126]. The patterns of CNVs and breakpoint junctions differ from chromothripsis,  
12 and the breakpoint sequence analysis suggested repeated template switching at  
13 microhomologies as the process leading to the observed rearrangements.

14

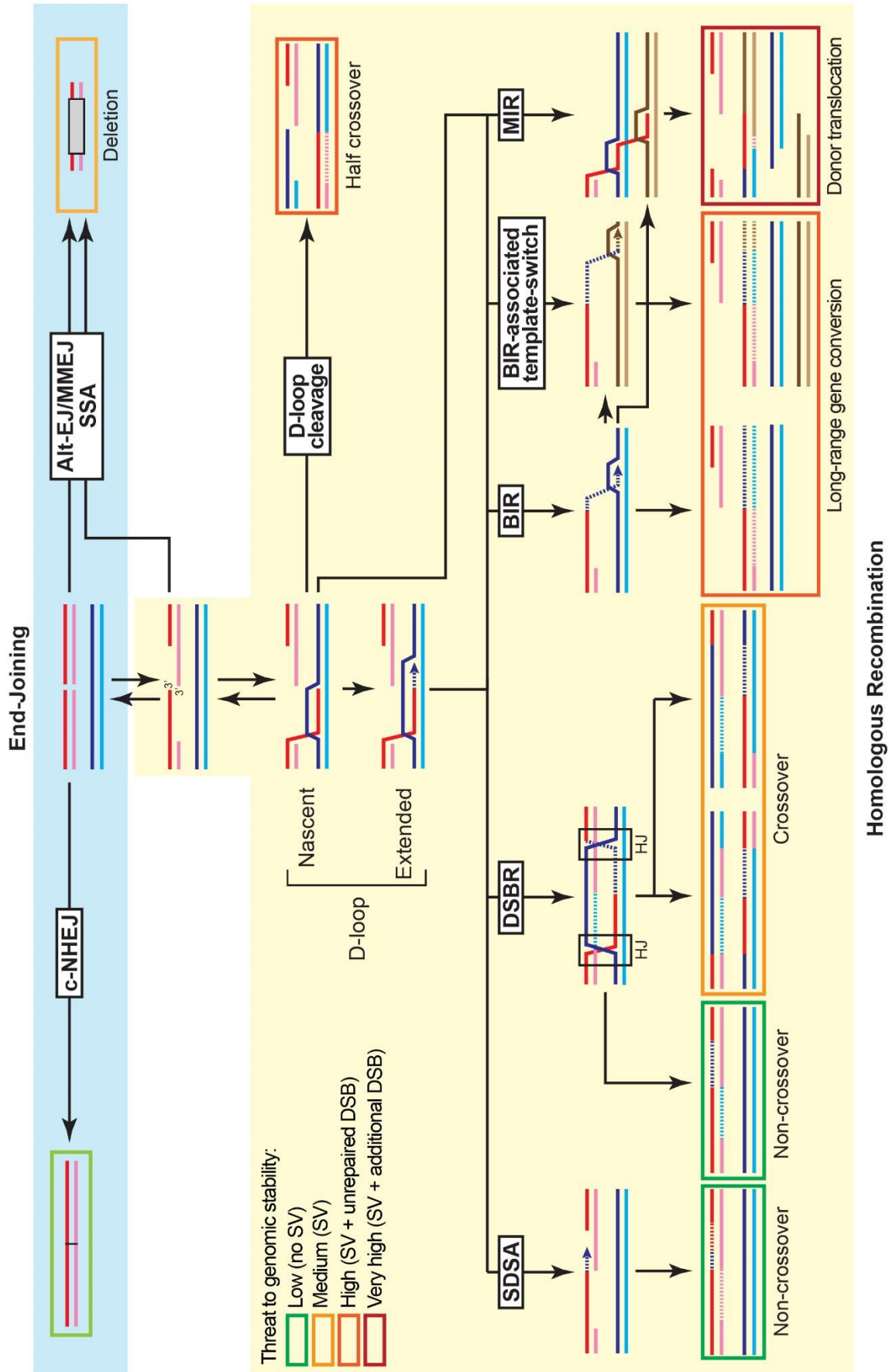
15 **Chromoplexy** describes yet another pattern of complex genome rearrangements involving a  
16 connected chain of translocations which first discovered in prostate tumors [127].

