

Review

Centromeric Transcription: A Conserved Swiss-Army Knife

Ganesan Arunkumar  and Daniël P. Melters * 

Chromatin Structure and Epigenetic Mechanisms, Laboratory of Receptor Biology and Gene Expression, Center for Cancer Research, NCI, NIH, Bethesda, MD 20892, USA; arunkumar.ganesan@nih.gov

* Correspondence: daniel.melters@nih.gov

Received: 22 July 2020; Accepted: 7 August 2020; Published: 9 August 2020



Abstract: In most species, the centromere is comprised of repetitive DNA sequences, which rapidly evolve. Paradoxically, centromeres fulfill an essential function during mitosis, as they are the chromosomal sites wherein, through the kinetochore, the mitotic spindles bind. It is now generally accepted that centromeres are transcribed, and that such transcription is associated with a broad range of functions. More than a decade of work on this topic has shown that centromeric transcripts are found across the eukaryotic tree and associate with heterochromatin formation, chromatin structure, kinetochore structure, centromeric protein loading, and inner centromere signaling. In this review, we discuss the conservation of small and long non-coding centromeric RNAs, their associations with various centromeric functions, and their potential roles in disease.

Keywords: centromere; transcription; non-coding RNAs; nucleosomes; chromatin; epigenetics

1. Introduction

Centromeres are essential for ensuring accurate chromosome segregation in eukaryotes. Centromeric chromatin is characterized by the enrichment of nucleosomes containing the centromere-specific histone H3 variant CENP-A/CENH3, which in turn, directly and indirectly, recruit kinetochore components. Despite their essential function, these centromeric and kinetochore proteins are fast-evolving [1–3]. This rapid evolution is also a characteristic of centromeric DNAs [4,5], which are highly repetitive in nature. This discrepancy between the essential function of centromeres and the rapid evolution of its components is known as the centromere paradox [5].

Despite the variation in centromere structure and sequence, a common theme has recently emerged. Whether a point centromere or a regional centromere, centromeric transcription is critical for centromere function. Interestingly, centromeric transcripts (cenRNAs) are found in two forms: long non-coding RNAs (lncRNAs) (>200 nt) and small RNAs (<200 nt). Indeed, the rapid evolution of centromeric DNA is mirrored by the rapid evolution of lncRNAs [6,7]. In general, lncRNAs can be processed based on their nuclear localization [8,9], but they can also be processed into siRNAs by dicer [9,10]. As the functions of more lncRNAs are being dissected, their roles in modulating mRNA cleavage, translational repression, and regulation of alternative splicing, have become apparent [8,9,11]. Although many lncRNAs are fast evolving, several lncRNAs are conserved at both the sequence and synteny (physical co-localization on genetic locus) levels [6,7], suggesting that some lncRNAs might be conserved at the functional level across species. Surprisingly, a recent study [12] showed that a conserved lncRNA expressed in human and mouse embryonic stem cells displayed surprisingly divergent subcellular localization and functions. The human version of the lncRNA was spliced more often, resulting in preferential cytoplasmic localization, whereas the mouse version was exclusively nuclear [12].

Although centromeres are transcribed, it remains unclear what features of these transcripts drive their associated functions. Centromeric DNA sequences evolve rapidly and only show limited

conservation, even among closely related species [4]. In addition, centromeres are known to reposition on chromosomes [13]. This creates a scenario wherein centromeric lncRNAs are conserved based on their function rather than on their sequence and synteny. In this review, we highlight recent advances in our understanding of centromeric transcription, how common centromeric transcription is, how transcripts may impact centromere biology, and how aberrant centromeric transcription contributes to disease.

2. Centromeric Transcription

RNA polymerase 2 (RNAP2) is capable of transcribing most DNA sequences without strong sequence-specificity [14,15]. Consequently, as long as transcriptional initiation and elongation occur, RNAP2 will transcribe most underlying DNA [14,16–18]. Sequence-specific factors, such as TBP, help direct where the pre-initiation complex is assembled and thus where transcription occurs [16,18]. In contrast, for highly repetitive sequences, such as centromeric DNA, no strong promoter activity has been reported [19,20]. Nevertheless, centromeric transcripts have been observed across a broad range of species (Table 1, Figure 1). The transcripts identified so far fall into two main categories: long non-coding RNAs (>200 nt) and small RNAs (<200 nt).

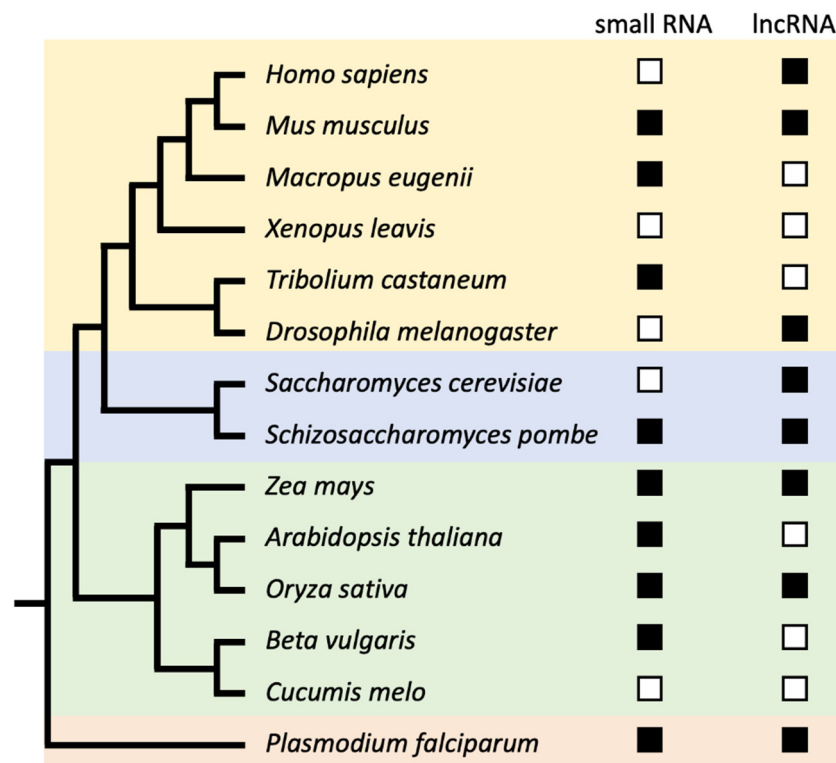


Figure 1. Cladogram showing which types of transcripts have been found in which species. The presence of centromeric small RNAs (<200 nt) and centromeric lncRNA (>200 nt) is shown behind each species described in Table 1. The colors yellow, blue, green, and red represent the Animalia, Fungi, Plantae, and Chromista kingdoms, respectively.

Table 1. Summary of centromeric transcription as reported in various eukaryotic species, including the name of the transcribed sequence, where in the cell cycle transcription occurs, the reported lengths of the transcriptional products, any proteins they are known to interact with, whether these transcripts are 5' capped (5'), polyadenylated (pA), and/or spliced (S), and whether these sequences have been described to act cis or trans.

Species	Sequence Name	Cell Cycle	Transcript Length	Interacting Proteins	5'	pA	S	cis/trans	Ref.
Humans (<i>Homo sapiens</i>)	α -satellite *, L1 *	early G1, mitosis	0.5–2.45 kb	CENP-A, CENP-B, CENP-C, HJURP, SGO1, Aurora B, DHX38, SUV39H1		+		cis	[21–33]
Mouse (<i>Mus musculus</i>)	minor satellite *, major satellite #	S, G2/M	0.1–5 kb, 120 nt	CENP-A, Aurora B, Survivin, INCENP, WDHD1, Dicer, SUV39H2		+			[34–44]
Tammar wallaby (<i>Macropus eugenii</i>)	KERV-1 #, sat23 #		34–42 nt	CENP-B					[45,46]
<i>Xenopus laevis</i> , <i>Xenopus tropicalis</i>	cen-RNA *, fcr1 *			CENP-C, Aurora B, INCENP, Borealin			+	trans	[47–49]
Red flour beetle (<i>Tribolium castaneum</i>)	TCAST *, #		21–26 nt				+		[50]
Fruit fly (<i>Drosophila melanogaster</i>)	satellite III *	mitosis	~1.3 kb	CENP-C				trans	[51–53]
Budding yeast (<i>Saccharomyces cerevisiae</i>)	cenRNA *	S phase	462–1754 nt			+		trans	[54–56]
Fission yeast (<i>Schizosaccharomyces pombe</i>)	Otr #, imr1 #, cnt *	S phase	0.5–10 kb	Rdp1, Ago1, Chp1, Clr4, Swi6	+	+	+	cis	[57–68]
Corn (<i>Zea mays</i>)	CentC *, CRM *		40–900 nt	CENP-A, CENP-C			+		[69–72]
<i>Arabidopsis thaliana</i>	cen180 *		24 nt						[73]
Rice (<i>Oryza sativa</i>)	CentO *, CRR *		4–15 kb; 21–25 nt			+	+		[74,75]
Sugar beet (<i>β vulgaris</i>)	pBV *, pEV satellite *		24 nt						[76]
Melon (<i>Cucumis melo</i>)	CmSat162 *, CmSat189 *								[77]
<i>Plasmodium falciparum</i>	Cen2 *, Cen3 *		75 nt, 175 nt						[78]

* centromere core region and # pericentromeric region. The yellow, blue, green, and red colors represent the Animalia, Fungi, Plantae, and Chromista kingdoms, respectively.

2.1. Centromeric Long Non-Coding Transcripts

The first reported centromeric transcripts came from mouse satellite DNA in the late 1960s [36,37], but it took several more decades before the functional implication of centromeric transcription was appreciated. Today, centromeric transcripts have been observed in Apicomplexa, plants, fungi, and animals (Table 1, Figure 1). Their sequences and transcript lengths remain to be determined for most species. A fundamental problem is that most centromeres do not appear to have well-defined lncRNA genes with distinctive genic features such as promoters, splice sites, and polyadenylation signals. Consequently, with short-read sequencing technologies, although they effective at assessing the transcription and splicing of annotated genes, it is technically difficult to assemble de novo transcripts from centromeric RNAs [79,80]. An alternative method to assess the length and abundance of centromeric transcripts is to perform a Northern blot from isolated total RNAs. Indeed, such efforts have been made in various species with varying results (Figure 1). In human cell lines, one study used a consensus α -satellite probe and found that human centromeric transcripts are \approx 1.3 kb in length [21]. However, another study used probes specific to centromeric D17Z1, D17Z1-B, and DXZ1 arrays, and observed a smear ranging from 0.3 to 2 kb [22]. In mouse cell lines different sizes of satellite RNA (0.12–4 kb) were found. Curiously, the 120-nt minor satellite transcripts were sensitive to different growing conditions [35]. This shows that centromeric transcription can be regulated by external stimuli. Uncovering the functional consequences of altering the transcriptional output of centromeres is important for understanding how centromeric transcription is related to disease.

In fission yeast, the central centromere core *Cnt1* produces transcripts of \approx 0.5 kb in length [57]. Interestingly, different RNA-processing and kinetochore mutants changed the abundance of this transcript [57]. Many plants have retrotransposons at their centromeres, which are expressed (Table 1). In rice, in addition to the smear of 4–15 kb, a distinct band at \approx 3.1 kb was observed for various centromeric retrotransposons [74]. In maize, the centromere-specific retrotransposon CRM-derived transcripts are 40–900 nt in length [70]. Whereas most species have a regional centromere that is commonly comprised of tandem repeat arrays and transposable elements, budding yeast has a point centromere [81]. Despite the seemingly stark differences in DNA sequence organization and DNA binding capabilities, a pivotal study demonstrated that budding yeast centromeres are also transcribed into lncRNAs and that these lncRNAs are important for centromere homeostasis [54,55]. That a broad range of species produce lncRNAs from centromeric DNA suggests that centromeric lncRNAs serve a functional role in centromere biology. A key missing part of the puzzle is whether centromeric transcription also occurs in holocentric species. Work on the nematode *Caenorhabditis elegans* [82] and lepidopteran *Bombyx mori* [83] indicate anti-correlations between active transcription and the localizations of CENP-A and CENP-T chromatin, respectively. These data were obtained in asynchronized cells, which might not reveal low-level transcription at centromeric chromatin. It will be interesting to learn whether holocentric transcription indeed does occur, despite the complicated nature of studying centromeres with a diffuse phenotype.

