Published online 9 May 2023

NAR Cancer, 2023, Vol. 5, No. 2 **1**
https://doi.org/10.1093/narcan/zcad019

Molecular mechanisms protecting centromeres from self-sabotage and implications for cancer therapy

Rim Nassar[1,†], Lily Thompson[1,2,†] and Elise Fouquerel[1,*]

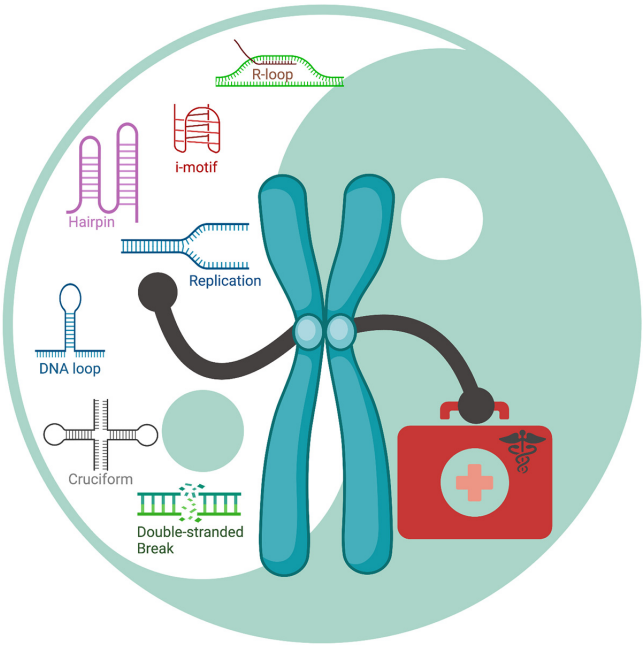
[1]UPMC Hillman Cancer Center, Department of Pharmacology and Chemical Biology, University of Pittsburgh Cancer Institute, Pittsburgh, PA 15232, USA and [2]Department of Biochemistry and Molecular Biology, Thomas Jefferson University, Philadelphia, PA 19107, USA

Received January 27, 2023; Revised March 27, 2023; Editorial Decision April 17, 2023; Accepted April 20, 2023

ABSTRACT

Centromeres play a crucial role in DNA segregation by mediating the cohesion and separation of sister chromatids during cell division. Centromere dysfunction, breakage or compromised centromeric integrity can generate aneuploidies and chromosomal instability, which are cellular features associated with cancer initiation and progression. Maintaining centromere integrity is thus essential for genome stability. However, the centromere itself is prone to DNA breaks, likely due to its intrinsically fragile nature. Centromeres are complex genomic loci that are composed of highly repetitive DNA sequences and secondary structures and require the recruitment and homeostasis of a centromere-associated protein network. The molecular mechanisms engaged to preserve centromere inherent structure and respond to centromeric damage are not fully understood and remain a subject of ongoing research. In this article, we provide a review of the currently known factors that contribute to centromeric dysfunction and the molecular mechanisms that mitigate the impact of centromere damage on genome stability. Finally, we discuss the potential therapeutic strategies that could arise from a deeper understanding of the mechanisms preserving centromere integrity.

GRAPHICAL ABSTRACT



INTRODUCTION: CENTROMERE STRUCTURE AND FUNCTION

The centromere is a specialized region of the chromosome at which the kinetochore assembles. It is essential for faithful chromosomal segregation in both mitotic and meiotic cells (1,2). In humans, centromeric regions contain a highly repetitive DNA sequence composed of 171-bp alpha-satellite monomers that are arranged head-to-tail, forming homogeneous higher order repeat (HOR) arrays that can span several megabases of DNA (3) (Figure 1A). However, DNA sequence is neither necessary nor sufficient for kinetochore assembly and chromosome segregation. Rather, the centromere is defined by both genomic and epigenetic mechanisms. A specialized centromeric histone termed centromeric protein A (CENP-A) replaces the canonical

*To whom correspondence should be addressed. Tel: +1 412 623 3237; Email: Fouquerele3@upmc.edu
†The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