17

18 **Kataegis** describes localized hypermutation of single nucleotide changes identified in mice  
19 and in some cancer genomes [101, 128]. Kataegis is the C-to-T mutagenic signature of  
20 APOBEC deaminases that act on single-stranded DNA, suggesting that kataegis signals  
21 regions of long-lived single-stranded DNA intermediates [40, 101, 129].

22

1 **Figures**



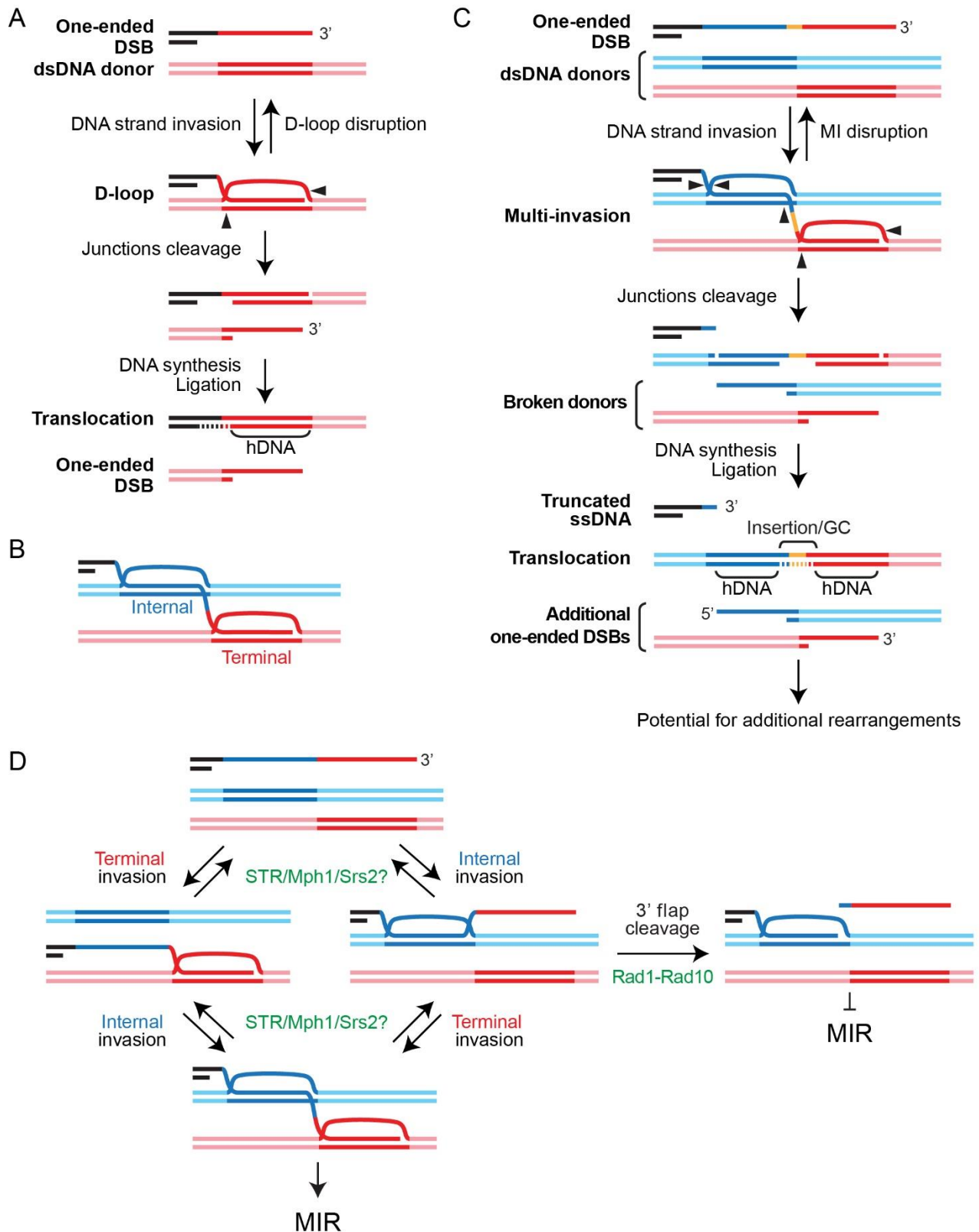
2

1 **Figure 1: Overview of the DSB repair pathways and their associated risk for genomic**  
2 **stability.** The repair products are boxed. The box color (from green to red) indicates the threat  
3 to genomic stability posed by the product of each pathway and sub-pathway. DNA synthesis is  
4 indicated by an arrow and newly-synthesized DNA by a dotted line. The long ssDNA associated  
5 with BIR has the potential to undergo MIR. More detailed mechanisms for D-loop cleavage  
6 and MIR are provided in **Figures 2A** and **2C**, respectively. The reversibility of DSB resection  
7 provided by the Shieldin complex with fill-in by the DNA polymerase  $\alpha$ -primase complex  
8 provides an unanticipated degree of flexibility in the choice of DSB repair between HR and EJ  
9 mechanisms [130]. Hence, EJ pathways are available not only on unresected or minimally  
10 resected DSB, they can also be engaged after extensive resection and subsequent fill-in.