2.2. Centromeric Small RNAs

In addition to lncRNAs, small RNAs (<200 nt in length) from centromeric DNA have also been found across a broad range of species (Table 1, Figure 1). Indeed, centromeric small RNAs were first isolated from *Arabidopsis* [73], after strong clues came from fission yeast studies that showed that deletion of components of the RNAi machinery resulted in aberrant accumulation of pericentromeric transcripts [64,65]. The centromeric structure of fission yeast is distinctly different from that of most plants and animals. The centromere core, where CENP-A nucleosomes are localized, is flanked by pericentric inverted repeats [84]. These repeats are transcribed and essential for pericentric heterochromatin formation [85–87]. This clear-cut demarcation between the functional centromere and pericentromere is much fuzzier in repeat-rich regional centromeres. Nevertheless, a similar increase in the abundance of centromeric small RNAs was observed in *Arabidopsis* mutants of the RNAi machinery and histone deacetylases [73]. Shortly after the identification of centromeric small RNAs in *Arabidopsis*,

small RNAs were isolated from non-mutant rice plants [75]. In addition, centromeric small RNAs have also been identified in the *Plasmodium* parasite [78], sugar beet [76], red flour beetle [50], and Tamar wallaby [45]. These findings provide further evidence that small RNAs are produced from centromeric DNA across kingdoms and that these small RNAs are strongly associated with the RNAi machinery.

2.3. Post-Transcriptional Processing of Centromeric Transcripts

It has long been appreciated that more RNA is produced than is strictly needed to maintain all cellular functions. In addition to pre-mRNA being processed into a mature mRNA, many RNAs are actively degraded [88,89]. Although genes can be embedded in centromeric chromatin [90], repetitive centromeric DNA is not known to encode proteins. Nevertheless, polyadenylation isolation protocols consistently detect centromeric transcripts across kingdoms (Table 1), providing evidence that cenRNAs can be processed like lncRNAs and mRNAs. These two latter groups of transcripts also tend to undergo splicing events to produce mature transcripts [91,92]. Indeed, centromeric RNAs are associated with splicing factors (Table 1). In the case of fission yeast, the pericentric non-coding *dg* RNA contains an intron that plays an important role in the recruitment of the RNAi-machinery needed for pericentric heterochromatin formation [63]. Human α -satellite has been found to associate with the RNA helicase DHX38 [32], which is thought to be important for pre-mRNA splicing [93]. Although the data are sparse thus far, they do provide tantalizing hints that centromeric RNAs interact with various RNA processing components. Whether splicing itself is an important maturation step for cenRNAs or whether the splicing machinery serves as a physical link with other nuclear components remains to be determined.

In several species, both centromeric lncRNA and small RNAs have been found (Table 1). Thus far though, in most species, only one of the two RNAs has been identified. One technical reason for this may be that different RNA isolation methods do not isolate all types of RNA equally, especially if the RNA types are not equally abundant or stable [94]. Identification of centromeric transcripts depends on knowing the sequence composition of the actively transcribing centromere in each species. Understanding the relationship between centromeric lncRNAs and small RNA production is important. Recently, the sequence of the fruit fly centromere was deciphered [95] and this has opened up a new avenue to study the evolution of centromeric transcription in a highly tractable system. Indeed, in *Drosophila* species centromeric DNA is highly repetitive and fast-evolving [96]. In other words, it will be of interest to understand the evolutionary conservation of the production of both centromeric lncRNAs and small RNAs. Furthermore, uncovering when and how centromeric transcripts are processed and with which biological function this processing is associated is an exciting avenue for future research.

3. Functions of Centromeric Transcription

A logical prediction for the existence of the two classes of centromeric RNAs is that the lncRNA is the precursor of the small RNAs to ultimately produce pericentromeric heterochromatin (Figure 2). Yet non-centromeric, long non-coding RNAs have been implicated in a much broader range of cellular functions, including chromatin architecture, chromatin remodeling, transcriptional regulation, formation of nuclear bodies, and translational regulation [8,97]. Overall, lncRNA can function as a guide, scaffold, decoy, or signal [6]. In fact, well-known chromatin binding factors, such as CTCF [98] and the polycomb complex [99], not only bind distinct DNA motifs but are also functionally associated with RNA. This raises the question, what is the full breadth of the functional implications of centromeric transcription?

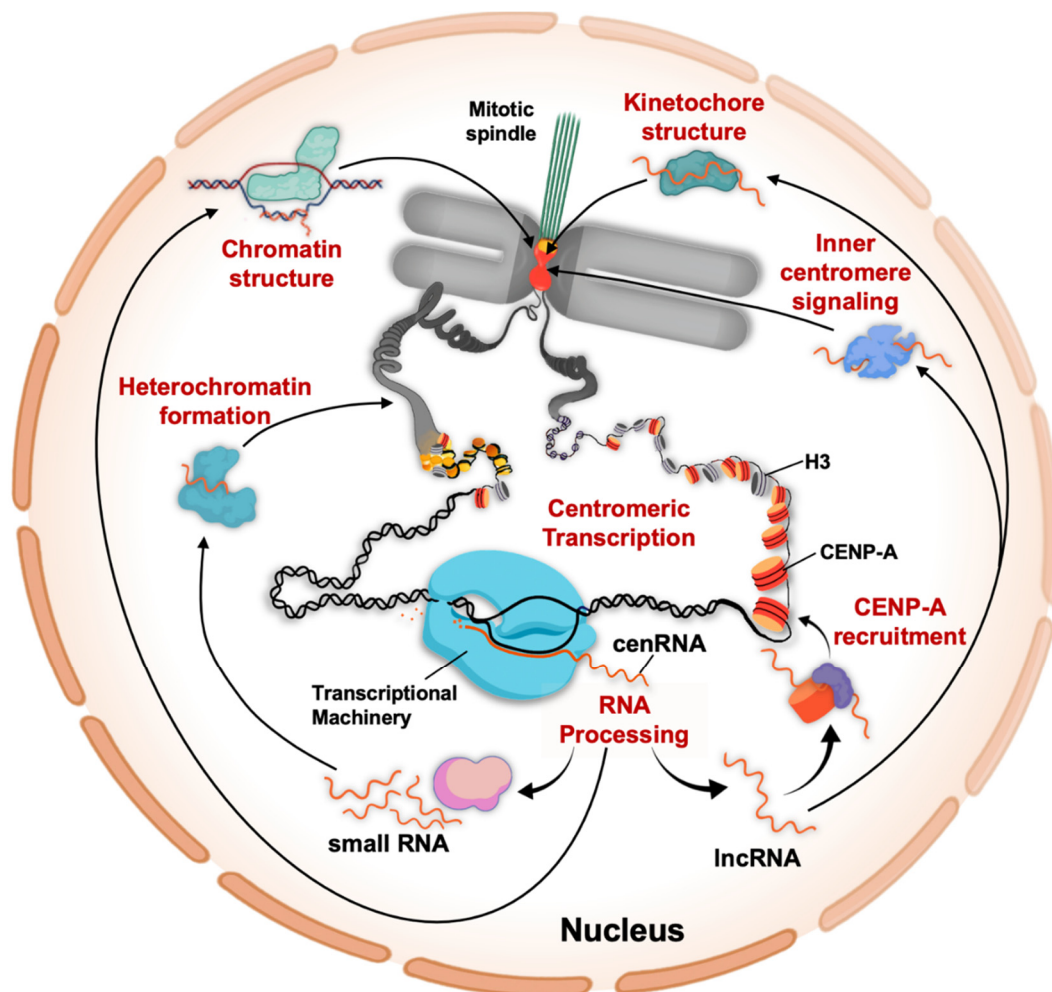


Figure 2. Biological functions of centromeric transcripts. The centromere-specific nucleosomes are distributed in a species-specific manner. For instance, in fission yeast, CENP-A^{Cnp1} nucleosomes are restricted to the centromere core, whereas in humans CENP-A chromatin is interspersed with H3 chromatin at the centromere. Despite these different distribution patterns of CENP-A chromatin, centromeric transcription has been observed in a broad range of species. Centromeric transcripts are processed and are found as either lncRNAs or small RNAs. Through the argonaunt/dicer machinery, centromeric small RNAs facilitate the pericentric heterochromatin formation. Centromeric chromatin structure can be modulated through DNA-RNA hybrids as well as through CENP-B, both of which associate with centromeric transcripts. Centromeric lncRNAs facilitate recruitment and loading of new CENP-A nucleosomes, whereas the kinetochore component CENP-C pulls down centromeric transcripts as well, contributing to kinetochore formation and structure. Finally, Aurora B signaling at the inner centromere is critical for faithful chromosome segregation, and centromeric transcription is important for this process.

3.1. Recruitment and Loading of CENP-A

During every cell division, the total number of centromeric CENP-A nucleosomes is halved [100]. To guarantee continued faithful chromosome segregation over many cell cycles, the number of centromeric CENP-A nucleosomes must be maintained to guarantee that the kinetochore can be formed. Thus, the centromeric CENP-A nucleosome pool must be replenished with new CENP-A nucleosomes. Indeed, CENP-A nucleosomes are loaded on centromeric chromatin in a cell cycle-specific manner [54,100–103]. In the last several years it has been shown that centromeric transcription plays critical roles in various aspects of loading of new CENP-A nucleosomes. The first clues came from studies in fission yeast; various mutants of the RNAP2 machinery, including transcription initiation and

elongation factors, showed reduced levels of CENP-A^{Cnp1} [57,104]. Direct evidence came from work in mammalian cell lines and *Xenopus* oocyte extracts where knock-down of centromeric transcripts resulted in reduced CENP-A levels at the centromere [21,24–26,33,48]. Indeed, recent work in mammalian and fruit fly cell lines showed that chemical inhibition of activated RNAP2 resulted in the loss of centromeric CENP-A^{CID} chromatin [21,52]; and the elongation factor Spt6 facilitates maintenance of centromeric CENP-A^{CID} [105]. These lines of investigation strongly hint at a distinct role for both the act of transcription and the production of an RNA species in the loading of CENP-A.

It is not very likely that new CENP-A is loaded exclusively on naked DNA, rather than on chromatin. Indeed, in human and fruit fly cell lines, active transcription is needed to remove so-called placeholder H3.3 nucleosomes prior to new CENP-A loading at the centromere [53,105,106]. Similarly, in fission yeast, H3 nucleosomes function as placeholder nucleosomes [61]. None of these experiments addressed how CENP-A:H4 is recruited to the centromere. In many species, a dedicated CENP-A chaperone has been identified that performs this function. Human cells use HJURP [107,108] and RbAp48 [109], fruit flies use CAL1 [110], fission and budding yeast uses SCM3 [111–114], the holocentric roundworm *Caenorhabditis elegans* uses RbAp46/48^{LIN-53} [115], and *Arabidopsis thaliana* uses NASP^{SIM3} [116]. Although knock-down of fruit fly 359-bp satellite RNA resulted in reduced levels of CAL1 at the centromere [51], HJURP has been shown to directly interact with α -satellite lncRNAs [21]. Additionally, soluble pre-assembled HJURP/CENP-A complexes were observed, and knock-down of these α -satellite lncRNAs resulted in reduced levels of CENP-A and HJURP at the centromere [21]. Altogether, a picture emerges wherein centromeric transcription plays multiple active roles in both recruitment and incorporation of new CENP-A nucleosomes, including the eviction of placeholder nucleosomes. This begs the question: What are the precise roles of each step of transcription of centromeric DNA and subsequent post-transcriptional RNA processing in establishing and maintaining centromere identity?

3.2. Recruitment of CENP-C

Recent work showed that knock-down of CENP-C resulted in increased centromeric transcription [117], whereas knock-down of centromeric transcripts resulted in increased levels of centromeric CENP-C [31]. These data suggest that there is a functional interaction between CENP-C and centromeric transcription. The question remains as to whether this is an indirect effect, driven by CENP-A, or a direct effect. A direct association between CENP-C and centromeric transcripts is possible, as CENP-C has extensive nucleic acid binding activity [24,69,118,119], and both budding yeast and human CENP-C associates with AT-rich DNA [118,119]. Interestingly, *in vitro*, a 24-nt single-stranded centromeric RNA facilitates the binding of maize CENP-C to centromeric DNA [69]. Outside of the centromere, RNA–DNA triplexes are stabilized by nucleosomes *in vitro*; and in human cells, these RNA–DNA triplex structures were enriched at active regulatory sites [120]. This raises the possibility that RNA has a pivotal function in stabilizing CENP-C at the centromere, either in a CENP-A-dependent or CENP-A-independent manner. It will be interesting to dissect how centromeric RNA contributes to CENP-C localization and stabilization on centromeric DNA in other eukaryotic species.