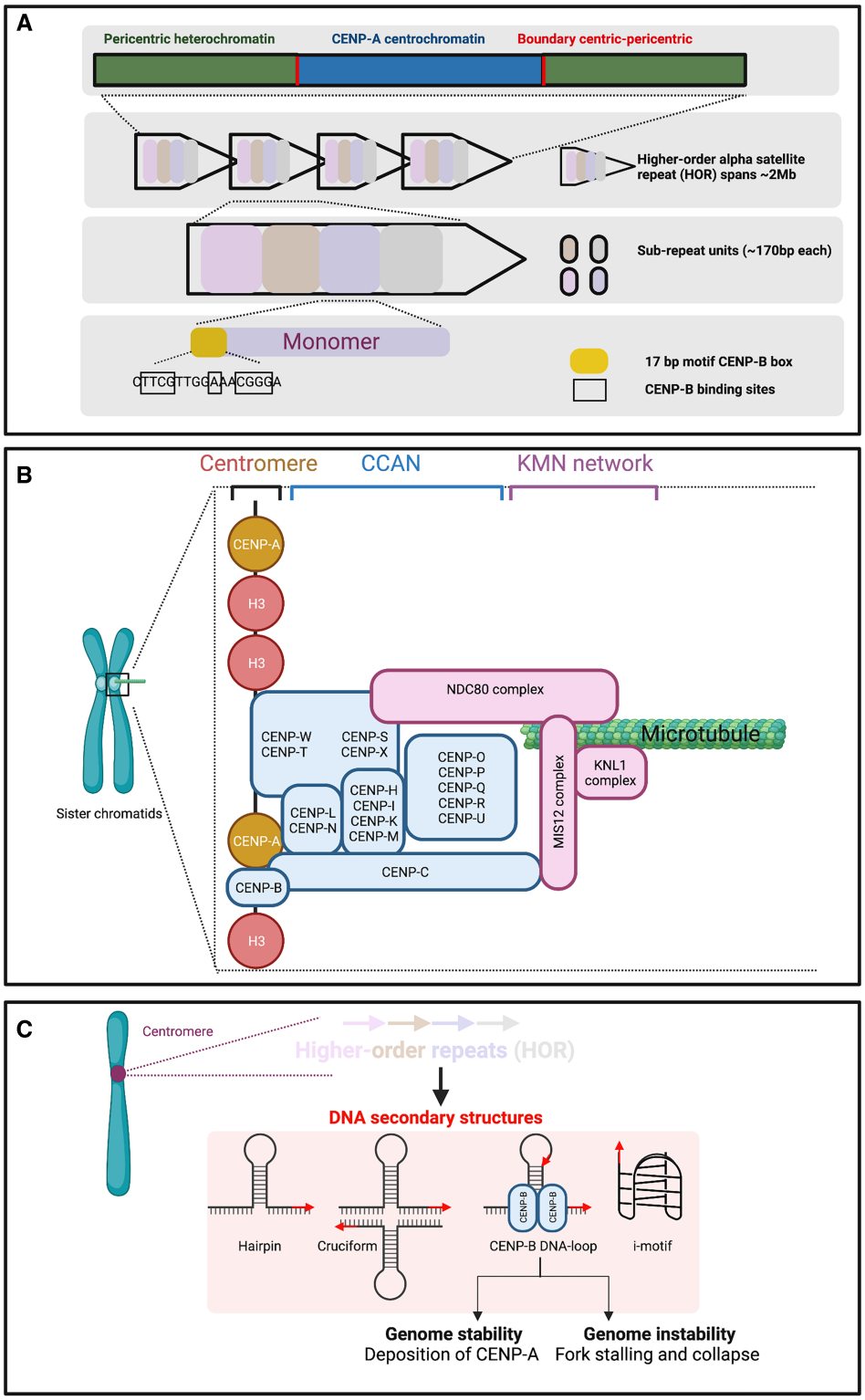


Figure 1. Human centromere structure and kinetochore organization. (**A**) Human centromeres contain arrays of HORs made of ∼171-bp alpha-satellite repeat monomers that are primarily but not completely identical as illustrated by the different color boxes (light pink, brown, purple and gray) and arranged in a head-to-tail manner. Some of these alpha satellites harbor a CENP-B protein binding motif called the CENP-B box (17-bp motif) (yellow box). Peri-centromeric regions (dark green) that flank the centromere region (dark blue) are enriched in heterochromatin and contain alpha satellites in a less ordered fashion. (**B**) The CCAN complex creates a linkage between the centromere and microtubules in the kinetochore. The CCAN comprises five subcomplexes CENP-C, CENP-L/N, CENP-H/I/K/M/L/N, CENP-T/W/S/X and CENP-O/P/Q/R. CENP-B stabilizes the interaction between the chromatin and the kinetochore, via a direct interaction with CENP-C. The CCAN functions as a platform to recruit the KMN network (mis12/NDC80/KNL1) that directly connects the kinetochores. (**C**) Non-B-form DNA secondary structures formed at alpha satellites, such as cruciforms, Z-DNA hairpins or i-motifs. CENP-B also mediates the formation of DNA secondary structures.

histone H3 and characterizes the centromere (1,2). In addition to CENP-A, the centromere is surrounded by the constitutive centromere-associated network (CCAN), which is a subcomplex of the kinetochore that localizes to the centromere throughout the cell cycle and functions as the base linking the centromere to microtubules (Figure 1B). The CENP-A containing centromere core is flanked by inactive pericentromeric domains. Pericentromeric domains contain repeats that are less ordered, which include smaller arrays of diverged alpha-satellite monomers, transposable elements and non-alpha-satellite repeat families (4,5). A fraction of the alpha-satellite monomers contain a 17-bp motif termed a CENP-B box that is recognized and bound by the centromere binding protein B (CENP-B), which is the only known sequence-specific centromeric binding protein and is present at all centromeres except that of the Y chromosome (6). CENP-B is primarily known to stabilize the nucleator of kinetochore assembly CENP-C at the centromere (6). CENP-B can also shape centromere structure by promoting larger DNA loop formation between repeats (7). Upon CENP-B dimerization, two different CENP-B boxes are brought together, leading to compaction of the centromere (Figure 1C). CENP-B-mediated DNA looping and subsequent compaction promotes recruitment of CENP-C and deposition of CENP-A (7). In addition, non-B-form DNA structures form at centromeres, such as cruciforms, Z-DNA hairpins or i-motifs (8–10). The function of these non-B-form DNA structures at centromeres remains debated, yet some speculate a direct role in specifying centromere identity. It is possible that they mediate the deposition of CENP-A at the centromere. For instance, the CENP-A chaperone HJURP, which can recognize Holliday junctions, could bind the four-way junctions of a cruciform (11–13). i-Motifs have been identified to form in the CENP-B box and in alpha satellites (10,14). Although no studies have yet shed light on their function, it is most likely that they play a role in the structural organization of centromeres. These structures are therefore believed to be essential for the maintenance of centromere structure and function.

CENTROMERES ARE SOURCES OF GENOME INSTABILITY AND TUMORIGENESIS

Genome instability refers to a range of alterations spanning from DNA base mutations and breaks to chromosome rearrangements and chromosome instability (CIN), which is itself defined by the gain or loss of chromosomes during each cell division (15). More than 95% of cancers display genome instability and CIN, which results in aneuploidy or polyploidy (16,17). Mathematical models indicate that CIN can initiate tumorigenesis prior to tumor suppressor gene inactivation (18), underlying the importance of understanding the mechanisms leading to CIN. Years of research have led to the conclusion that defects in centromere assembly or maintenance are a cause of the numerical aberrations observed in cancer cells (1,19,20). Indeed, these studies have identified centromeres as the location of breaks in multiple cancer types as evidenced by high percentages of whole-chromosome arm gain, loss or translocation (colorectal, oral and squamous cell carcinomas) (1,19,20). These studies demonstrate that centromeric breakage can directly contribute to CIN and further exacerbate cellular transformation.

Although the impact of centromere dysfunctions on genome stability and tumorigenesis is now well established, the molecular mechanisms driving them are not fully uncovered and remain a topic of investigation. Such studies will reveal important insights into the roles of these genomic loci in cancer progression. Here, we review the current knowledge related to the inherent factors contributing to centromeric dysfunction and the molecular mechanisms mitigating its impact on genome stability. Lastly, we discuss how centromere dysfunction and protection mechanisms can be leveraged for the development of promising anticancer therapeutic strategies.

Secondary DNA structures at the centromere

Secondary structures are sources of centromere fragility during replication. Centromeres are intrinsically fragile, which means that they are prone to breakage especially during replication. Centromere breakage can directly promote chromosomal rearrangements, lagging chromosomes, chromosomal bridges, aneuploidy, formation of micronuclei and chromothripsis, all hallmarks of cancer cells (21–27). This fragility is proposed to be due to their highly repetitive sequences and the presence of secondary DNA structures and loops that can make their replication challenging. However, the contribution of these structures to centromere fragility is not fully understood. Given that they promote DNA compaction and centromeric epigenetic identity (7,11,13), it is most likely that they are crucial to ensure proper cell division but need to be resolved during the replication process. Indeed, preventing DNA looping through CENP-B depletion triggers DNA decondensation and chromosome breakage during mitosis, as evidenced by γH2Ax and 53BP1 recruitment (7). Thus, during mitosis, centromeric loops seem to favor genome stability (Figure 1B). Yet, studies performed in *Xenopus laevis* egg extracts using bacterial artificial chromosomes (BACs) containing human alpha-satellite DNA showed that the secondary structures can slow down replication when compared to BACs harboring repeat-free DNA with similar GC base content (28). Slowing of replication machinery can contribute to fork stalling and collapse and result in replication stress-mediated double-strand break (DSB) formation, thus highlighting the need for the recruitment of specialized enzymes involved in the resolution of these structures.