11

12



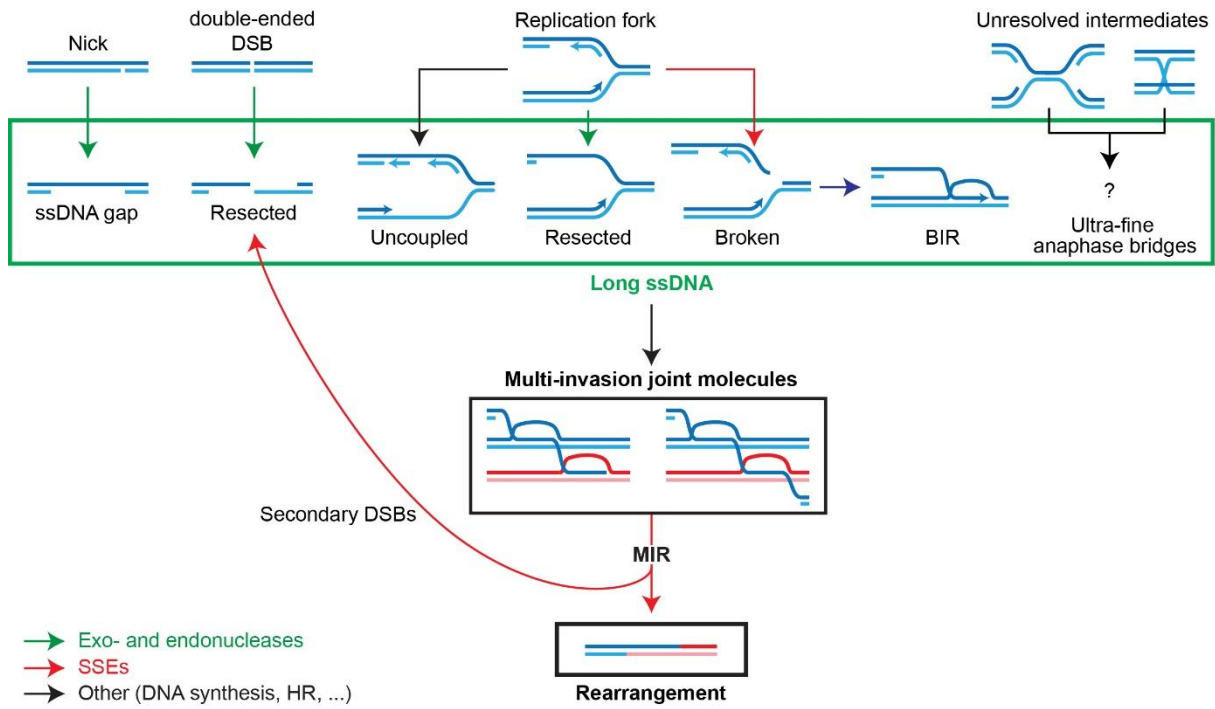


1

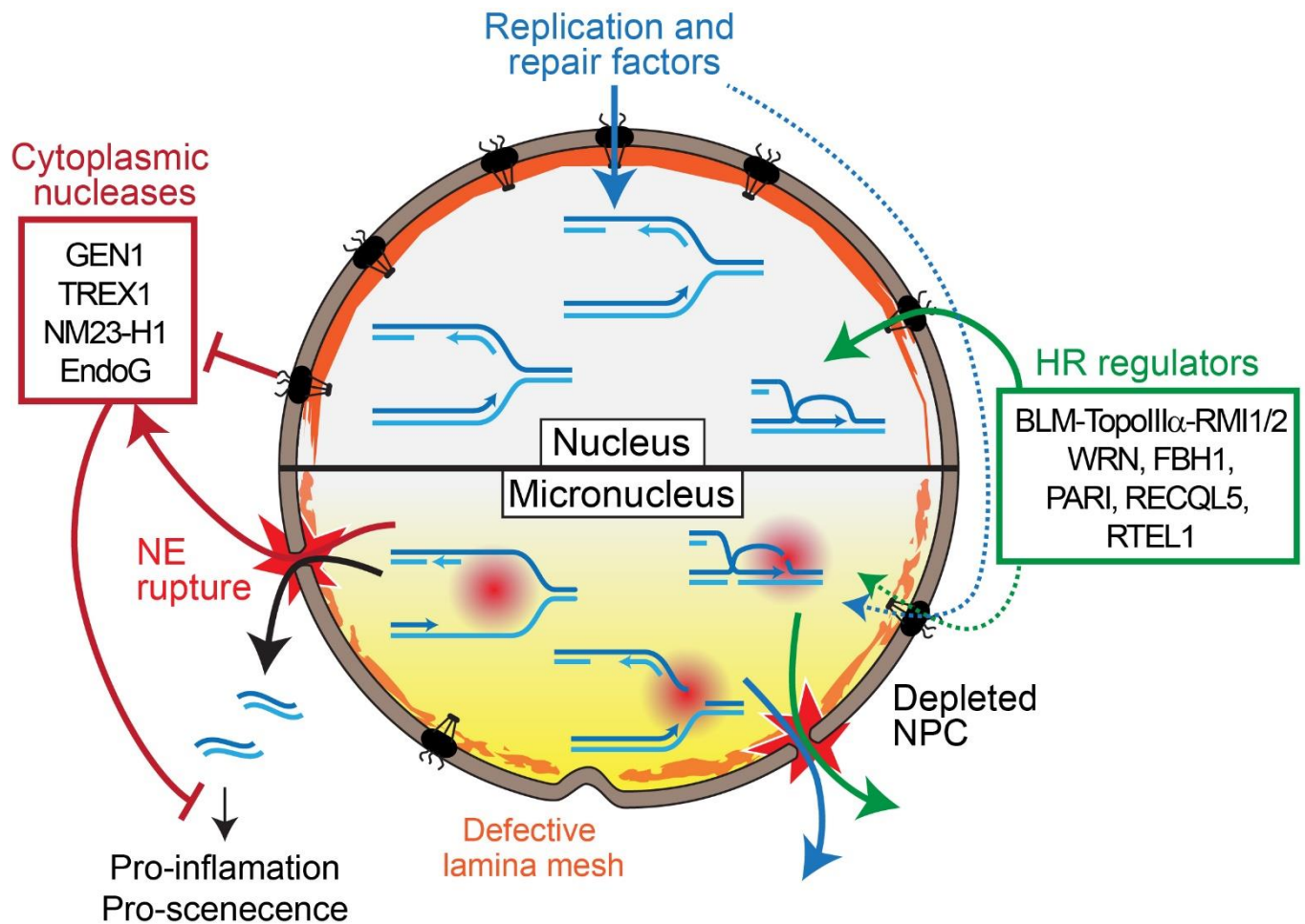
2 **Figure 2: Single and multi-invasion joint molecules are at-risk HR intermediates for SSE-**  
 3 **mediated genomic instability.** (A) Model for half-crossover formation from cleavage of a  
 4 single D-loop. (B) Multi-invasions (MI) joint molecules are formed when a Rad51-ssDNA  
 5 filament invades two independent donors along its length. It features an internal and a terminal