3.3. Centromeric RNA, CENP-B, and Centromeric Chromatin Structure

Interestingly, the DNA binding centromeric protein CENP-B has also been found to associate with centromeric RNAs (Table 1). CENP-B binds to the 17-bp CENP-B box sequence [121], which is found in a subpopulation of α -satellite DNA, which can be occupied by either CENP-A or H3 nucleosomes [122–124]. In addition, CENP-B binds at the DNA entry site of CENP-A nucleosomes, helping to phase CENP-A nucleosomes on centromeric DNA [125]. On α -satellite arrays on human artificial chromosomes, the presence of CENP-B is critical for the recruitment of CENP-A nucleosomes [126,127]. Similarly, when endogenous CENP-A is degraded using the auxin-inducible system, CENP-B still partially allows centromeres to form a functional kinetochore [128]. Interestingly,

a recent study showed that CENP-B serves as a “beacon” for H3.3 incorporation [129]. Furthermore, in a recent fluorescence microscopy-based interaction-trap assay, components of both heterochromatin (SUV39H1 and HP1) and open chromatin (including histone methyltransferases ASH1L and NSD1) were recruited by CENP-B [130]. These results point towards a role for CENP-B in organizing centromeric chromatin structures to be competent for both new CENP-A loading and heterochromatin formation. Nevertheless, it remains poorly understood whether or not, and if so how, centromeric transcription facilitates CENP-B’s role in chromatin organization. Therefore, understanding whether or not, and if so how, centromeric transcripts associated with CENP-B help to distinguish between the formation of these two mutually exclusive chromatin domains is important. From a chromatin organizational perspective, CENP-B can potentially cross-link neighboring chromatin fibers through its dimerization domain [131]. In addition, RNA–DNA triplexes [132], such as R-loops, have also been observed at centromeres, and these R-loops were associated with the ATR kinase, which is important in safeguarding genome stability [133]. In maize, circular RNA derived from centromeric CRM retroelements potentially associated with two R-loops, and thereby formatted chromatin loops. Knock-down of these circular RNAs resulted in the loss of chromatin loops and reduced levels of CENP-A^{CENH3} [72]. These latter studies imply that R-loops might be involved in organizing local chromatin structures which might contribute to the regulation of ATM/ATR kinase activity. Thus, understanding how centromeric chromatin conformation contributes to both centromere organization and centromere maintenance is an important question to explore.

3.4. Inner Centromere Signaling

In addition to the functions described above, centromeric transcripts have another critical role in centromere biology. Prior to anaphase, the mitotic checkpoint must be met, guaranteeing that all chromosomes are properly oriented and attached to the mitotic spindles. At the chromatin region (inner centromere) between the inter-sister chromatids, the chromosomal passenger complex (CPC) accumulates [134] and senses and responds to the pulling forces generated at the kinetochores [135,136]. The CPC is comprised of the mitotic kinase Aurora B, INCENP, survivin, and borealin [134]. Recently, CPC was reported to form coacervates, which are thought to be functionally important [137]. Interestingly, in various vertebrate species, CPC components have been shown to pull down centromeric transcripts (Table 1). In particular, when Aurora B is bound to centromeric RNA, it regulates both its activity and its localization [47]. Shugoshin (SGO1) is thought to protect centromeric cohesion from cleavage during prophase [138], but SGO1 also brings RNAP2 to the centromere [29] and indeed associates with centromeric RNA as well [29]. Mitotic transcription specifically is important for the spatiotemporal functioning of CPC at the inner centromeres [47]. Interestingly, using a nucleosome affinity library a recent study showed that the CPC interacts with the nucleosome acid patch [139] and the acid patch is known to modulate higher-order chromatin structure [140]. Thus, how centromere chromatin structures dictate the recruitment of SGO1 and CPC; and how CPC modifies the centromere chromatin to facilitate mitotic progression are exciting avenues for future studies.

3.5. Pericentromeric Heterochromatin

In fission yeast, the importance of pericentromeric transcription for heterochromatin formation has been extensively studied (reviewed here [86]). Beyond fission yeast, only a few studies have addressed directly whether centromeric transcription plays an important role in pericentromeric heterochromatin formation [23,43,141]. Why are CENP-A nucleosomes deposited where they are and not somewhere else on the chromosome? In addition, ectopic CENP-A^{CID} prefers to seed at euchromatin-heterochromatin boundaries [142]. This latter observation suggests that CENP-A chromatin prefer to be right next to a heterochromatin domain. If centromeric transcription can create a chromatin environment that is conducive to both heterochromatin formation and kinetochore establishment, it would be the ideal double-edged sword. Two recent studies in human and mouse cells [23,141] showed that centromeric RNAs are important for heterochromatin formation. In the

paternal mouse pronucleus, pericentric transcripts suppress SUV39H2 activity, resulting in subsequent reduced H3K9me3 levels [141], potentially through the formation of RNA–DNA triplexes near nucleosomes [43]. In contrast, in human cells, SUV39H1 needs to bind α -satellite RNA to establish constitutive heterochromatin [23]. The latter scenario is more reminiscent of what is observed in fission yeast, wherein, through a dicer-mediated pathway, constitutive heterochromatin is established and maintained [86]. Indeed, in mouse cell lines, dicer has been associated with centromeric RNA [41]. It is therefore conceivable that centromeric small RNAs might have specific developmental roles in establishing where and when constitutive heterochromatin is formed. The make-up of the respective protein complexes with which centromere small RNAs associate might provide important clues.

4. Centromeric Transcription in Disease

Given the biological functions with which centromeric transcripts are thought to be associated, misregulation seems almost inevitable. Indeed, the misregulation of mouse minor satellite RNAs leads to impaired centromere function and defective chromosome segregation [35]. In addition, work in budding yeast has focused on how CENP-A^{Cse4} loading is limited to the centromere [143–147]. Whether the removal of ectopic CENP-A in non-Ascomycota species happens by the same mechanism with the same efficiency remains to be determined. The fact that centromeric DNA is not conserved strongly argues against a model wherein the sequence is the driving factor in the functional consequences of centromeric transcription. Thus, the question of whether transcription at non-centromeric loci can hijack centromere-like functions is at the heart of ectopic centromeric chromatin formation.

4.1. Ectopic CENP-A and Neocentromeres

In many cancers, various centromere and kinetochore components are overexpressed, including CENP-A and its dedicated chaperone HJURP [148]. Overexpressed CENP-A has been found to not only associate with HJURP but also with the H3.3 chaperones DAXX/ARTX and HIRA [149–152], resulting in ectopic incorporation of CENP-A nucleosomes. One potential consequence of ectopic CENP-A chromatin sites is the formation of dicentric chromosomes, rendering the chromosome unstable [153]. Indeed, ectopic CENP-A sites promote the formation of ectopic kinetochores and subsequent mitotic defects [150]. Similarly, artificially overexpressed CENP-A^{CID} in Schneider S2 cells was also able to form functional ectopic kinetochores [142]. Similarly, in fission yeast [154] and budding yeast [155,156], ectopic localization of CENP-A occurs when CENP-A is overexpressed. In another experimental setup, ectopic CENP-A was found when the innate centromere was deleted [157].

These results give the impression that CENP-A readily goes ectopic, consistently posing the risk of creating dicentric chromosomes. One would expect that mechanisms have evolved to reduce this risk to a minimum without impacting the native centromere functions. In the case of the point centromere of budding yeast, which is thought to contain only of a single CENP-A^{Cse4} nucleosome [158], SUMOylation of N-terminal tail of CENP-A^{Cse4} is a driving force in preventing CENP-A^{Cse4} mislocalization [143–147]. In plants, active removal of CENP-A^{Cse4} has also been observed [159]. In contrast, a recent study in human cell lines suggested that ectopic CENP-A nucleosomes are removed during replication [160]. Nevertheless, in various cancers, ectopic CENP-A nucleosomes are readily detected [149,161]. This either means that, in the case of the human cells, replication has a limited capacity to correct ectopic localization of CENP-A nucleosomes, or an alternative mechanism is actively recruiting CENP-A to ectopic sites. As centromeric DNA evolves very quickly, it is unlikely that a strong sequence signal is driving the functional associations between centromeric transcripts and new CENP-A loading. Nevertheless, neocentromeres and ectopic CENP-A sites are not randomly distributed across the human genome [149], implying that certain loci have centromeric potential, whereas other loci do not. Studies on human artificial chromosomes showed that a chromatin environment containing a distinct set and distribution of histone modifications was important for CENP-A chromatin formation as well [162–164]. Using the marker chromosome mardel(10), L1 retrotransposon transcripts have been found to facilitate the formation of CENP-A

chromatin [165]. Given these results, it is therefore tempting to speculate that some non-centromeric regions might have the capacity to mimic transcriptional conditions of the centromere.

4.2. Centromeric Transcription in Cancer

What happens when centromeric transcription is misregulated? One may predict that when centromeric transcription is suppressed, essential centromeric functions during mitosis and CENP-A chromatin maintenance is impaired, resulting in either cell death or senescence. On the other hand, overexpression of centromeric transcripts may result in aberrant regulation of mitosis, resulting in genomic instability. Indeed, in various human cancers, α -satellite transcripts are overexpressed [166], and overexpression of these transcripts strongly correlates with a poor clinical prognosis [148]. In human and murine cell lines centromeric transcripts can accumulate as a result of DNA demethylation, heat shock, or the induction of apoptosis [167]. In these cases, and overexpression of centromeric RNA, genomic instability is observed [167,168], and genome instability is one of the hallmarks of cancer [169].

In human mammary epithelial cells derived from several patients, artificial overexpression of α -satellite sequences resulted in both hypomethylation of centromeric DNA and a significant increase in segregation errors, especially in chromosomes 8 and 20 [170]. Furthermore, in mouse and human breast tumors lacking BRCA1, satellite RNAs are highly expressed [171,172], as BRCA1 facilitates the monoubiquitination of histones associated with satellite DNA, thereby suppressing transcription [171]. When satellite RNAs are overexpressed the DNA damage response pathways are activated, inducing aneuploidy, which, in mice, is sufficient for tumor development in mammary glands [172]. Interestingly, some centromeric transcripts remained associated with distinct genomic loci, even throughout mitosis [173,174], promoting the localization of chromatin factors, such as PRC1 and MeCP2 [175]. Furthermore, centromeric transcripts have also been associated with repeat expansion [176]. Whether misregulation of centromeric transcription also facilitates the expansion of the pericentric HERV-K retrovirus [177] remains to be determined. In a recent study in fission yeast, mutations in RNAP2 or TFIIS^{Tfs1}, a transcription factor that facilitates RNAP2 restarting after backtracking, increased the incidence of chromosomal rearrangements [178]. All in all, these studies hint at the possibility that genomic instability as a consequence of aberrant centromeric transcription is caused by a broad range of nuclear functions.

Thus, what are the functional consequences of overexpressing of centromeric transcripts that can be directly attributed to these transcripts? One clue how a non-coding RNA can impact nuclear organization and function comes from the overexpression of the unique lncRNA: ChRO1. ChRO1 expressed specifically during the terminal myoblast differentiation and is involved in the reorganization of constitutive heterochromatin [179]. Inhibition of ChRO1 leads to aberrant chromatin organization and mislocalization of histone marks and chromatin binding factors in murine cells [179]. In particular, ChRO1 mediates centromeric RNA accumulation at centromeric sites through DAXX/ARTX/H3.3 [179]. Interestingly, in fission yeast, a link between the maintenance and nuclear localization of pericentric heterochromatin and histone chaperone complexes was also observed [180].

As noted above, overexpressed CENP-A can associate with H3.3 chaperones [149–152]. Therefore, in many cancers, both CENP-A and centromeric transcripts are frequently overexpressed. Whether these two observations are dependent on each other or have a cumulative or even synergistic effect on tumor development is still an unanswered question. How the mislocalization of histone variants is driven by the expression of non-coding RNAs and how histone chaperone complexes are guided by non-coding RNAs remain of subjects of intense investigation. It is interesting to speculate that the role of histone chaperones go beyond assisting histone localization and nucleosomes dynamics [181]. Maybe, histone chaperones function as a critical factor contributing to chromatin domain formation, long-distance chromatin interactions, and even mechanosensing [182] by bringing together nucleosomes, non-coding RNAs, and other chromatin binding factors. Consequently, misregulation would lead to a broad range of chromatin effects, which are commonly observed in cancers.

5. Conclusions and Future Perspective

Centromeres are genetically different across species and they do not follow synteny [5]. In addition, frequent ectopic incorporation of CENP-A and the formation of neocentromeres [28,149,150] clearly indicate that centromere formation is not primarily driven by genetic factors, but rather a complex interplay between various epigenetic mechanisms. In most species [183], nucleosomes containing centromere-specific histone variants and their cognate binding partners form the proteinaceous basis of the kinetochore. Over the last few decades, it has become apparent that centromeric transcription is found across the eukaryotic tree (Figure 1). Although not every function has been shown to exist in every species studied so far (Table 1), it is reasonable to predict that these transcripts fulfill a multitude of roles (Figure 2), as if centromeric transcription were a swiss-army knife.