Mechanisms preventing replication-associated centromere fragility. Centromeres appear to surround themselves with several DNA repair factors during replication (28). For instance, proteomic analysis of centromeric chromatin revealed the presence of the mismatch repair factors MSH2–6, which could promote the resolution of hairpins and other secondary structures potentially formed behind the fork on single-stranded DNA (ssDNA). Another study also found the presence of the nuclease and helicase DNA2. Because loss of DNA2 was found to activate ATR, it was

hypothesized to resolve secondary structures ahead of the fork, preventing RPA recruitment and subsequent ATR response signaling (Figure 2A) (28,29). Interestingly, triggering replication stress using the replication inhibitor aphidicolin does not elicit RPA loading and subsequent TopBP1/ATR signaling activation, contrary to what is observed in other regions of the genome (28,29). This unexpected response is attributed to higher order structures embodied by large double-stranded DNA loops formed during replication. These loops are promoted by topoisomerase-mediated positive supercoiling and stabilized by the condensin complex subunits SMC2–4 (28) (Figure 2A). Because treatment with the topoisomerase inhibitor topotecan restores RPA loading, one role of centromere loop formation may be to prevent activation of ATR signaling behind the replication fork (28,29), thereby facilitating the replication process (Figure 2A). Additional studies will be necessary to further uncover the precise timing and dynamics involved in the formation and resolution of centromeric DNA loops, but their formation upon centromeric DNA replication may be an initial step for the chromatin condensation that occurs during the early stages of mitosis. Because these DNA loops are also mediated by CENP-B and are involved in CENP-A deposition (7), their function expands beyond the G2/M phases of the cell cycle, suggesting that they are a stable component of centromere organization. Thus, it implies that their resolution prior to replication is necessary to not impede DNA synthesis. A rapid response could be ensured by the ADP-ribose transferase PARP1, which is also enriched at centromeres (28,30). PARP1 is a DNA binding protein whose poly(ADP-ribosyl)ation (PARylation) activity is triggered upon the recognition of DNA breaks, stalled replication forks and also several non-canonical DNA structures, including hairpins (31,32). At centromeres, PARP1 activity may be directly elicited by secondary structures and DNA loops as well as by the stalled replication forks. Intriguingly, PARP1 interacts with and PARylates both CENP-A and CENP-B (33,34). During single-strand break (SSB) repair, PARylation of histones allows for their local release and subsequent chromatin relaxation (35). Similarly, the negatively charged PAR could affect the binding of the centromeric proteins to promote a local unlooping of centromeric DNA. A study on the replication of alpha-satellite DNA has revealed that depletion of CENP-B triggers an enrichment of proteins of the pre-replication complex (36). Because slowdown of replication forks by exogenous or endogenous impediments elicits the activation of dormant replication origins, one can speculate that removal of CENP-B from the chromatin may ensure the completion of replication in a timely manner. Finally, MSH2–6 and DNA2 are known PAR binders (37,38). These interactions could therefore orchestrate their recruitment at centromeres. The enrichment of proteins implicated in DNA repair and replication stress pathways during centromeric replication highlights the fragility of these genomic loci. It also demonstrates the extent of the mechanisms that have evolved to mediate this fragility. Gaining a deeper appreciation of the challenges caused by inherent centromere structure and their resolutions will certainly advance our understanding of CIN-mediated genome instability.

Centromeric R-loops

Beneficial and harmful roles of centromeric R-loops. R-loops are nucleic acid structures consisting of a DNA–RNA hybrid and a displaced ssDNA. Centromeric repeats are actively transcribed by RNA polymerase II into non-coding cenRNAs that are part of the centromeric chromatin and participate in kinetochore assembly (39–42). Several studies have shown that cenRNAs associate with centromeric DNA in *cis* and form R-loops at centromeres of yeast, worms and synchronized or unsynchronized mammalian cells (43–46). A recent study also highlights a role for a non-centromeric long non-coding RNA (lnRNA) containing the oncogene *c-Myc* in forming R-loops and promoting CENP-A recruitment at an ectopic locus in colon cancer (47). These studies imply that R-loops, centromeric or not, may have an active role in centromere specification. Like in the rest of the genome, centromeric R-loops have beneficial roles in the maintenance of centromere functions but can also be detrimental depending on the phase of the cell cycle. Their accumulation impairs kinetochore bi-orientation and CENP-A localization, as well as increases chromosomal breaks and micronuclei formation, likely by acting as physical barriers to the progression of replication forks (45,46). Indeed, during replication, R-loop persistence triggers the formation of DSBs, followed by recombination and translocation events at centromeres (48). On the other hand, during early mitosis, R-loops promote the recruitment of the kinase Aurora B and the chromosome passenger complex (43,49). Aurora B orchestrates major steps of mitosis such as spindle assembly checkpoint, sister chromatid cohesion and attachment of microtubules to the kinetochores. Aurora B activation is dependent on the stabilization of R-loops by the ATR kinase and its downstream effector Chk1 (43). These mitotic R-loops accumulate during prophase but need to be resolved during mitotic progression (49). Collectively, perturbation of R-loop homeostasis causes defects in cohesion, kinetochore integrity and overall mitotic fidelity (43,49).

Mechanisms of resolution. Cells have developed several strategies to prevent or resolve R-loops, including factors involved in RNA biogenesis and stability and DNA–RNA helicases as well as RNase H that digests the RNA strand of the RNA–DNA hybrid (50). Studies addressing the functionality of known factors involved in R-loop metabolism within centromeric chromatin specifically are still in their initial stages. A recent study highlighted a role for the DSB repair protein BRCA1, which was previously described to process R-loops at transcription pausing sites (51,52). BRCA1 recruitment at centromeres of undamaged cells in interphase is indeed R-loop-dependent and protects centromeres from R-loop-induced DSBs, possibly by mediating the recruitment of the RNA–DNA helicase senataxin (Figure 2B). Importantly, the absence of BRCA1 unleashes Rad52-dependent recombination between satellite repeats leading to subsequent chromosome missegregation and micronuclei formation (46). The NER endonucleases XPG and XPF also promote the removal of centromeric R-loops (Figure 2B). However, their activity causes the formation of DSBs, which are at the origin of centromere-related aberrations in patients with immunodeficiency–centromeric