1 D-loop. (C) Model for multi-invasion-induced rearrangement (MIR). (D) Several activities  
2 inhibit MIR in a reversible (Sgs1-Top2-Rmi1 (STR), Mph1, Srs2) or irreversible (Rad1-Rad10)  
3 fashion. The specificity of STR, Mph1 and Srs2 as well as the precise nature of their substrates  
4 has not been established.

5



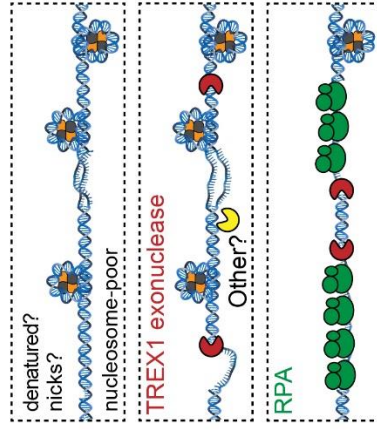
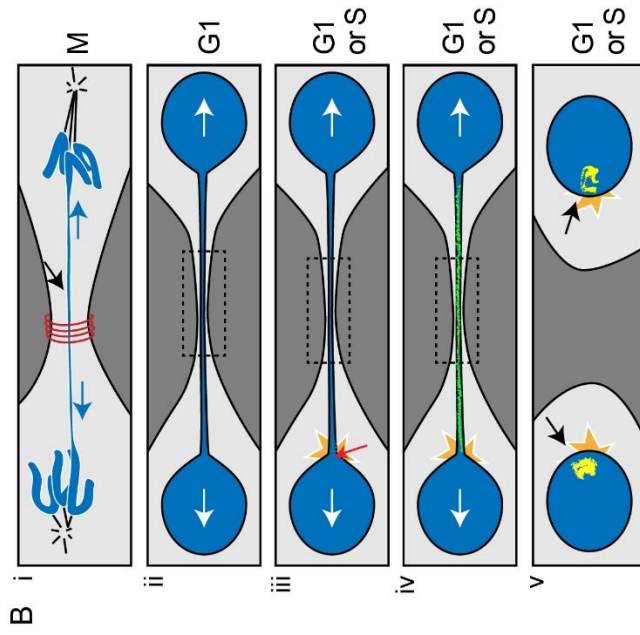
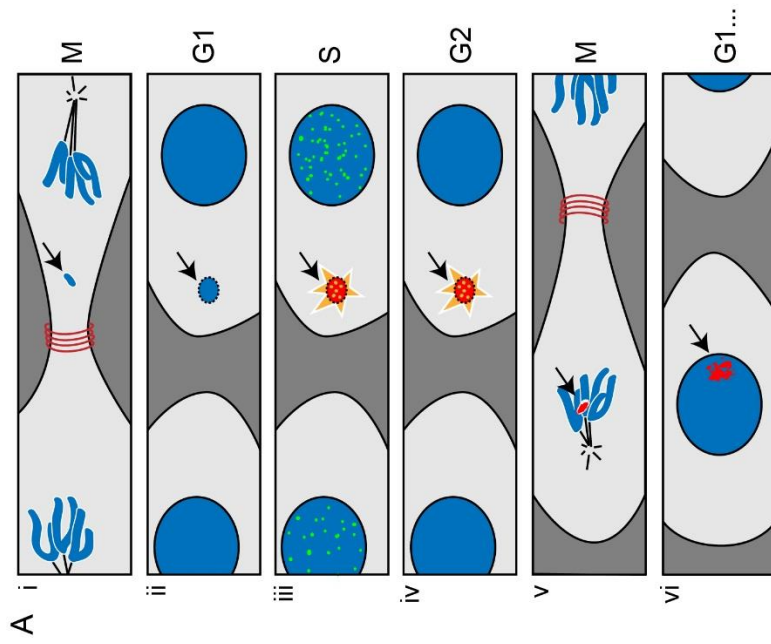
**Figure 3: Nuclease involvement in MIR.** Exo- and endonucleases involved in resection (green) generate long recombinogenic substrates, while SSEs (red) process MI joint molecules into rearrangements and additional resected secondary single-ended DSBs (indicated by backwards arrow to resected DSBs).