Several big questions still await answering—not only how well conserved these functions are, but more importantly, how centromeric transcripts precisely perform their functions. What attracts centromeres and neocentromeres to be adjacent to heterochromatin? Interestingly, it is paradoxical that transcriptionally silent heterochromatin requires transcription to be formed. Is heterochromatin-associated transcription linked to the loading of new CENP-A nucleosomes? Thus, an obvious question is whether lncRNAs and small RNAs are produced from the same transcript (Figure 3A). If this were the case, this would allow for precise spatiotemporal regulation of centromeric RNAs. One primary function of centromeric transcripts appears to be loading of both CENP-A and CENP-C, setting the stage to form a functional kinetochore. RNA–DNA triplexes have been implicated in loading both proteins, thereby unraveling the role of RNA–DNA triplexes in both CENP-A and CENP-C loading (Figure 3B). Whether these loading events are synergistic is important for understanding how CENP-A chromatin is formed and how it spreads. Indeed, tantalizing evidence shows a positive feedback loop between CENP-C^{CID}, dCENP-C, and CAL1 [184]. Subsequently, whether RNA has a structural role in the formation and/or maintenance of kinetochore is also an unexplored question (Figure 3C). It would be intriguing if different RNases could be linked to different kinetochore components, allowing the study of the structural role of RNA at the kinetochore. At the inner centromere, CPC is localized and activated through centromeric transcription during mitosis. Another unanswered question is whether centromeric transcription is sensitive to mitotic pulling forces (Figure 3D), potentially linking mitotic biomechanical events to direct transcriptional outputs [182]. In cancers, ectopic CENP-A nucleosomes are commonly found, yet it remains unclear how precisely CENP-A is recruited to these sites. One interesting possibility is that ectopic sites can mimic structural features of centromeric RNA, thereby hijacking CENP-A away from the centromere (Figure 3E). This could also have potential evolutionary implications if RNA folding motifs are a driving force in the potential of a sequence to become a centromere. Finally, the holy grail in studying centromeres is the ability to be able to genetically manipulate centromeres. Being able to recruit components to the centromere at will, and visualize and track transcriptional activity, have not been possible because of the centromere's highly repetitive nature. One possibility might be to utilize the *Tal1* LTR retrotransposon from *Arabidopsis lyrata* (Figure 3F). This LTR is absent in *A. thaliana*, yet when introduced, it is still specifically incorporated at *A. thaliana*'s centromeres [185]. If this system could be engineered to study centromere biology, including centromeric transcription, a range of previously untestable hypotheses could finally be tested. The recent publication of the first telomere-to-telomere sequence of a human chromosome [186,187] might even make human centromeres a target for such studies.

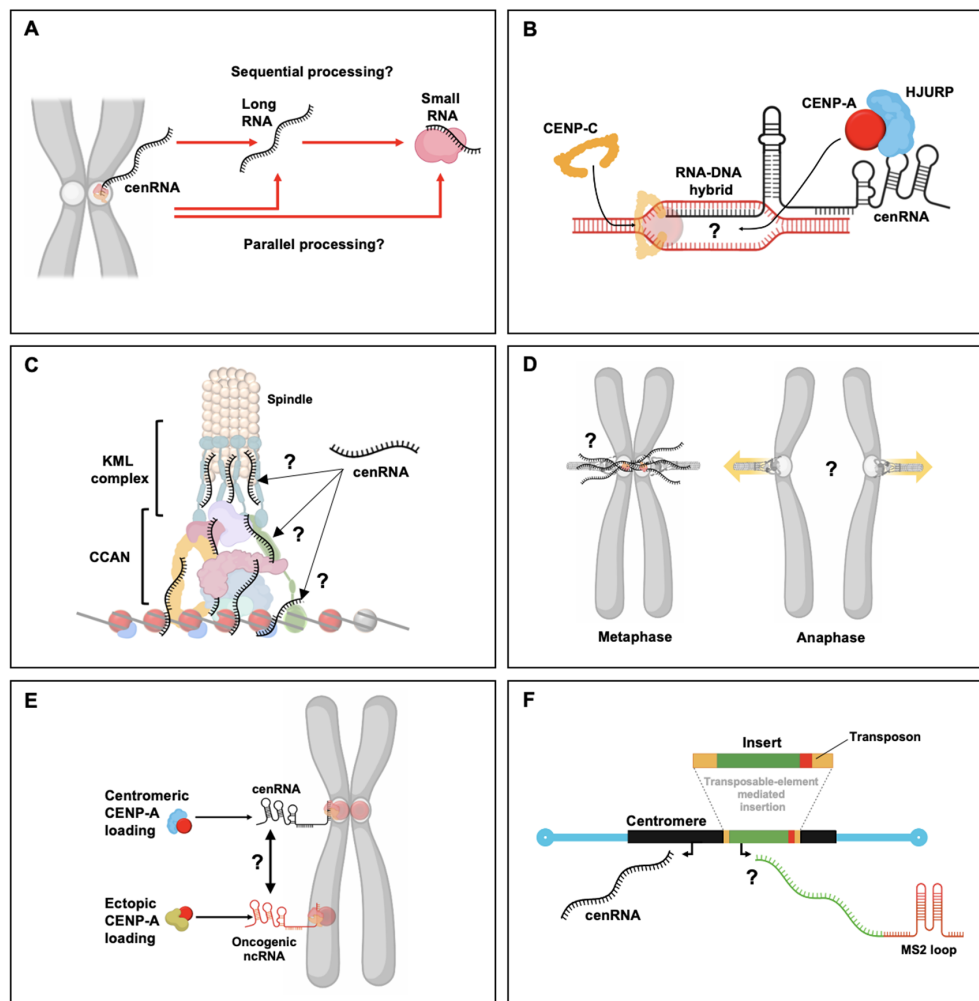


Figure 3. Open questions about the functional consequences of centromeric transcription. (A) Even though two different types of RNAs are produced from the centromere, namely, lncRNAs and small RNAs, it remains unclear whether these two types of RNAs are produced sequentially or in parallel. (B) Both CENP-A and CENP-B associate with centromeric transcripts and both CENP-A loading and CENP-C loading have been linked to RNA–DNA triplex formation. Additionally, in maize, CENP-C binding to centromeric DNA is associated with small RNA, whereas in humans CENP-A loading has been tied to lncRNAs. It remains unclear whether the same transcript can recruit and stabilize both CENP-A and CENP-C at centromeric chromatin. (C) As CENP-A and CENP-C loading is cenRNA-dependent, one intriguing question is whether the kinetochore structure as a whole is dependent on the presence of centromeric transcripts. If so, which kinetochore components drive this dependence? (D) The localization and activity of CPC at the inner centromere are dependent on centromeric transcripts. It will be of great interest to understand how mitotic pulling forces modulate centromeric transcription, especially during the transition from metaphase to anaphase. (E) CENP-A nucleosomes have been found both at the centromere and ectopically. Especially in cancers, ectopic CENP-A accumulates. Little sequence conservation exists between ectopic CENP-A sites and the centromere. This leaves the possibility that specific secondary and tertiary RNA structures exist that are found similar to/in common between transcripts derived from the centromere and ectopic sites, allowing ectopic sites to hijack CENP-A recruitment. In addition, this might also provide critical insights into the evolution of centromere DNA from a functional perspective as a CENP-A recruitment motif. (F) Whereas various loci on the chromosome arms have been extensively studied using genetics tools such as LacO operons and MS2/PP7 stem-loops, the centromere has been recalcitrant to genetic manipulations because of its highly repetitive nature. In *Arabidopsis thaliana*, the *A. lyrata* Tal1 LTR retrotransposon specifically integrates into the *A. thaliana* centromere. This opens the door to finally genetically modify the centromere to be able to study transcription in an inducible and tractable manner, and recruit proteins of interest.

All in all, the centromere transcription field is still in its early days, as many critical questions remain unanswered. Being able to learn how a fast-evolving sequence can still perform such a broad range of essential functions is especially critical for understanding how the genome integrity is regulated.

Funding: This research received no external funding.

Acknowledgments: We thank Yamini Dalal and Minh Bui for critical reading of the manuscript and David Posner for spelling and grammar suggestions. G.A.K. and D.P.M. are supported by the Intramural Research Program of the National Institutes of Health and did not receive external funding.

Conflicts of Interest: The authors declare no competing or financial interests.

References

1. Malik, H.S.; Henikoff, S. Adaptive evolution of Cid, a centromere-specific histone. *Dros. Genet.* **2001**, *157*, 1293–1298.
2. Malik, H.S.; Vermaak, D.; Henikoff, S. Recurrent evolution of DNA-binding motifs in the drosophila centromeric histone. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1449–1454. [[CrossRef](#)] [[PubMed](#)]
3. Cooper, J.L.; Henikoff, S. Adaptive evolution of the histone fold domain in centromeric histones. *Mol. Biol. Evol.* **2004**, *21*, 1712–1718. [[CrossRef](#)] [[PubMed](#)]
4. Melters, D.P.; Bradnam, K.R.; Young, H.A.; Telis, N.; May, M.R.; Ruby, J.G.; Sebra, R.; Peluso, P.; Eid, J.; Rank, D.; et al. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol.* **2013**, *14*, R10. [[CrossRef](#)]
5. Henikoff, S.; Ahmad, K.; Malik, H.S. The centromere paradox: Stable inheritance with rapidly evolving DNA. *Science* **2001**, *293*, 1098–1102. [[CrossRef](#)] [[PubMed](#)]
6. Hezroni, H.; Koppstein, D.; Schwartz, M.G.; Avrutin, A.; Bartel, D.P.; Ulitsky, I. Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep.* **2015**, *11*, 1110–1122. [[CrossRef](#)] [[PubMed](#)]
7. Ulitsky, I. Evolution to the rescue: Using comparative genomics to understand long non-coding RNAs. *Nat. Rev. Genet.* **2016**, *17*, 601–614. [[CrossRef](#)]
8. Yao, R.W.; Wang, Y.; Chen, L.L. Cellular functions of long noncoding RNAs. *Nat. Cell Biol.* **2019**, *21*, 542–551. [[CrossRef](#)]
9. Zhang, X.; Wang, W.; Zhu, W.; Dong, J.; Cheng, Y.; Yin, Z.; Shen, F. Mechanisms and functions of long non-coding RNAs at multiple regulatory levels. *Int. J. Mol. Sci.* **2019**, *20*, 5573. [[CrossRef](#)]
10. Ulitsky, I. Interactions between short and long noncoding RNAs. *FEBS Lett.* **2018**, *592*, 2874–2883. [[CrossRef](#)]
11. Barrios, N.R.; Legascue, M.F.; Benhamed, M.; Ariel, F.; Crespi, M. Splicing regulation by long noncoding RNAs. *Nucleic Acids Res.* **2018**, *46*, 2169–2184. [[CrossRef](#)] [[PubMed](#)]
12. Guo, C.J.; Ma, X.K.; Xing, Y.H.; Zheng, C.C.; Xu, Y.F.; Shan, L.; Zhang, J.; Wang, S.; Wang, Y.; Carmichael, G.G.; et al. Distinct processing of lncRNAs contributes to non-conserved functions in stem cells. *Cell* **2020**, *181*, 621–636.e22. [[CrossRef](#)] [[PubMed](#)]
13. Montefalcone, G.; Tempesta, S.; Rocchi, M.; Archidiacono, N. Centromere repositioning. *Genome Res.* **1999**, *9*, 1184–1188. [[CrossRef](#)] [[PubMed](#)]
14. Djebali, S.; Davis, C.A.; Merkel, A.; Dobin, A.; Lassmann, T.; Mortazavi, A.; Tanzer, A.; Lagarde, J.; Lin, W.; Schlesinger, F.; et al. Landscape of transcription in human cells. *Nature* **2012**, *489*, 101–108. [[CrossRef](#)]
15. Dunham, I.; Kundaje, A.; Aldred, S.F.; Collins, P.J.; Davis, C.A.; Doyle, F.; Epstein, C.B.; Frietze, S.; Harrow, J.; Kaul, R.; et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **2012**, *489*, 57–74.
16. Haberle, V.; Stark, A. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 621–637. [[CrossRef](#)]
17. Kwak, H.; Lis, J.T. Control of transcriptional elongation. *Annu. Rev. Genet.* **2013**, *47*, 483–508. [[CrossRef](#)]
18. Petrenko, N.; Jin, Y.; Dong, L.; Wong, K.H.; Struhl, K. Requirements for RNA polymerase II preinitiation complex formation in vivo. *Elife* **2019**, *8*, e43654. [[CrossRef](#)]
19. McKenzie, E.A.; Kent, N.A.; Dowell, S.J.; Moreno, F.; Bird, L.E.; Mellor, J. The centromere and promoter factor 1, CPFI, of *Saccharomyces cerevisiae* modulates gene activity through a family of factors including SPT21, RPD1 (SIN3), RPD3 and CCR4. *MGG Mol. Gen. Genet.* **1993**, *240*, 374–386. [[CrossRef](#)]