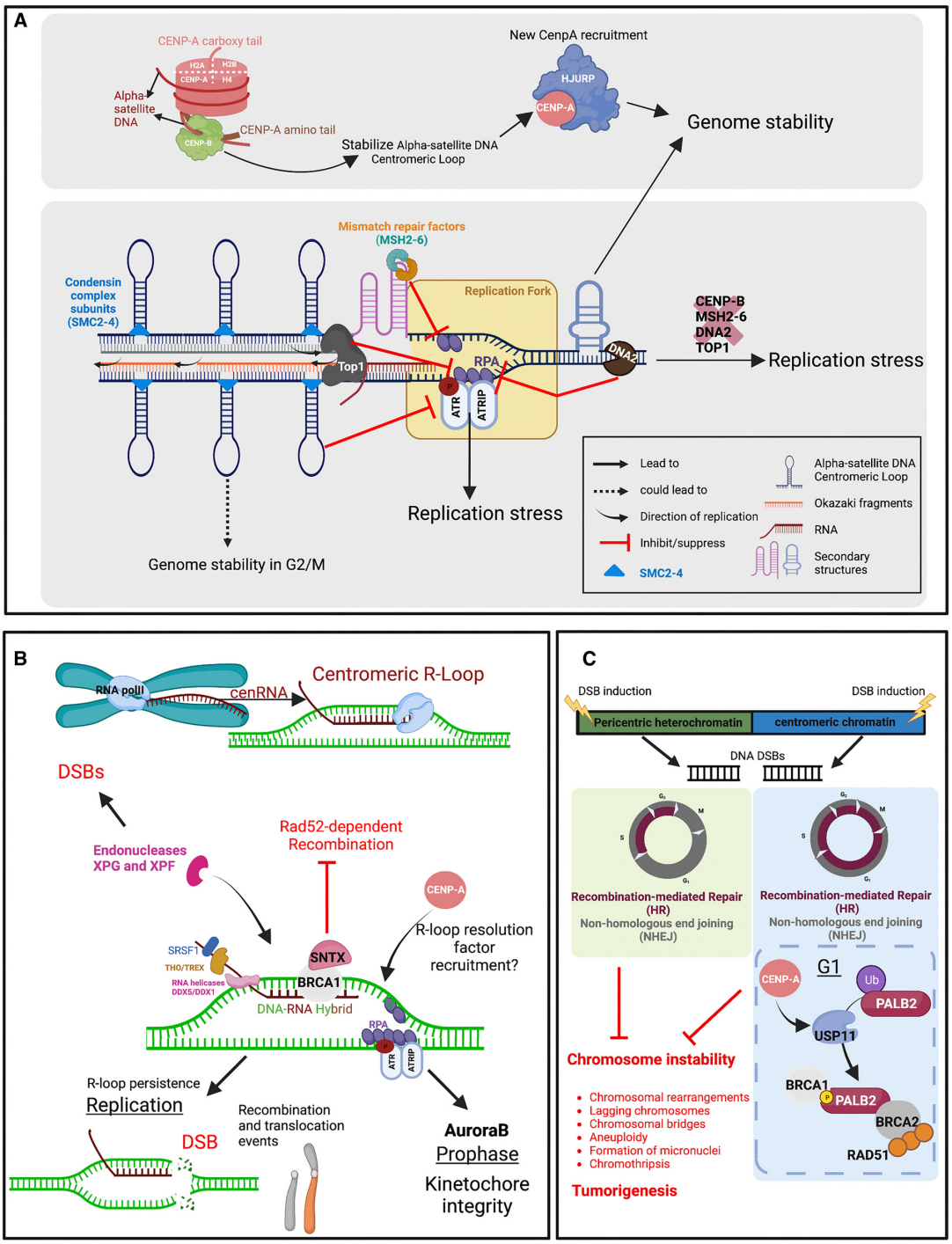


Figure 2. Structural challenges and their resolution. (**A**) Alpha-satellite DNA loops are predicted to be stabilized by CENP-B binding and promote CENP-A deposition by the CENP-A chaperone HJURP. These loops are promoted by topoisomerase and stabilized by the condensin complex subunits SMC2–4. The helicase DNA2 may resolve secondary structures ahead of the fork and prevent RPA recruitment and ATR response signaling, while the mismatch repair factors MSH2–6 may promote the resolution of secondary structures formed behind the fork. Slowdown of replication machinery can contribute to fork stalling and collapse and result in replication stress. (**B**) Non-coding centromeric RNAs (cenRNAs) are transcribed by RNA polymerase II and can form R-loops. R-loop accumulation during replication triggers the formation of DSBs, followed by recombination and translocation events at centromeres. During mitosis, R-loops are stabilized by the ATR kinase and its downstream effector Chk1, promoting the recruitment of the kinase Aurora B and contributing to kinetochore integrity. BRCA1 recruitment prevents R-loop-induced DSBs at centromeres, while the nucleotide excision repair (NER) endonucleases XPG and XPF generate DSBs to trigger removal of centromeric R-loops. Finally, the RNA helicases (DDX5, DDX1), as well as the splicing factor SRSF1 (ASF/SF2), prevent R-loop formation and ensure genome stability. CENP-A is suggested to mediate the recruitment of R-loop resolution factors. (**C**) While centromeric DSBs can recruit both homologous recombination (HR) and non-homologous end joining (NHEJ) repair factors throughout the cell cycle, DSBs at pericentric heterochromatin recruit HR factors only during S/G2. Failure to repair centromeric DSBs can directly promote chromosomal rearrangements, lagging chromosomes, chromosomal bridges, aneuploidy, micronuclei formation and chromothripsis, all hallmarks of cancer cells.

region instability–facial anomalies (ICF) syndrome (53). Another study showed that the deletion of *Hpr1* in budding yeast, a component of the RNA biogenesis and processing factor THO/TREX complex, leads to an accumulation of centromeric R-loops that can be resolved by exogenous expression of RNase H1 (45). Finally, the centromeric protein interactome includes several DEAD-box RNA helicases, some of them shown to unwind RNA–DNA hybrids (DDX5, DDX1), as well as the splicing factor SRSF1 (ASF/SF2) whose recruitment on nascent RNA transcripts prevents R-loop formation and ensures genome stability (28,54).

Outstanding questions. These studies suggest that centromeric R-loops are most likely processed through canonical pathways. Nonetheless, it cannot be excluded that the unique chromatin environment of centromeres necessitates the recruitment of specific factors. Outstanding questions also remain regarding the relationship of these factors with centromeric proteins and the dynamics of these protein interactions during the different phases of the cell cycle. For instance, because the removal of CENP-A during S phase is followed by an accumulation of R-loops in late S phase, CENP-A is suggested to mediate the recruitment of R-loop resolution factors during DNA synthesis to prevent genomic instability during mitosis (48). Future studies analyzing the CENP-A protein interactome during S phase may enable confirmation of this hypothesis. Moreover, we have recently obtained evidence that the DNA repair enzyme PARP1 can associate with R-loops *in vitro* and in cells and that this binding triggers its PARylation activity. Additionally, PARP1 activity promotes the association of the RNA–DNA helicase DDX18 with R-loops (55). Because PARP1 is enriched at centromeres during DNA synthesis (2), it is possible that it also orchestrates R-loop resolution at centromeres during S phase by ensuring the recruitment of resolution factors.