1

2 **Figure 4: S-phase-specific DNA damage from defective nucleoplasm isolation and DNA**  
 3 **metabolism in micronuclei.** The nuclear envelope (NE) of micronuclei is depleted for nuclear  
 4 pore complex and features abnormal lamina deposition. These dysregulations lead to abnormal  
 5 nuclear protein import and frequent NE rupture, respectively. NE rupture leads to the  
 6 penetration of cytoplasmic components to the micronucleus and the leakage of soluble nuclear  
 7 components and DNA.

8



1 **Figure 5: Cellular models for chromothripsis.** (A) Chromothripsis by isolation and  
2 breakage of chromosomes in micronuclei [92, 93]. (i-ii) Mis-segregation and physical  
3 separation of a large chromatin fragment leads to micronucleus formation. (ii-iii) Micronuclei  
4 are defective for the nuclear import and barrier function (**Fig. 4**). (iii) Upon S-phase entry, NE  
5 rupture leads to the formation of massive DNA damage, which features co-localizing DSBs  
6 (green) and  $\gamma$ H2AX foci [95] (iv-v) The micronuclear DNA reaches mitosis under-replicated  
7 and fragmented. (v-vi) Micronuclear DNA fragments can be re-incorporated into the main  
8 nucleus at subsequent mitosis where it may undergo final repair. (B) Chromothripsis originating  
9 from attempted segregation of dicentric chromosomes [94]. (i) Spindle-exerted tensions can  
10 stretch but not break the central portion of a dicentric chromosome. (ii) NE forms around the  
11 daughter nuclei and the intervening chromatin bridge, depleted for lamin and nuclear pore  
12 complex. Right inset: the bridge is partially depleted for nucleosomes. (iii) Tensions exerted on  
13 the bridge cause NE rupture at the base of the bridge in G1/S phase. NE rupture leads to leaking  
14 of nuclear components and penetration of cytoplasmic proteins such as the TREX1 exonuclease  
15 (right inset). (iv) TREX1 exploits nicks or ssDNA gaps present in the bridge or generated by  
16 other cytoplasmic endonuclease to resect the chromatin bridge (right inset). This leads to  
17 accumulation of RPA foci (green) and (v) culminates in the resolution of the bridge. In addition  
18 to RPA, the resected DNA causes the formation of 53BP1,  $\gamma$ H2AX and Mre11 foci (yellow) in  
19 the daughter nuclei.

20

21