20. Kent, N.A.; Eibert, S.M.; Mellor, J. Cbf1p is required for chromatin remodeling at promoter-proximal CACGTG motifs in yeast. *J. Biol. Chem.* **2004**, *279*, 27116–27123. [[CrossRef](#)]
21. Quénet, D.; Dalal, Y. A long non-coding RNA is required for targeting centromeric protein a to the human centromere. *Elife* **2014**, *3*, e03254. [[CrossRef](#)] [[PubMed](#)]
22. McNulty, S.M.; Sullivan, L.L.; Sullivan, B.A. Human centromeres produce chromosome-specific and array-specific alpha satellite transcripts that are complexed with CENP-A and CENP-C. *Dev. Cell* **2017**, *42*, 226–240. [[CrossRef](#)] [[PubMed](#)]
23. Johnson, W.L.; Yewdell, W.T.; Bell, J.C.; McNulty, S.M.; Duda, Z.; O'Neill, R.J.; Sullivan, B.A.; Straight, A.F. RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *Elife* **2017**, *6*, e25299. [[CrossRef](#)] [[PubMed](#)]
24. Wong, L.H.; Moore, K.H.B.; Chan, L.; Quach, J.M.; Anderson, M.A.; Northrop, E.L.; Hannan, R.; Saffery, R.; Shaw, M.L.; Williams, E.; et al. Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res.* **2007**, *17*, 1146–1160. [[CrossRef](#)] [[PubMed](#)]
25. Chan, F.L.; Marshall, O.J.; Saffery, R.; Kim, B.W.; Earle, E.; Choo, K.H.A.; Wong, L.H. Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1979–1984. [[CrossRef](#)]
26. Saffery, R.; Sumer, H.; Hassan, S.; Wong, L.H.; Craig, J.M.; Todokoro, K.; Anderson, M.; Stafford, A.; Choo, K.H.A. Transcription within a functional human centromere. *Mol. Cell* **2003**, *12*, 509–516. [[CrossRef](#)]
27. Palozola, K.C.; Donahue, G.; Liu, H.; Grant, G.R.; Becker, J.S.; Cote, A.; Yu, H.; Raj, A.; Zaret, K.S. Mitotic transcription and waves of gene reactivation during mitotic exit. *Science* **2017**, *358*, 119–122. [[CrossRef](#)]
28. Marshall, O.J.; Chueh, A.C.; Wong, L.H.; Choo, K.H.A. Neocentromeres: New insights into centromere structure, disease development, and karyotype evolution. *Am. J. Hum. Genet.* **2008**, *82*, 261–282. [[CrossRef](#)]
29. Liu, H.; Qu, Q.; Warrington, R.; Rice, A.; Cheng, N.; Yu, H. Mitotic transcription installs sgo1 at centromeres to coordinate chromosome segregation. *Mol. Cell* **2015**, *59*, 426–436. [[CrossRef](#)]
30. Ideue, T.; Cho, Y.; Nishimura, K.; Tani, T. Involvement of satellite I noncoding RNA in regulation of chromosome segregation. *Genes Cells* **2014**, *19*, 528–538. [[CrossRef](#)]
31. Bury, L.; Moodie, B.; McKay, L.S.; Miga, K.H.; Cheeseman, I.M. Alpha-satellite RNA transcripts are repressed by centromere-nucleolus associations. *bioRxiv* **2020**. [[CrossRef](#)]
32. Nishimura, K.; Cho, Y.; Tokunaga, K.; Nakao, M.; Tani, T.; Ideue, T. DEAH box RNA helicase DHX38 associates with satellite I noncoding RNA involved in chromosome segregation. *Genes Cells* **2019**, *24*, 585–590. [[CrossRef](#)] [[PubMed](#)]
33. Bergmann, J.H.; Rodríguez, M.G.; Martins, N.M.C.; Kimura, H.; Kelly, D.A.; Masumoto, H.; Larionov, V.; Jansen, L.E.T.; Earnshaw, W.C. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO J.* **2011**, *30*, 328–340. [[CrossRef](#)] [[PubMed](#)]
34. Probst, A.V.; Okamoto, I.; Casanova, M.; Marjou, F.E.; Baccon, P.L.; Almouzni, G. A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev. Cell* **2010**, *19*, 625–638. [[CrossRef](#)] [[PubMed](#)]
35. Segard, H.B.; Guais, A.; Francastel, C. Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8709–8714. [[CrossRef](#)]
36. Karslioglu, A.B.; Perrera, V.; Scaranaro, M.; Velazquez, I.A.D.L.R.; Nobelen, S.V.D.; Shukeir, N.; Popow, J.; Gerle, B.; Opravil, S.; Pagani, M.; et al. A transcription factor-based mechanism for mouse heterochromatin formation. *Nat. Struct. Mol. Biol.* **2012**, *19*, 1023–1030. [[CrossRef](#)]
37. Harel, J.; Hanania, N.; Tapiero, H.; Harel, L. RNA replication by nuclear satellite DNA in different mouse cells. *Biochem. Biophys. Res. Commun.* **1968**, *25*, 696–701. [[CrossRef](#)]
38. Cohen, A.K.; Huh, T.Y.; Helleiner, C.W. Transcription of satellite DNA in mouse L-cells. *Can. J. Biochem.* **1973**, *51*, 529–532. [[CrossRef](#)]
39. Ferri, F.; Segard, H.B.; Velasco, G.; Hubé, F.; Francastel, C. Non-coding murine centromeric transcripts associate with and potentiate aurora B kinase. *Nucleic Acids Res.* **2009**, *37*, 5071–5080. [[CrossRef](#)]
40. Mallm, J.P.; Rippe, K. Aurora kinase B regulates telomerase activity via a centromeric RNA in stem cells. *Cell Rep.* **2015**, *11*, 1667–1678. [[CrossRef](#)]
41. Lu, J.; Gilbert, D.M. Proliferation-dependent and cell cycle-regulated transcription of mouse pericentric heterochromatin. *J. Cell Biol.* **2007**, *179*, 411–421. [[CrossRef](#)] [[PubMed](#)]

42. Kanellopoulou, C.; Muljo, S.A.; Kung, A.L.; Ganesan, S.; Drapkin, R.; Jenuwein, T.; Livingston, D.M.; Rajewsky, K. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* **2005**, *19*, 489–501. [[CrossRef](#)] [[PubMed](#)]
43. Hsieh, C.L.; Lin, C.L.; Liu, H.; Chang, Y.J.; Shih, C.J.; Zhong, C.Z.; Lee, S.C.; Tan, B.C.M. WDHD1 modulates the post-transcriptional step of the centromeric silencing pathway. *Nucleic Acids Res.* **2011**, *39*, 4048–4062. [[CrossRef](#)] [[PubMed](#)]
44. Camacho, O.V.; Galan, C.; Rosowska, K.S.; Ching, R.; Gamalinda, M.; Karabiber, F.; Velazquez, I.D.L.R.; Engist, B.; Koschorz, B.; Shukeir, N.; et al. Major satellite repeat RNA stabilize heterochromatin retention of Suv39h enzymes by RNA-nucleosome association and RNA:DNA hybrid formation. *Elife* **2017**, *6*, e25293. [[CrossRef](#)]
45. Carone, D.M.; Longo, M.S.; Ferreri, G.C.; Hall, L.; Harris, M.; Shook, N.; Bulazel, K.V.; Carone, B.R.; Obergfell, C.; O'Neill, M.J.; et al. A new class of retroviral and satellite encoded small RNAs emanates from mammalian centromeres. *Chromosoma* **2009**, *118*, 113–125. [[CrossRef](#)]
46. Carone, D.M.; Zhang, C.; Hall, L.E.; Obergfell, C.; Carone, B.R.; O'Neill, M.J.; O'Neill, R.J. Hypermorphic expression of centromeric retroelement-encoded small RNAs impairs CENP-A loading. *Chromosom. Res.* **2013**, *21*, 49–62. [[CrossRef](#)]
47. Blower, M.D. Centromeric transcription regulates aurora-b localization and activation. *Cell Rep.* **2016**, *15*, 1624–1633. [[CrossRef](#)]
48. Grenfell, A.W.; Heald, R.; Strzelecka, M. Mitotic noncoding RNA processing promotes kinetochore and spindle assembly in *Xenopus*. *J. Cell Biol.* **2016**, *214*, 133–141. [[CrossRef](#)]
49. Jambhekar, A.; Emerman, A.B.; Schweidenback, C.T.H.; Blower, M.D. RNA stimulates aurora B kinase activity during mitosis. *PLoS ONE* **2014**, *9*, e100748. [[CrossRef](#)]
50. Pezer, Ž.; Ugarkovic, D. Satellite DNA-associated siRNAs as mediators of heat shock response in insects. *RNA Biol.* **2012**, *9*, 587–595. [[CrossRef](#)]
51. Rošić, S.; Köhler, F.; Erhardt, S. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J. Cell Biol.* **2014**, *207*, 335–349. [[CrossRef](#)] [[PubMed](#)]
52. Bobkov, G.O.M.; Gilbert, N.; Heun, P. Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation. *J. Cell Biol.* **2018**, *217*, 1957–1972. [[CrossRef](#)] [[PubMed](#)]
53. Chen, C.C.; Bowers, S.; Lipinszki, Z.; Palladino, J.; Trusiak, S.; Bettini, E.; Rosin, L.; Przewłoka, M.R.; Glover, D.M.; O'Neill, R.J.; et al. Establishment of centromeric chromatin by the CENP-A assembly factor CAL1 requires FACT-mediated transcription. *Dev. Cell* **2015**, *34*, 73–84. [[CrossRef](#)] [[PubMed](#)]
54. Ling, Y.H.; Yuen, K.W.Y. Point centromere activity requires an optimal level of centromeric noncoding RNA. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 6270–6279. [[CrossRef](#)] [[PubMed](#)]
55. Ohkuni, K.; Kitagawa, K. Endogenous transcription at the centromere facilitates centromere activity in budding yeast. *Curr. Biol.* **2011**, *21*, 1695–1703. [[CrossRef](#)] [[PubMed](#)]
56. Marczyński, G.T.; Jaehning, J.A. A transcription map of a yeast centromere plasmid: Unexpected transcripts and altered gene expression. *Nucleic Acids Res.* **1985**, *13*, 8487–8506. [[CrossRef](#)]
57. Choi, E.S.; Strålfors, A.; Castillo, A.G.; Dubief, M.D.; Ekwall, K.; Allshire, R.C. Identification of noncoding transcripts from within CENP-A chromatin at fission yeast centromeres. *J. Biol. Chem.* **2011**, *286*, 23600–23607. [[CrossRef](#)]
58. Chen, E.S.; Zhang, K.; Nicolas, E.; Cam, H.P.; Zofall, M.; Grewal, S.I.S. Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* **2008**, *451*, 734–737. [[CrossRef](#)]
59. Catania, S.; Pidoux, A.L.; Allshire, R.C. Sequence features and transcriptional stalling within centromere DNA promote establishment of CENP-A chromatin. *PLoS Genet.* **2015**, *11*, e1004986. [[CrossRef](#)]
60. Djupedal, I.; Portoso, M.; Spåhr, H.; Bonilla, C.; Gustafsson, C.M.; Allshire, R.C.; Ekwall, K. RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes Dev.* **2005**, *19*, 2301–2306. [[CrossRef](#)]
61. Shukla, M.; Tong, P.; White, S.A.; Singh, P.P.; Reid, A.M.; Catania, S.; Pidoux, A.L.; Allshire, R.C. Centromere DNA destabilizes H3 nucleosomes to promote CENP-A deposition during the cell cycle. *Curr. Biol.* **2018**, *28*, 3924–3936. [[CrossRef](#)] [[PubMed](#)]
62. Singh, P.P.; Shukla, M.; White, S.A.; Lafos, M.; Tong, P.; Auchynnikava, T.; Spanos, C.; Rappsilber, J.; Pidoux, A.L.; Allshire, R.C. Hap2-Ino80-facilitated transcription promotes de novo establishment of CENP-A chromatin. *Genes Dev.* **2020**, *34*, 226–238. [[CrossRef](#)] [[PubMed](#)]