R-loops have recently emerged as important genome stability regulators. The duality of their function at centromeres underscores the importance of the tight regulation of their formation and resolution. This invites questions regarding the existence of mechanisms that readily prevent R-loop accumulation during DNA synthesis while allowing their formation during early mitosis and under conditions of DNA damage induction. Uncovering these mechanisms may offer exciting new research directions in the field.

DNA damage induction and repair at centromeres

Sources of DNA damage and their impact on centromeric DNA. Sources of DNA damage are pervasive and include endogenously and exogenously produced reactive oxygen species, UV and ionizing radiations, or various chemicals. These give rise to DNA base lesions, SSBs and DSBs, the latter being the most deleterious form of DNA damage as their unfaithful repair can lead to chromosomal fusions, translocations or deletions. There is a large gap of knowledge regarding the impact of DNA damaging agents on centromeres. Moreover, the common use of irradiations or oxidizing and alkylating chemicals that broadly impact the genome prevents the attribution of cellular and molecular responses to centromeric DNA damage specifically. This is particularly relevant if the centromeric chromatin environment offers protection to the centromeric DNA. Indeed, centromeres are composed of heterochromatin and this compaction could help protect the DNA from damage. Remarkably, telomeres, which are also composed of compact chromatin (56,57), exhibit 2-fold fewer UV photoproducts than the rest of the genome, which was attributed to a shielding of the DNA by telomeric DNA binding proteins (58). Centromeric proteins could therefore play a similar role in protecting DNA, but whether centromeric DNA is less vulnerable to genotoxic insults has not yet been investigated.

Spontaneous DNA breaks. Spontaneous DNA breaks within the centromere have been mainly attributed to active DNA replication. However, a new study has recently highlighted that they can also be induced *de novo* during quiescence (59). These breaks were initiated by topoisomerase IIβ activity and resolved by RAD51 recombinase. This study provides insights into the centromere paradox that stems from the observation that centromere DNA sequences evolve rapidly and are prone to recombination but can maintain their primary functionality. Moreover, because DSB enrichment was associated with CENP-A occupancy, it also brings evidence that innate centromere fragility may be involved in the epigenetic centromere identity. Therefore, despite the risks of inducing CIN, centromere fragility seems to have beneficial roles at centromeres.

Mechanisms of centromeric DNA DSB repair in the context of chromatin. Although centromeres are intrinsically fragile and represent hotspots for chromosomal breaks, the molecular mechanisms engaged to repair DSBs are not yet fully understood. Like its impact on damage occurrence, the specific centromeric chromatin environment can influence the repair mechanisms. Indeed, while chromatin compaction can help protect DNA from extensive damage, when damage does occur, it could limit the access of repair proteins to the lesion. For example, the repair kinetics of heterochromatin DSBs is slower than that of euchromatin DSBs (60). Although the impact of PARP activities on centromeric proteins and repair rates remains to be tested, one can speculate that the PARylation of CENP-A, CENP-B and BUB3 by PARP1 and PARP2 reported upon induction of DNA strand breaks by γ-irradiation could possibly be one way to promote their release from DNA and subsequent local chromatin relaxation as well as to the recruitment of DNA repair factors (33,34).

Recently, using CRISPR/Cas9 and guide RNA targeting the minor satellite repeats, the Soutoglou laboratory has overcome the issues that arise when using broad DNA damaging agents and shed some light on the centromeric DSB repair mechanisms in the context of chromatin. With this precision tool, they highlighted differences in the mode of repair between DSBs arising within the centromeric and the pericentric heterochromatin. While centromeric DSBs can recruit both HR and NHEJ repair factors throughout the cell cycle and even in G1 phase, DSBs at pericentric heterochromatin recruit HR factors only during S/G2

phases (61) (Figure 2C). Additionally, the presence of the active chromatin marks H3K4me2 at centromeres promotes transcription upon DSB induction and the formation of R-loops, which facilitates DNA end resection (62). Strikingly, the centromere histone variant CENP-A enables the activation of HR in G1 by mediating the recruitment of the deubiquitinase USP11, which subsequently enables the formation of the RAD51–BRCA1–BRCA2 HR complex by deubiquitinating PALB2 (62) (Figure 2C). Interestingly, while NHEJ repair during G1 is positionally stable, repair by HR requires end resection and DSB relocation to the periphery of heterochromatin where RAD51 can operate. Failure to relocate engages repair of the breaks through RAD52-dependent single-strand annealing, which uses homologous repeats to bridge DSB ends but causes deletions and rearrangements between the repeats (61,62). Along with the observation by the Esashi group that RAD51 depletion leads to a loss of CENP-A in both quiescent and cycling cells (59), this work demonstrates that recombination events can be beneficial for centromere stability. It also brings evidence that centromeric heterochromatin is not refractory to repair when it comes to DSB induction. Finally, it highlights that the chromatin environment, which differs between pericentromeres and centromeres, can dictate the repair pathway choice. Whether these observations can be made for other types of DNA damage remains to be tested. Future research efforts on the use and development of tools that can target DNA damage to the centromeres specifically will prove crucial to answering these outstanding questions.

Centromeres are promising therapeutic targets against tumorigenesis

Current anticancer drugs targeting centromere function are mostly antimitotic compounds that destabilize or stabilize microtubules to prevent cell proliferation and promote cell death. However, these drugs are also associated with cytotoxicity of non-cancer cells and a possibility for multiple resistance mechanisms to arise (63), further underscoring the need for new strategies. The recent knowledge acquired on the roles of R-loops and centromeric and DNA repair proteins in preserving centromeric DNA structure, as well as the impact of the centromeric chromatin environment on DNA repair pathway choice, offers potentially promising pathways to target.