63. Mutazono, M.; Morita, M.; Tsukahara, C.; Chinen, M.; Nishioka, S.; Yumikake, T.; Dohke, K.; Sakamoto, M.; Ideue, T.; Nakayama, J.I.; et al. The intron in centromeric noncoding RNA facilitates RNAi-mediated formation of heterochromatin. *PLoS Genet.* **2017**, *13*, e1006606. [[CrossRef](#)] [[PubMed](#)]
64. Volpe, T.A.; Kidner, C.; Hall, I.M.; Teng, G.; Grewal, S.I.S.; Martienssen, R.A. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science* **2002**, *297*, 1833–1837. [[CrossRef](#)] [[PubMed](#)]
65. Provost, P.; Silverstein, R.A.; Dishart, D.; Walfridsson, J.; Djupedal, I.; Kniola, B.; Wright, A.; Samuelsson, B.; Rådmark, O.; Ekwall, K. Dicer is required for chromosome segregation and gene silencing in fission yeast cells. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 16648–16653. [[CrossRef](#)] [[PubMed](#)]
66. Holoch, D.; Moazed, D. Small-RNA loading licenses argonaute for assembly into a transcriptional silencing complex. *Nat. Struct. Mol. Biol.* **2015**, *22*, 328–335. [[CrossRef](#)]
67. Ishida, M.; Shimojo, H.; Hayashi, A.; Kawaguchi, R.; Ohtani, Y.; Uegaki, K.; Nishimura, Y.; Nakayama, J.I. Intrinsic nucleic acid-binding activity of Chp1 chromodomain is required for heterochromatic gene silencing. *Mol. Cell* **2012**, *47*, 228–241. [[CrossRef](#)]
68. Keller, C.; Adaixo, R.; Stunnenberg, R.; Woolcock, K.J.; Hiller, S.; Bühler, M. HP1(Swi6) mediates the recognition and destruction of heterochromatic RNA transcripts. *Mol. Cell* **2012**, *47*, 215–227. [[CrossRef](#)]
69. Du, Y.; Topp, C.N.; Dawe, R.K. DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. *PLoS Genet.* **2010**, *6*, e1000835. [[CrossRef](#)]
70. Topp, C.N.; Zhong, C.X.; Dawe, R.K. Centromere-encoded RNAs are integral components of the maize kinetochore. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15986–15991. [[CrossRef](#)]
71. Koo, D.H.; Zhao, H.; Jiang, J. Chromatin-associated transcripts of tandemly repetitive DNA sequences revealed by RNA-FISH. *Chromosom. Res.* **2016**, *24*, 467–480. [[CrossRef](#)] [[PubMed](#)]
72. Liu, Y.; Su, H.; Zhang, J.; Liu, Y.; Feng, C.; Han, F. Back-spliced RNA from retrotransposon binds to centromere and regulates centromeric chromatin loops in maize. *PLoS Biol.* **2020**, *18*, e3000582. [[CrossRef](#)] [[PubMed](#)]
73. May, B.P.; Lippman, Z.B.; Fang, Y.; Spector, D.L.; Martienssen, R.A. Differential regulation of strand-specific transcripts from arabidopsis centromeric satellite repeats. *PLoS Genet.* **2005**, *1*, e79. [[CrossRef](#)] [[PubMed](#)]
74. Neumann, P.; Yan, H.; Jiang, J. The centromeric retrotransposons of rice are transcribed and differentially processed by RNA interference. *Genetics* **2007**, *176*, 749–761. [[CrossRef](#)] [[PubMed](#)]
75. Lee, H.R.; Neumann, P.; Macas, J.; Jiang, J. Transcription and evolutionary dynamics of the centromeric satellite repeat CentO in rice. *Mol. Biol. Evol.* **2006**, *23*, 2505–2520. [[CrossRef](#)]
76. Zakrzewski, F.; Weisshaar, B.; Fuchs, J.; Bannack, E.; Minoche, A.E.; Dohm, J.C.; Himmelbauer, H.; Schmidt, T. Epigenetic profiling of heterochromatic satellite DNA. *Chromosoma* **2011**, *120*, 409–422. [[CrossRef](#)]
77. Setiawan, A.B.; Teo, C.H.; Kikuchi, S.; Sassa, H.; Kato, K.; Koba, T. Centromeres of cucumis melo L. Comprise centromeric and two novel repeats, CmSat162 and CmSat189. *PLoS ONE* **2020**, *15*, e0227578. [[CrossRef](#)]
78. Li, F.; Sonbuchner, L.; Kyes, S.A.; Epp, C.; Deitsch, K.W. Nuclear non-coding RNAs are transcribed from the centromeres of plasmodium falciparum and are associated with centromeric chromatin. *J. Biol. Chem.* **2008**, *283*, 5692–5698. [[CrossRef](#)]
79. Jain, M.; Koren, S.; Miga, K.H.; Quick, J.; Rand, A.C.; Sasani, T.A.; Tyson, J.R.; Beggs, A.D.; Dilthey, A.T.; Fiddes, I.T.; et al. Nanopore sequencing and assembly of a human genome with ultra-long reads. *Nat. Biotechnol.* **2018**, *36*, 338–345. [[CrossRef](#)]
80. Quénet, D.; Sturgill, D.; Olson, M.; Dalal, Y. CENP-A associated lncRNAs influence chromosome segregation in human cells. *BioRxiv* **2018**, 097956. [[CrossRef](#)]
81. Ohkuni, K.; Kitagawa, K. Role of transcription at centromeres in budding yeast. *Transcription* **2012**, *3*, 193–197. [[CrossRef](#)] [[PubMed](#)]
82. Gassmann, R.; Rechtsteiner, A.; Yuen, K.W.; Muroyama, A.; Egelhofer, T.; Gaydos, L.; Barron, F.; Maddox, P.; Essex, A.; Monen, J.; et al. An inverse relationship to germline transcription defines centromeric chromatin in *C. elegans*. *Nature* **2012**, *484*, 534–537. [[CrossRef](#)] [[PubMed](#)]
83. Senaratne, A.P.; Muller, H.; Fryer, K.A.; Drinnenberg, I.A. The molecular architecture of CenH3-deficient holocentromeres in Lepidoptera is dependent on transcriptional and chromatin dynamics. *BioRxiv* **2020**. [[CrossRef](#)]
84. Tong, P.; Pidoux, A.L.; Toda, N.R.T.; Ard, R.; Berger, H.; Shukla, M.; Garcia, J.T.; Müller, C.A.; Nieduszynski, C.A.; Allshire, R.C. Interspecies conservation of organisation and function between nonhomologous regional centromeres. *Nat. Commun.* **2019**, *10*, 2343. [[CrossRef](#)]

85. Alper, B.J.; Lowe, B.R.; Partridge, J.F. Centromeric heterochromatin assembly in fission yeast—balancing transcription, RNA interference and chromatin modification. *Chromosom. Res.* **2012**, *20*, 521–534. [[CrossRef](#)]
86. Allshire, R.C.; Ekwall, K. Epigenetic regulation of chromatin states in *Schizosaccharomyces pombe*. *Cold Spring Harb. Perspect. Biol.* **2015**, *7*, a018770. [[CrossRef](#)]
87. Djupedal, I.; Braun, I.C.K.; Mosher, R.A.; Söderholm, N.; Simmer, F.; Hardcastle, T.J.; Fender, A.; Heidrich, N.; Kagansky, A.; Bayne, E.; et al. Analysis of small RNA in fission yeast; centromeric siRNAs are potentially generated through a structured RNA. *EMBO J.* **2009**, *28*, 3832–3844. [[CrossRef](#)]
88. Houseley, J.; Tollervey, D. The many pathways of RNA degradation. *Cell* **2009**, *136*, 763–776. [[CrossRef](#)]
89. Bresson, S.; Tollervey, D. Surveillance-ready transcription: Nuclear RNA decay as a default fate. *Open Biol.* **2018**, *8*, 170270. [[CrossRef](#)]
90. Nagaki, K.; Cheng, Z.; Ouyang, S.; Talbert, P.B.; Kim, M.; Jones, K.M.; Henikoff, S.; Buell, C.R.; Jiang, J. Sequencing of a rice centromere uncovers active genes. *Nat. Genet.* **2004**, *36*, 138–145. [[CrossRef](#)]
91. Wilkinson, M.E.; Charenton, C.; Nagai, K. RNA Splicing by the Spliceosome. *Annu. Rev. Biochem.* **2020**, *89*, 359–388. [[CrossRef](#)] [[PubMed](#)]
92. Thillainadesan, G.; Xiao, H.; Holla, S.; Dhakshnamoorthy, J.; Jenkins, L.M.M.; Wheeler, D.; Grewal, S.I.S. Conserved protein Pir2ARS2 mediates gene repression through cryptic introns in lncRNAs. *Nat. Commun.* **2020**, *11*, 1–11. [[CrossRef](#)] [[PubMed](#)]
93. Vijayakumari, D.; Sharma, A.K.; Bawa, P.S.; Kumar, R.; Srinivasan, S.; Vijayraghavan, U. Early splicing functions of fission yeast Prp16 and its unexpected requirement for gene silencing is governed by intronic features. *RNA Biol.* **2019**, *16*, 754–769. [[CrossRef](#)] [[PubMed](#)]
94. Brown, R.A.M.; Epis, M.R.; Horsham, J.L.; Kabir, T.D.; Richardson, K.L.; Leedman, P.J. Total RNA extraction from tissues for microRNA and target gene expression analysis: Not all kits are created equal. *BMC Biotechnol.* **2018**, *18*, 16. [[CrossRef](#)] [[PubMed](#)]
95. Chang, C.H.; Chavan, A.; Palladino, J.; Wei, X.; Martins, N.M.C.; Santinello, B.; Chen, C.C.; Erceg, J.; Beliveau, B.J.; Wu, C.T.; et al. Islands of retroelements are major components of drosophila centromeres. *PLoS Biol.* **2019**, *17*, e3000241. [[CrossRef](#)] [[PubMed](#)]
96. Bracewell, R.; Chatla, K.; Nalley, M.J.; Bachtrog, D. Dynamic turnover of centromeres drives karyotype evolution in drosophila. *Elife* **2019**, *8*, e49002. [[CrossRef](#)] [[PubMed](#)]
97. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914. [[CrossRef](#)]
98. Hansen, A.S.; Hsieh, T.H.S.; Cattoglio, C.; Pustova, I.; Meyer, R.S.; Reinberg, D.; Darzacq, X.; Tjian, R. Distinct classes of chromatin loops revealed by deletion of an RNA-binding region in CTCF. *Mol. Cell* **2019**, *76*, 395–411. [[CrossRef](#)]
99. Zhang, Q.; McKenzie, N.J.; Thomson, R.W.; Gail, E.H.; Flanigan, S.F.; Owen, B.M.; Lauman, R.; Levina, V.; Garcia, B.A.; Schittenhelm, R.B.; et al. RNA exploits an exposed regulatory site to inhibit the enzymatic activity of PRC2. *Nat. Struct. Mol. Biol.* **2019**, *26*, 237–247. [[CrossRef](#)]
100. Jansen, L.E.T.; Black, B.E.; Foltz, D.R.; Cleveland, D.W. Propagation of centromeric chromatin requires exit from mitosis. *J. Cell Biol.* **2007**, *176*, 795–805. [[CrossRef](#)]
101. Schuh, M.; Lehner, C.F.; Heidmann, S. Incorporation of drosophila CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. *Curr. Biol.* **2007**, *17*, 237–243. [[CrossRef](#)] [[PubMed](#)]
102. Lermontova, I.; Schubert, V.; Fuchs, J.; Klatt, S.; Macas, J.; Schubert, I. Loading of arabidopsis centromeric histone CENH3 occurs mainly during G2 and requires the presence of the histone fold domain. *Plant Cell* **2006**, *18*, 2443–2451. [[CrossRef](#)] [[PubMed](#)]
103. Gonzalez, M.; He, H.; Sun, S.; Li, C.; Li, F. Cell cycle-dependent deposition of CENP-A requires the Dos1/2-Cdc20 complex. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 606–611. [[CrossRef](#)] [[PubMed](#)]
104. Choi, E.S.; Strålfors, A.; Catania, S.; Castillo, A.G.; Svensson, J.P.; Pidoux, A.L.; Ekwall, K.; Allshire, R.C. Factors that promote H3 chromatin integrity during transcription prevent promiscuous deposition of CENP-ACnp1 in fission yeast. *PLoS Genet.* **2012**, *8*, e1002985. [[CrossRef](#)] [[PubMed](#)]
105. Bobkov, G.O.M.; Huang, A.; Berg, S.J.W.V.D.; Mitra, S.; Anselm, E.; Lazou, V.; Schunter, S.; Feederle, R.; Imhof, A.; Lusser, A.; et al. Spt6 is a maintenance factor for centromeric CENP-A. *Nat. Commun.* **2020**, *11*, 1–14. [[CrossRef](#)]
106. Dunleavy, E.M.; Almouzni, G.; Karpen, G.H. H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G phase. *Nucleus* **2011**, *2*, 146–157. [[CrossRef](#)]