Targeting centromeric R-loops. A growing interest in R-loops as targets in cancer therapy has recently emerged, owing to several studies demonstrating their dual role in genome stability. Dysregulated transcription programs in cancer cells have indeed been correlated to an increase in R-loops and subsequent replication stress. Whether centromeric R-loops are particularly more abundant in some cancer cells remains to be established. However, the role of centromeric R-loops in ensuring proper mitotic division is a unique aspect of R-loop biology that could be exploited in cancer therapy to counteract cell proliferation. While strategies that directly perturb R-loop formation and/or resolution may be useful to increase the replication stress burden to deadly levels in cancer cells, targeting centromeric R-loops specifically could therefore confer an additional advantage. Existing drugs such as inhibitors targeting the *de novo* DNA methyltransferase 3b (DNMT3b), which was shown to protect centromeres against the deleterious effects of R-loops (53), could be exploited in this context. DNMT3b is a DNA methyltransferase that is constitutively present at centromeres to modulate methylation in this region (64). DNMT3b loss-of-function mutation is specifically found in patients suffering from the ICF syndrome, whose cells harbor a high level of centromeric R-loops and DNA breaks. Accordingly, DNMT3b deletion in human carcinoma cells HCT116 leads to reduced level of centromeric R-loop and subsequent DSBs at pericentromeric regions (53). Because DNMT3b is also the most commonly overexpressed DNA methyltransferase in cancer cells, its inhibition may primarily affect diseased cells and is therefore a promising therapeutic avenue to investigate. Alternatively, slowing down cancer cell proliferation by blocking centromeric R-loop formation during mitosis could be another interesting approach to examine. This could be achieved by preventing cenRNA stabilization by inhibiting ATR or by using antisense oligonucleotides (ASOs) that have already been successfully used to target lnRNAs in some human diseases (65). Finally, the recent data describing a role for a non-centromeric lnRNA in promoting CENP-A deposition at an ectopic locus in cancer cells further highlight the importance of focusing on the development of sequence-specific ASOs or small-molecule inhibitors disrupting this interaction (47).

Targeting centromeric proteins. Targeting centromeric proteins is also a promising avenue from which novel treatments may emerge. In 2016, a group developed the centromere and kinetochore gene expression score (CES) and demonstrated that high CES often correlates with increased levels of genomic instability and poor prognosis for several types of cancers (66). CENP-A, in particular, is overexpressed in ∼20 different cancer types in which ectopic deposition is increased and contributes to CIN (67). Additionally, both HJURP and CENP-A are upregulated in p53-null human tumors (68). Thus, elevated centromeric protein expression is proposed to be a biomarker for disease progression and patient outcome. Targeting levels of highly expressed centromeric proteins could therefore be a tactic that could confer a therapeutic advantage. Interestingly, high CES also correlates with improved response to radiotherapy, cisplatin and topoisomerase inhibitors, due to reduced tolerance of these high CES cells to additional genotoxic stress (66). Thus, determining the centromeric protein expression profile in patients could help in deciding whether traditional treatments should be encouraged to improve personalized treatments. The manipulation of centromeric proteins, especially those involved in shaping the centromeric chromatin structures, could also offer a way to modulate the sensitivity of DNA to DNA damaging agents. Chromatin remodeling influences various cell functions and, when dysfunctional, can promote tumorigenesis. Whether centromeric chromatin undergoes drastic remodeling in cancer is not reported. Yet, recent observation on the importance of chromatin state in the choice of DNA repair pathway offers the possibility

of exploring the impact of chromatin remodeling-related drugs in combination with DNA damaging agents.

CONCLUDING REMARKS

The molecular mechanisms of centromeric instability represent a relatively recent area of research that has gained momentum in the past decade. The dual nature of centromeres as essential mediators of cell division but also fragile regions prone to breaks and fork stalling is an intriguing dichotomy that highlights the importance of protective mechanisms at the centromere. Because centromeres are vital for faithful transmission of the genome, compromised centromeric integrity can be especially deleterious as evidenced by the prevalence of cancers presenting with aneuploidies that originate from breaks at centromeric regions. Thus, understanding and further characterizing the molecular mechanisms of genomic instability at centromere loci is imperative to informing and identifying novel therapeutic approaches that can curb cancer occurrence and improve disease outcome.

DATA AVAILABILITY

All data generated and presented in this review are available from the corresponding author upon request.

ACKNOWLEDGMENTS

We thank Natalie Laspata for helpful edits on the manuscript.

FUNDING

National Institutes of Health [R35GM142982]; Start-up fundings from UPMC Hillman Cancer Center (to E.F.).
Conflict of interest statement. None declared.