107. Dunleavy, E.M.; Roche, D.; Tagami, H.; Lacoste, N.; Gallet, D.R.; Nakamura, Y.; Daigo, Y.; Nakatani, Y.; Pettinitti, G.A. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* **2009**, *137*, 485–497. [[CrossRef](#)]
108. Foltz, D.R.; Jansen, L.E.T.; Bailey, A.O.; Iii, J.R.Y.; Bassett, E.A.; Wood, S.; Black, B.E.; Cleveland, D.W. Centromere-specific assembly of CENP-A nucleosomes is mediated by HJURP. *Cell* **2009**, *137*, 472–484. [[CrossRef](#)]
109. Furuyama, T.; Dalal, Y.; Henikoff, S. Chaperone-mediated assembly of centromeric chromatin in vitro. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 6172–6177. [[CrossRef](#)]
110. Chen, C.C.; Dechassa, M.L.; Bettini, E.; Ledoux, M.B.; Belisario, C.; Heun, P.; Luger, K.; Mellone, B.G. CAL1 is the drosophila CENP-A assembly factor. *J. Cell Biol.* **2014**, *204*, 313–329. [[CrossRef](#)]
111. Camahort, R.; Li, B.; Florens, L.; Swanson, S.K.; Washburn, M.P.; Gerton, J.L. Scm3 is essential to recruit the histone H3 variant Cse4 to centromeres and to maintain a functional kinetochore. *Mol. Cell* **2007**, *26*, 853–865. [[CrossRef](#)] [[PubMed](#)]
112. Stoler, S.; Rogers, K.; Weitze, S.; Morey, L.; Hayes, M.F.; Baker, R.E. Scm3, an essential *Saccharomyces cerevisiae* centromere protein required for G2/M progression and Cse4 localization. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10571–10576. [[CrossRef](#)] [[PubMed](#)]
113. Mizuguchi, G.; Xiao, H.; Wisniewski, J.; Smith, M.M.; Wu, C. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell* **2007**, *129*, 1153–1164. [[CrossRef](#)] [[PubMed](#)]
114. Williams, J.S.; Hayashi, T.; Yanagida, M.; Russell, P. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Mol. Cell* **2009**, *33*, 287–298. [[CrossRef](#)] [[PubMed](#)]
115. Lee, B.C.H.; Lin, Z.; Yuen, K.W.Y. RbAp46/48LIN-53 is required for holocentromere assembly in *Caenorhabditis elegans*. *Cell Rep.* **2016**, *14*, 1819–1828. [[CrossRef](#)] [[PubMed](#)]
116. Goff, S.L.; Keçeli, B.N.; Jeřábková, H.; Heckmann, S.; Rutten, T.; Cotterell, S.; Schubert, V.; Roitinger, E.; Mechtler, K.; Franklin, F.C.H.; et al. The H3 histone chaperone NASPSIM3 escorts CenH3 in *Arabidopsis*. *Plant. J.* **2020**, *101*, 71–86. [[CrossRef](#)]
117. Melters, D.P.; Pitman, M.; Rakshit, T.; Dimitriadis, E.K.; Bui, M.; Papoian, G.A.; Dalal, Y. Intrinsic elasticity of nucleosomes is encoded by histone variants and calibrated by their binding partners. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 24066–24074. [[CrossRef](#)]
118. Politi, V.; Perini, G.; Trazzi, S.; Pliss, A.; Raska, I.; Earnshaw, W.C.; Valle, G.D. CENP-C binds the alpha-satellite DNA in vivo at specific centromere domains. *J. Cell Sci.* **2002**, *115*, 2317–2327.
119. Xiao, H.; Wang, F.; Wisniewski, J.; Shaytan, A.K.; Ghirlando, R.; Fitzgerald, P.C.; Huang, Y.; Wei, D.; Li, S.; Landsman, D.; et al. Molecular basis of CENP-C association with the CENP-A nucleosome at yeast centromeres. *Genes Dev.* **2017**, *31*, 1958–1972. [[CrossRef](#)]
120. Maldonado, R.; Schwartz, U.; Silberhorn, E.; Längst, G. Nucleosomes stabilize ssRNA-dsDNA triple helices in human cells. *Mol. Cell* **2019**, *73*, 1243–1254. [[CrossRef](#)]
121. Masumoto, H.; Masukata, H.; Muro, Y.; Nozaki, N.; Okazaki, T. A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J. Cell Biol.* **1989**, *109*, 1963–1973. [[CrossRef](#)] [[PubMed](#)]
122. Fachinetti, D.; Han, J.S.; McMahon, M.A.; Ly, P.; Abdullah, A.; Wong, A.J.; Cleveland, D.W. DNA sequence-specific binding of CENP-B enhances the fidelity of human centromere function. *Dev. Cell* **2015**, *33*, 314–327. [[CrossRef](#)] [[PubMed](#)]
123. Yoda, K.; Ando, S.; Okuda, A.; Kikuchi, A.; Okazaki, T. In vitro assembly of the CENP-B/ α -satellite DNA/core histone complex: CENP-B causes nucleosome positioning. *Genes Cells* **1998**, *3*, 533–548. [[CrossRef](#)] [[PubMed](#)]
124. Ohzeki, J.I.; Nakano, M.; Okada, T.; Masumoto, H. CENP-B box is required for de novo centromere chromatin assembly on human alphoid DNA. *J. Cell Biol.* **2002**, *159*, 765–775. [[CrossRef](#)] [[PubMed](#)]
125. Fujita, R.; Otake, K.; Arimura, Y.; Horikoshi, N.; Miya, Y.; Shiga, T.; Osakabe, A.; Tachiwana, H.; Ohzeki, J.I.; Larionov, V.; et al. Stable complex formation of CENP-B with the CENP-A nucleosome. *Nucleic Acids Res.* **2015**, *43*, 4909–4922. [[CrossRef](#)]
126. Ikeno, M.; Grimes, B.; Okazaki, T.; Nakano, M.; Saitoh, K.; Hoshino, H.; McGill, N.I.; Cooke, H.; Masumoto, H. Construction of YAC-based mammalian artificial chromosomes. *Nat. Biotechnol.* **1998**, *16*, 431–439. [[CrossRef](#)]
127. Harrington, J.J.; Bokkelen, G.V.; Mays, R.W.; Gustashaw, K.; Willard, H.F. Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat. Genet.* **1997**, *15*, 345–355. [[CrossRef](#)]

128. Hoffmann, S.; Dumont, M.; Barra, V.; Ly, P.; Arbely, Y.N.; McMahon, M.A.; Hervé, S.; Cleveland, D.W.; Fachinetti, D. CENP-A is dispensable for mitotic centromere function after initial centromere/kinetochore assembly. *Cell Rep.* **2016**, *17*, 2394–2404. [[CrossRef](#)]
129. Morozov, V.M.; Giovanazzi, S.; Ishov, A.M. CENP-B protects centromere chromatin integrity by facilitating histone deposition via the H3.3-specific chaperone Daxx. *Epigenetics Chromatin* **2017**, *10*, 63. [[CrossRef](#)]
130. Otake, K.; Ohzeki, J.; Shono, N.; Kugou, K.; Okazaki, K.; Nagase, T.; Yamakawa, H.; Kouprina, N.; Larionov, V.; Kimura, H.; et al. CENP-B creates alternative epigenetic chromatin states permissive for CENP-A or heterochromatin assembly. *J. Cell Sci.* **2020**, 243303. [[CrossRef](#)]
131. Tawaramoto, M.S.; Park, S.; Tanaka, Y.; Nureki, O.; Kurumizaka, H.; Yokoyama, S. Crystal structure of the human centromere protein B (CENP-B) dimerization domain at 1.65-Å resolution *. *Biochemistry* **2003**, *278*, 51454–51461.
132. Li, Y.; Syed, J.; Sugiyama, H. RNA-DNA triplex formation by long noncoding RNAs. *Cell Chem. Biol.* **2016**, *23*, 1325–1333. [[CrossRef](#)] [[PubMed](#)]
133. Kabeche, L.; Nguyen, H.D.; Buisson, R.; Zou, L. A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science* **2018**, *359*, 108–114. [[CrossRef](#)] [[PubMed](#)]
134. Trivedi, P.; Stukenberg, P.T. A centromere-signaling network underlies the coordination among mitotic events. *Trends Biochem. Sci.* **2016**, *41*, 160–174. [[CrossRef](#)] [[PubMed](#)]
135. Bloom, K.S. Centromeric heterochromatin: The primordial segregation machine. *Annu. Rev. Genet.* **2014**, *48*, 457–484. [[CrossRef](#)] [[PubMed](#)]
136. Abdelkrim, J.; Robertson, B.C.; Stanton, J.L.; Gemmell, N.J. Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *Biotechniques* **2009**, *46*, 185–192. [[CrossRef](#)] [[PubMed](#)]
137. Trivedi, P.; Palomba, F.; Niedzialkowska, E.; Digman, M.A.; Gratton, E.; Stukenberg, P.T. The inner centromere is a biomolecular condensate scaffolded by the chromosomal passenger complex. *Nat. Cell Biol.* **2019**, *21*, 1127–1137. [[CrossRef](#)]
138. McGuinness, B.E.; Hirota, T.; Kudo, N.R.; Peters, J.M.; Nasmyth, K. Shugoshin prevents dissociation of cohesin from centromeres during mitosis in vertebrate cells. *PLoS Biol.* **2005**, *3*, e86. [[CrossRef](#)]
139. Skrajna, A.; Goldfarb, D.; Kedziora, K.M.; Cousins, E.M.; Grant, G.D.; Spangler, C.J.; Barbour, E.H.; Yan, X.; Hathaway, N.A.; Brown, N.G.; et al. Comprehensive nucleosome interactome screen establishes fundamental principles of nucleosome binding. *Nucleic Acids Res.* **2020**. [[CrossRef](#)]
140. Kalashnikova, A.A.; Goff, M.E.P.; Muthurajan, U.M.; Luger, K.; Hansen, J.C. The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J. R. Soc. Interface* **2013**, *10*, 20121022. [[CrossRef](#)]
141. Burton, A.; Brochard, V.; Galan, C.; Morales, E.R.R.; Rovira, Q.; Terrones, D.R.; Kruse, K.; Gras, S.L.; Udayakumar, V.S.; Chin, H.G.; et al. Heterochromatin establishment during early mammalian development is regulated by pericentromeric RNA and characterized by non-repressive H3K9me3. *Nat. Cell Biol.* **2020**, *22*, 767–778. [[CrossRef](#)] [[PubMed](#)]
142. Olszak, A.M.; Essen, D.V.; Pereira, A.J.; Diehl, S.; Manke, T.; Maiato, H.; Sacconi, S.; Heun, P. Heterochromatin boundaries are hotspots for de novo kinetochore formation. *Nat. Cell Biol.* **2011**, *13*, 799–808. [[CrossRef](#)] [[PubMed](#)]
143. Au, W.C.; Zhang, T.; Mishra, P.K.; Eisenstatt, J.R.; Walker, R.L.; Ocampo, J.; Dawson, A.; Warren, J.; Costanzo, M.; Baryshnikova, A.; et al. Skp, cullin, F-box (SCF)-Met30 and SCF-Cdc4- mediated proteolysis of CENP-A prevents mislocalization of CENP-A for chromosomal stability in budding yeast. *PLoS Genet.* **2020**, *16*, e1008597. [[CrossRef](#)] [[PubMed](#)]
144. Eisenstatt, J.R.; Boeckmann, L.; Au, W.C.; Garcia, V.; Bursch, L.; Ocampo, J.; Costanzo, M.; Weinreich, M.; Sclafani, R.A.; Baryshnikova, A.; et al. Dbf4-dependent kinase (DDK)-mediated proteolysis of CENP-A prevents mislocalization of CENP-A in *Saccharomyces cerevisiae*. *G3 Genes Genomes Genet.* **2020**, *10*, 2057–2068. [[CrossRef](#)]
145. Yilmaz, S.C.; Au, W.C.; Mishra, P.K.; Eisenstatt, J.R.; Chang, J.; Dawson, A.R.; Zhu, I.; Rahman, M.; Bilke, S.; Costanzo, M.; et al. A genome-wide screen reveals a role for the HIR histone chaperone complex in preventing mislocalization of budding yeast CENP-A. *Genetics* **2018**, *210*, 203–218. [[CrossRef](#)]
146. Ohkuni, K.; Myers, R.L.; Warren, J.; Au, W.C.; Takahashi, Y.; Baker, R.E.; Basrai, M.A. N-terminal sumoylation of centromeric histone H3 variant Cse4 regulates its proteolysis to prevent mislocalization to non-centromeric chromatin. *G3 Genes Genomes Genet.* **2018**, *8*, 1215–1223. [[CrossRef](#)]