REFERENCES

1. Barra,V. and Fachinetti,D. (2018) The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nat. Commun.*, **9**, 4340.
2. McKinley,K.L. and Cheeseman,I.M. (2016) The molecular basis for centromere identity and function. *Nat. Rev. Mol. Cell Biol.*, **17**, 16–29.
3. Rudd,M.K., Schueler,M.G. and Willard,H.F. (2003) Sequence organization and functional annotation of human centromeres. *Cold Spring Harbor Symp. Quant. Biol.*, **68**, 141–149.
4. Altemose,N., Logsdon,G.A., Bzikadze,A.V., Sidhwani,P., Langley,S.A., Caldas,G.V., Hoyt,S.J., Uralsky,L., Ryabov,F.D., Shew,C.J. *et al.* (2022) Complete genomic and epigenetic maps of human centromeres. *Science*, **376**, eabl4178.
5. Smurova,K. and De Wulf,P. (2018) Centromere and pericentromere transcription: roles and regulation … in sickness and in health. *Front. Genet.*, **9**, 674.
6. Fachinetti,D., Han,J.S., McMahon,M.A., Ly,P., Abdullah,A., Wong,A.J. and Cleveland,D.W. (2015) DNA sequence-specific binding of CENP-B enhances the fidelity of human centromere function. *Dev. Cell*, **33**, 314–327.
7. Chardon,F., Japaridze,A., Witt,H., Velikovsky,L., Chakraborty,C., Wilhelm,T., Dumont,M., Yang,W., Kikuti,C., Gangnard,S. *et al.* (2022) CENP-B-mediated DNA loops regulate activity and stability of human centromeres. *Mol. Cell*, **82**, 1751–1767.
8. Jonstrup,A.T., Thomsen,T., Wang,Y., Knudsen,B.R., Koch,J. and Andersen,A.H. (2008) Hairpin structures formed by alpha satellite DNA of human centromeres are cleaved by human topoisomerase IIalpha. *Nucleic Acids Res.*, **36**, 6165–6174.
9. Nonin-Lecomte,S. and Leroy,J.L. (2001) Structure of a C-rich strand fragment of the human centromeric satellite III: a pH-dependent intercalation topology. *J. Mol. Biol.*, **309**, 491–506.
10. Garavis,M., Mendez-Lago,M., Gabelica,V., Whitehead,S.L., Gonzalez,C. and Villasante,A. (2015) The structure of an endogenous *Drosophila* centromere reveals the prevalence of tandemly repeated sequences able to form i-motifs. *Sci. Rep.*, **5**, 13307.
11. Kasinathan,S. and Henikoff,S. (2018) Non-B-form DNA is enriched at centromeres. *Mol. Biol. Evol.*, **35**, 949–962.
12. Talbert,P. and Henikoff,S. (2022) Centromeres organize (epi)genome architecture. *Cell*, **185**, 3083–3085.
13. Talbert,P.B. and Henikoff,S. (2022) The genetics and epigenetics of satellite centromeres. *Genome Res.*, **32**, 608–615.
14. Garavis,M., Escaja,N., Gabelica,V., Villasante,A. and Gonzalez,C. (2015) Centromeric alpha-satellite DNA adopts dimeric i-motif structures capped by AT Hoogsteen base pairs. *Chemistry*, **21**, 9816–9824.
15. Aguilera,A. and Gomez-Gonzalez,B. (2008) Genome instability: a mechanistic view of its causes and consequences. *Nat. Rev. Genet.*, **9**, 204–217.
16. Mitelman,F., Johansson,B. and Mertens,F. (2007) The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer*, **7**, 233–245.
17. Cimini,D. (2008) Merotelic kinetochore orientation, aneuploidy, and cancer. *Biochim. Biophys. Acta*, **1786**, 32–40.
18. Nowak,M.A., Komarova,N.L., Sengupta,A., Jallepalli,P.V., Shih,I.-M., Vogelstein,B. and Lengauer,C. (2002) The role of chromosomal instability in tumor initiation. *Proc. Natl Acad. Sci. U.S.A.*, **99**, 16226–16231.
19. Knutsen,T., Padilla-Nash,H.M., Wangsa,D., Barenboim-Stapleton,L., Camps,J., McNeil,N., Difilippantonio,M.J. and Ried,T. (2010) Definitive molecular cytogenetic characterization of 15 colorectal cancer cell lines. *Genes Chromosomes Cancer*, **49**, 204–223.
20. Mackinnon,R.N. and Campbell,L.J. (2011) The role of dicentric chromosome formation and secondary centromere deletion in the evolution of myeloid malignancy. *Genet. Res. Int.*, **2011**, 643628.
21. Mellone,B.G. and Fachinetti,D. (2021) Diverse mechanisms of centromere specification. *Curr. Biol.*, **31**, R1491–R1504.
22. Bakhoum,S.F. and Cantley,L.C. (2018) The multifaceted role of chromosomal instability in cancer and its microenvironment. *Cell*, **174**, 1347–1360.
23. Black,E.M. and Giunta,S. (2018) Repetitive fragile sites: centromere satellite DNA as a source of genome instability in human diseases. *Genes*, **9**, 615.
24. Balzano,E. and Giunta,S. (2020) Centromeres under pressure: evolutionary innovation in conflict with conserved function. *Genes*, **11**, 912.
25. Ly,P. and Cleveland,D.W. (2017) Rebuilding chromosomes after catastrophe: emerging mechanisms of chromothripsis. *Trends Cell Biol.*, **27**, 917–930.
26. Romeo,F., Falbo,L. and Costanzo,V. (2016) Replication, checkpoint suppression and structure of centromeric DNA. *Nucleus*, **7**, 540–546.
27. Ly,P., Brunner,S.F., Shoshani,O., Kim,D.H., Lan,W., Pyntikova,T., Flanagan,A.M., Behjati,S., Page,D.C., Campbell,P.J. *et al.* (2019) Chromosome segregation errors generate a diverse spectrum of simple and complex genomic rearrangements. *Nat. Genet.*, **51**, 705–715.
28. Aze,A., Sannino,V., Soffientini,P., Bachi,A. and Costanzo,V. (2016) Centromeric DNA replication reconstitution reveals DNA loops and ATR checkpoint suppression. *Nat. Cell Biol.*, **18**, 684–691.
29. Li,Z., Liu,B., Jin,W., Wu,X., Zhou,M., Liu,V.Z., Goel,A., Shen,Z., Zheng,L. and Shen,B. (2018) hDNA2 nuclease/helicase promotes centromeric DNA replication and genome stability. *EMBO J.*, **37**, e96729.
30. Earle,E. (2000) Poly(ADP-ribose) polymerase at active centromeres and neocentromeres at metaphase. *Hum. Mol. Genet.*, **9**, 187–194.
31. Bryant,H.E., Petermann,E., Schultz,N., Jemth,A.S., Loseva,O., Issaeva,N., Johansson,F., Fernandez,S., McGlynn,P. and Helleday,T. (2009) PARP is activated at stalled forks to mediate Mre11-dependent replication restart and recombination. *EMBO J.*, **28**, 2601–2615.
32. Lonskaya,I., Potaman,V.N., Shlyakhtenko,L.S., Oussatcheva,E.A., Lyubchenko,Y.L. and Soldatenkov,V.A. (2005) Regulation of poly(ADP-ribose) polymerase-1 by DNA structure-specific binding. *J. Biol. Chem.*, **280**, 17076–17083.