147. Ohkuni, K.; Takahashi, Y.; Fulp, A.; Lawrimore, J.; Au, W.C.; Pasupala, N.; Myers, R.L.; Warren, J.; Strunnikov, A.; Baker, R.E.; et al. SUMO-targeted ubiquitin ligase (STUbL) Slx5 regulates proteolysis of centromeric histone H3 variant Cse4 and prevents its mislocalization to euchromatin. *Mol. Biol. Cell* **2016**, *27*, 1500–1510. [[CrossRef](#)]
148. Zhang, W.; Mao, J.H.; Zhu, W.; Jain, A.K.; Liu, K.; Brown, J.B.; Karpen, G.H. Centromere and kinetochore gene misexpression predicts cancer patient survival and response to radiotherapy and chemotherapy. *Nat. Commun.* **2016**, *7*, 12619. [[CrossRef](#)]
149. Athwal, R.K.; Walkiewicz, M.P.; Baek, S.; Fu, S.; Bui, M.; Camps, J.; Ried, T.; Sung, M.-H.; Dalal, Y. CENP-A nucleosomes localize to transcription factor hotspots and subtelomeric sites in human cancer cells. *Epigenetics Chromatin* **2015**, *8*, 2. [[CrossRef](#)]
150. Nye, J.; Sturgill, D.; Athwal, R.; Dalal, Y. HJURP antagonizes CENP-A mislocalization driven by the H3.3 chaperones HIRA and DAXX. *PLoS ONE* **2018**, *13*, e0205948. [[CrossRef](#)]
151. Lacoste, N.; Woolfe, A.; Tachiwana, H.; Garea, A.V.; Barth, T.; Cantaloube, S.; Kurumizaka, H.; Imhof, A.; Almouzni, G. Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. *Mol. Cell* **2014**, *53*, 631–644. [[CrossRef](#)] [[PubMed](#)]
152. Shrestha, R.L.; Ahn, G.S.; Staples, M.I.; Sathyan, K.M.; Karpova, T.S.; Foltz, D.R.; Basrai, M.A. Mislocalization of centromeric histone H3 variant CENP-A contributes to chromosomal instability (CIN) in human cells. *Oncotarget* **2017**, *8*, 46781–46800. [[CrossRef](#)] [[PubMed](#)]
153. Stimpson, K.M.; Matheny, J.E.; Sullivan, B.A. Dicentric chromosomes: Unique models to study centromere function and inactivation. *Chromosom. Res.* **2012**, *20*, 595–605. [[CrossRef](#)] [[PubMed](#)]
154. Gonzalez, M.; He, H.; Dong, Q.; Sun, S.; Li, F. Ectopic centromere nucleation by CENP-A in fission yeast. *Genetics* **2014**, *198*, 1433–1446. [[CrossRef](#)]
155. Collins, K.A.; Camahort, R.; Seidel, C.; Gerton, J.L.; Biggins, S. The overexpression of a *Saccharomyces cerevisiae* centromeric histone H3 variant mutant protein leads to a defect in kinetochore biorientation. *Genetics* **2007**, *198*, 1433–1446. [[CrossRef](#)]
156. Hildebrand, E.M.; Biggins, S. Regulation of budding yeast CENP-A levels prevents misincorporation at promoter nucleosomes and transcriptional defects. *PLoS Genet.* **2016**, *12*, e1005930. [[CrossRef](#)]
157. Schotanus, K.; Heitman, J. Centromere deletion in *Cryptococcus deuterogattii* leads to neocentromere formation and chromosome fusions. *Elife* **2020**, *9*, e56026. [[CrossRef](#)]
158. Furuyama, S.; Biggins, S. Centromere identity is specified by a single centromeric nucleosome in budding yeast. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 14706–14711. [[CrossRef](#)]
159. Mérai, Z.; Chumak, N.; Aguilar, M.G.; Hsieh, T.F.; Nishimura, T.; Schoft, V.K.; Bindics, J.; Ślusarz, L.; Arnoux, S.; Opravil, S.; et al. The AAA-ATPase molecular chaperone Cdc48/p97 disassembles sumoylated centromeres, decondenses heterochromatin, and activates ribosomal RNA genes. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 16166–16171. [[CrossRef](#)]
160. Arbely, Y.N.; Miga, K.H.; Shoshani, O.; Aslanian, A.; McMahon, M.A.; Lee, A.Y.; Fachinetti, D.; Yates, J.R.; Ren, B.; Cleveland, D.W. DNA replication acts as an error correction mechanism to maintain centromere identity by restricting CENP-A to centromeres. *Nat. Cell Biol.* **2019**, *21*, 743–754. [[CrossRef](#)]
161. Saha, A.K.; Galindo, R.C.; Niknafs, Y.S.; Iyer, M.; Qin, T.; Padmanabhan, K.; Siddiqui, J.; Palande, M.; Wang, C.; Qian, B.; et al. The role of the histone H3 variant CENPA in prostate cancer. *J. Biol. Chem.* **2020**, *295*, 8537. [[CrossRef](#)] [[PubMed](#)]
162. Martins, N.M.C.; Soberanis, F.C.; Pesenti, E.; Kochanova, N.Y.; Shang, W.H.; Hori, T.; Nagase, T.; Kimura, H.; Larionov, V.; Masumoto, H.; et al. H3K9me3 maintenance on a human artificial chromosome is required for segregation but not centromere epigenetic memory. *J. Cell Sci.* **2020**, *133*, jcs242610. [[CrossRef](#)] [[PubMed](#)]
163. Molina, O.; Vargiu, G.; Abad, M.A.; Zhiteneva, A.; Jeyaprakash, A.A.; Masumoto, H.; Kouprina, N.; Larionov, V.; Earnshaw, W.C. Epigenetic engineering reveals a balance between histone modifications and transcription in kinetochore maintenance. *Nat. Commun.* **2016**, *7*, 13334. [[CrossRef](#)] [[PubMed](#)]
164. Martins, N.M.C.; Bergmann, J.H.; Shono, N.; Kimura, H.; Larionov, V.; Masumoto, H.; Earnshaw, W.C. Epigenetic engineering shows that a human centromere resists silencing mediated by H3K27me3/K9me3. *Mol. Biol. Cell* **2016**, *27*, 177–196. [[CrossRef](#)] [[PubMed](#)]
165. Chueh, A.C.; Northrop, E.L.; Moore, K.H.B.; Choo, K.H.A.; Wong, L.H. LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. *PLoS Genet.* **2009**, *5*, e1000354. [[CrossRef](#)]

166. Ting, D.T.; Lipson, D.; Paul, S.; Brannigan, B.W.; Akhavanfard, S.; Coffman, E.J.; Contino, G.; Deshpande, V.; Iafrate, A.J.; Letovsky, S.; et al. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **2011**, *331*, 593–596. [[CrossRef](#)]
167. McNulty, S.M.; Sullivan, B.A. Alpha satellite DNA biology: Finding function in the recesses of the genome. *Chromosom. Res.* **2018**, *26*, 115–138. [[CrossRef](#)]
168. Chan, D.Y.L.; Moralli, D.; Khoja, S.; Monaco, Z.L. Noncoding centromeric RNA expression impairs chromosome stability in human and murine stem cells. *Dis. Markers* **2017**, *2017*, 1–8. [[CrossRef](#)]
169. Negrini, S.; Gorgoulis, V.G.; Halazonetis, T.D. Genomic instability an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 220–228. [[CrossRef](#)]
170. Ichida, K.; Suzuki, K.; Fukui, T.; Takayama, Y.; Kakizawa, N.; Watanabe, F.; Ishikawa, H.; Muto, Y.; Kato, T.; Saito, M.; et al. Overexpression of satellite alpha transcripts leads to chromosomal instability via segregation errors at specific chromosomes. *Int. J. Oncol.* **2018**, *52*, 1685–1693. [[CrossRef](#)]
171. Zhu, Q.; Pao, G.M.; Huynh, A.M.; Suh, H.; Tonnu, N.; Nederlof, P.M.; Gage, F.H.; Verma, I.M. BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* **2011**, *477*, 179–184. [[CrossRef](#)] [[PubMed](#)]
172. Zhu, Q.; Hoong, N.; Aslanian, A.; Hara, T.; Benner, C.; Heinz, S.; Miga, K.H.; Ke, E.; Verma, S.; Soroczynski, J.; et al. Heterochromatin-encoded satellite RNAs induce breast cancer. *Mol. Cell* **2018**, *70*, 842–853. [[CrossRef](#)] [[PubMed](#)]
173. Jolly, C.; Metz, A.; Govin, J.; Vigneron, M.; Turner, B.M.; Khochbin, S.; Vourc'h, C. Stress-induced transcription of satellite III repeats. *J. Cell Biol.* **2004**, *164*, 25–33. [[CrossRef](#)] [[PubMed](#)]
174. Valgardsdottir, R.; Chiodi, I.; Giordano, M.; Rossi, A.; Bazzini, S.; Ghigna, C.; Riva, S.; Biamonti, G. Transcription of satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Res.* **2008**, *36*, 423–434. [[CrossRef](#)] [[PubMed](#)]
175. Hall, L.L.; Byron, M.; Carone, D.M.; Whitfield, T.W.; Pouliot, G.P.; Fischer, A.; Jones, P.; Lawrence, J.B. Demethylated HSATII DNA and HSATII RNA foci sequester PRC1 and MeCP2 into cancer-specific nuclear bodies. *Cell Rep.* **2017**, *18*, 2943–2956. [[CrossRef](#)]
176. Bersani, F.; Lee, E.; Kharchenko, P.V.; Xu, A.W.; Liu, M.; Xega, K.; MacKenzie, O.C.; Brannigan, B.W.; Wittner, B.S.; Jung, H.; et al. Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 15148–15153. [[CrossRef](#)]
177. Galindo, R.C.; Kaplan, M.H.; He, S.; Galindo, A.C.C.; Hernandez, M.J.G.; Kappes, F.; Dube, D.; Chan, S.M.; Robinson, D.; Meng, F.; et al. HIV infection reveals widespread expansion of novel centromeric human endogenous retroviruses. *Genome Res.* **2013**, *23*, 1505–1513. [[CrossRef](#)]
178. Okita, A.K.; Zafar, F.; Su, J.; Weerasekara, D.; Kajitani, T.; Takahashi, T.S.; Kimura, H.; Murakami, Y.; Masukata, H.; Nakagawa, T. Heterochromatin suppresses gross chromosomal rearrangements at centromeres by repressing Tfs1/TFIIS-dependent transcription. *Commun. Biol.* **2019**, *2*, 17. [[CrossRef](#)]
179. Park, J.; Lee, H.; Han, N.; Kwak, S.; Lee, H.T.; Kim, J.H.; Kang, K.; Youn, B.H.; Yang, J.H.; Jeong, H.J.; et al. Long non-coding RNA ChRO1 facilitates ATRX/DAXX-dependent H3.3 deposition for transcription-associated heterochromatin reorganization. *Nucleic Acids Res.* **2018**, *46*, 11759–11775. [[CrossRef](#)]
180. Holla, S.; Dhakshnamoorthy, J.; Folco, H.D.; Balachandran, V.; Xiao, H.; Sun, L.L.; Wheeler, D.; Zofall, M.; Grewal, S.I.S. Positioning heterochromatin at the nuclear periphery suppresses histone turnover to promote epigenetic inheritance. *Cell* **2020**, *180*, 150–164. [[CrossRef](#)]
181. Levin, Z.A.G.; Quivy, J.-P.; Almouzni, G. Histone chaperones: Assisting histone traffic and nucleosome dynamics. *Annu. Rev. Biochem.* **2014**, *83*, 487–517. [[CrossRef](#)] [[PubMed](#)]
182. Pitman, M.; Melters, D.P.; Dalal, Y. Job opening for nucleosome mechanic: Flexibility required. *Cells* **2020**, *9*, 580. [[CrossRef](#)] [[PubMed](#)]
183. Drinnenberg, I.A.; deYoung, D.; Henikoff, S.; Malik, H.S. Recurrent loss of CenH3 is associated with independent transitions to holocentricity in insects. *Elife* **2014**, *3*, e03676. [[CrossRef](#)] [[PubMed](#)]
184. Roure, V.; Pritchard, B.M.; Lazou, V.; Rago, L.; Anselm, E.; Venegas, D.; Jeyaprakash, A.A.; Heun, P. Reconstituting drosophila centromere identity in human cells. *Cell Rep.* **2019**, *29*, 464–479. [[CrossRef](#)] [[PubMed](#)]
185. Tsukahara, S.; Kawabe, A.; Kobayashi, A.; Ito, T.; Aizu, T.; Shini, T.; Toyoda, A.; Fujiyama, A.; Tarutani, Y.; Kakutani, T. Centromere-targeted de novo integrations of an LTR retrotransposon of arabidopsis lyrata. *Genes Dev.* **2012**, *26*, 705–713. [[CrossRef](#)] [[PubMed](#)]

186. Miga, K.H.; Koren, S.; Rhie, A.; Vollger, M.R.; Gershman, A.; Bzikadze, A.; Brooks, S.; Howe, E.; Porubsky, D.; Logsdon, G.A.; et al. Telomere-to-telomere assembly of a complete human X chromosome. *Nature* **2020**. [[CrossRef](#)]
187. Bzikadze, A.V.; Pevzner, P.A. Automated assembly of centromeres from ultra-long error-prone reads. *Nat. Biotechnol.* **2020**. [[CrossRef](#)]



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).