33. Saxena,A. (2002) Poly(ADP-ribose) polymerase 2 localizes to mammalian active centromeres and interacts with PARP-1, Cenpa, Cenpb and Bub3, but not Cenpc. *Hum. Mol. Genet.*, **11**, 2319–2329.
34. Saxena,A., Saffery,R., Wong,L.H., Kalitsis,P. and Choo,K.H.A. (2002) Centromere proteins Cenpa, Cenpb, and Bub3 interact with poly(ADP-ribose) polymerase-1 protein and are poly(ADP-ribosyl)ated. *J. Biol. Chem.*, **277**, 26921–26926.
35. Schreiber,V., Dantzer,F., Ame,J.-C. and de Murcia,G. (2006) Poly(ADP-ribose): novel functions for an old molecule. *Nat. Rev. Mol. Cell Biol.*, **7**, 517–528.
36. Erliandri,I., Fu,H., Nakano,M., Kim,J.H., Miga,K.H., Liskovykh,M., Earnshaw,W.C., Masumoto,H., Kouprina,N., Aladjem,M.I. *et al.* (2014) Replication of alpha-satellite DNA arrays in endogenous human centromeric regions and in human artificial chromosome. *Nucleic Acids Res.*, **42**, 11502–11516.
37. Dasovich,M., Beckett,M.Q., Bailey,S., Ong,S.E., Greenberg,M.M. and Leung,A.K.L. (2021) Identifying poly(ADP-ribose)-binding proteins with photoaffinity-based proteomics. *J. Am. Chem. Soc.*, **143**, 3037–3042.
38. Pleschke,J.M., Kleczkowska,H.E., Strohm,M. and Althaus,F.R. (2000) Poly(ADP-ribose) binds to specific domains in DNA damage checkpoint proteins. *J. Biol. Chem.*, **275**, 40974–40980.
39. Talbert,P.B., Kasinathan,S. and Henikoff,S. (2018) Simple and complex centromeric satellites in *Drosophila* sibling species. *Genetics*, **208**, 977–990.
40. Ferri,F., Bouzinba-Segard,H., Velasco,G., Hube,F. and Francastel,C. (2009) Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Res.*, **37**, 5071–5080.
41. Rosic,S., Kohler,F. and Erhardt,S. (2014) Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J. Cell Biol.*, **207**, 335–349.
42. Quenet,D. and Dalal,Y. (2014) A long non-coding RNA is required for targeting centromeric protein A to the human centromere. *eLife*, **3**, e03254.
43. Kabeche,L., Nguyen,H.D., Buisson,R. and Zou,L. (2018) A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science*, **359**, 108–114.
44. Castellano-Pozo,M., Garcia-Muse,T. and Aguilera,A. (2012) The *Caenorhabditis elegans* THO complex is required for the mitotic cell cycle and development. *PLoS One*, **7**, e52447.
45. Mishra,P.K., Chakraborty,A., Yeh,E., Feng,W., Bloom,K.S. and Basrai,M.A. (2021) R-loops at centromeric chromatin contribute to defects in kinetochore integrity and chromosomal instability in budding yeast. *Mol. Biol. Cell*, **32**, 74–89.
46. Racca,C., Britton,S., Hedouin,S., Francastel,C., Calsou,P. and Larminat,F. (2021) BRCA1 prevents R-loop-associated centromeric instability. *Cell Death Dis.*, **12**, 896.
47. Arunkumar,G., Baek,S., Sturgill,D., Bui,M. and Dalal,Y. (2022) Oncogenic lncRNAs alter epigenetic memory at a fragile chromosomal site in human cancer cells. *Sci. Adv.*, **8**, eabl5621.
48. Giunta,S., Herve,S., White,R.R., Wilhelm,T., Dumont,M., Scelfo,A., Gamba,R., Wong,C.K., Rancati,G., Smogorzewska,A. *et al.* (2021) CENP-A chromatin prevents replication stress at centromeres to avoid structural aneuploidy. *Proc. Natl Acad. Sci. U.S.A.*, **118**, e2015634118.
49. Moran,E.C., Liu,L., Zasadzinska,E., Kestner,C.A., Sarkeshik,A., DeHoyos,H., Yates,J.R., Foltz,D. and Stukenberg,P.T. (2021) Mitotic R-loops direct Aurora B kinase to maintain centromeric cohesion. bioRxiv doi: https://doi.org/10.1101/2021.01.14.426738, 15 January 2021, preprint: not peer reviewed.
50. Santos-Pereira,J.M. and Aguilera,A. (2015) R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.*, **16**, 583–597.
51. Hatchi,E., Skourti-Stathaki,K., Ventz,S., Pinello,L., Yen,A., Kamieniarz-Gdula,K., Dimitrov,S., Pathania,S., McKinney,K.M., Eaton,M.L. *et al.* (2015) BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol. Cell*, **57**, 636–647.
52. Zhang,X., Chiang,H.C., Wang,Y., Zhang,C., Smith,S., Zhao,X., Nair,S.J., Michalek,J., Jatoi,I., Lautner,M. *et al.* (2017) Attenuation of RNA polymerase II pausing mitigates BRCA1-associated R-loop accumulation and tumorigenesis. *Nat. Commun.*, **8**, 15908.
53. Shih,H.T., Chen,W.Y., Wang,H.Y., Chao,T., Huang,H.D., Chou,C.H. and Chang,Z.F. (2022) DNMT3b protects centromere integrity by restricting R-loop-mediated DNA damage. *Cell Death Dis.*, **13**, 546.
54. Li,X. and Manley,J.L. (2005) Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell*, **122**, 365–378.
55. Lin,W.L., Chen,J.K., Wen,X., He,W., Zarceno,G.A., Chen,Y., Chen,S., Paull,T.T. and Liu,H.W. (2022) DDX18 prevents R-loop-induced DNA damage and genome instability via PARP-1. *Cell Rep.*, **40**, 111089.
56. Nikitina,T. and Woodcock,C.L. (2004) Closed chromatin loops at the ends of chromosomes. *J. Cell Biol.*, **166**, 161–165.
57. Soman,A., Korolev,N. and Nordenskiold,L. (2022) Telomeric chromatin structure. *Curr. Opin. Struct. Biol.*, **77**, 102492.
58. Parikh,D., Fouquerel,E., Murphy,C.T., Wang,H. and Opresko,P.L. (2015) Telomeres are partly shielded from ultraviolet-induced damage and proficient for nucleotide excision repair of photoproducts. *Nat. Commun.*, **6**, 8214.
59. Saayman,X., Graham,E., Nathan,W.J., Nussenzweig,A. and Esashi,F. (2023) Centromeres as universal hotspots of DNA breakage, driving RAD51-mediated recombination during quiescence. *Mol. Cell*, **83**, 523–538.
60. Goodarzi,A.A., Noon,A.T. and Jeggo,P.A. (2009) The impact of heterochromatin on DSB repair. *Biochem. Soc. Trans.*, **37**, 569–576.
61. Tsouroula,K., Furst,A., Rogier,M., Heyer,V., Maglott-Roth,A., Ferrand,A., Reina-San-Martin,B. and Soutoglou,E. (2016) Temporal and spatial uncoupling of DNA double strand break repair pathways within mammalian heterochromatin. *Mol. Cell*, **63**, 293–305.
62. Yilmaz,D., Furst,A., Meaburn,K., Lezaja,A., Wen,Y., Altmeyer,M., Reina-San-Martin,B. and Soutoglou,E. (2021) Activation of homologous recombination in G1 preserves centromeric integrity. *Nature*, **600**, 748–753.
63. van Vuuren,R.J., Visagie,M.H., Theron,A.E. and Joubert,A.M. (2015) Antimitotic drugs in the treatment of cancer. *Cancer Chemother. Pharmacol.*, **76**, 1101–1112.
64. Gopalakrishnan,S., Sullivan,B.A., Trazzi,S., Della Valle,G. and Robertson,K.D. (2009) DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions. *Hum. Mol. Genet.*, **18**, 3178–3193.
65. Quemener,A.M., Bachelot,L., Forestier,A., Donnou-Fournet,E., Gilot,D. and Galibert,M.D. (2020) The powerful world of antisense oligonucleotides: from bench to bedside. *Wiley Interdiscip. Rev. RNA*, **11**, e1594.
66. Zhang,W., Mao,J.H., Zhu,W., Jain,A.K., Liu,K., Brown,J.B. and Karpen,G.H. (2016) Centromere and kinetochore gene misexpression predicts cancer patient survival and response to radiotherapy and chemotherapy. *Nat. Commun.*, **7**, 12619.
67. Mahlke,M.A. and Nechemia-Arbely,Y. (2020) Guarding the genome: CENP-A-chromatin in health and cancer. *Genes*, **11**, 810.
68. Filipescu,D., Naughtin,M., Podsypanina,K., Lejour,V., Wilson,L., Gurard-Levin,Z.A., Orsi,G.A., Simeonova,I., Toufektchan,E., Attardi,L.D. *et al.* (2017) Essential role for centromeric factors following p53 loss and oncogenic transformation. *Genes Dev.*, **31**, 463–480.