

Review

Transcribing Centromeres: Noncoding RNAs and Kinetochore Assembly

Paul B. Talbert^{1,*} and Steven Henikoff¹

Chromosome segregation depends on the attachment of spindle microtubules to sites on chromosomal DNA known as centromeres, through kinetochore protein complexes. Although RNA was found in kinetochores in the 1970s, only with recent investigations has evidence emerged that loading of the centromere-specific nucleosomes that form the foundation of the kinetochore may be coupled to centromeric transcription. Centromeric transcripts are bound by several kinetochore proteins that require them for stabilization or localization. At least some centromeres have promoter activity, and many have non-B form DNA that may facilitate their transcription. Whereas other noncoding RNAs regulate gene expression or silence transposons, cotranscriptional assembly of kinetochores is a novel function for noncoding RNAs.

Centromeric RNA Revisited

Chromosome segregation in mitosis and meiosis depends on the attachment of microtubules to the chromosomes at sites called **centromeres** (see [Glossary](#)) by means of protein complexes known as **kinetochores**. RNA was first observed at kinetochores in both plants and animals in electron microscopy studies in the 1970s [1,2]. Despite these early reports, centromeres came to be viewed as transcriptionally silent because they are commonly embedded in heterochromatin, and because transcription from a strong promoter next to a budding yeast centromere interfered with its function [3]. Additionally, evolutionarily new centromeres and human **neocentromeres** were found to occur in 'gene deserts' [4]. However, recent reports of noncoding centromeric RNAs both challenge the heterochromatic centromere model and suggest novel functions for these RNAs. In contrast with the roles of many other noncoding RNAs that regulate gene expression [5], centromeric transcripts affect the stability or activity of several kinetochore components, and centromeric transcription may be required for loading centromeric nucleosomes.

Transcription of Centromeres

A defining feature of most centromeres is the presence of nucleosomes containing a centromeric variant of histone H3 in place of canonical H3. Centromeric H3 variants (**cenH3s**) have different names in different organisms: **CENP-A** in animals and fission yeast, Cse4 in budding yeast, and cenH3 in plants and many protists. The perception that centromeres are heterochromatic was challenged when animal centromeres were found to comprise ~15–40 kb blocks or subdomains of CENP-A nucleosomes, alternating with subdomains of H3 nucleosomes [6] bearing the H3K4me2 mark of active transcription [7]. Similarly, the centromere region from rice *cen8* was found to contain active genes [8]. More precise mapping of *cen8* revealed that most of the genes were in H3 subdomains within the centromere region, while the cenH3 subdomains had only very few active genes and pseudogenes [9]. This suggests that gene transcription is generally, but not entirely, incompatible with centromere function.

Highlights

Centromeres are transcribed at a low level and transcripts are incorporated into centromeric chromatin, where they serve essential functions.

Several kinetochore proteins bind centromeric transcripts, which may be necessary to stabilize or localize the proteins.

Loading of centromere-specific nucleosomes may be coupled to centromeric transcription.

Some centromeres have known promoter activity and most centromeres are enriched in non-B form DNA that may facilitate transcription or loading of centromere-specific nucleosomes.

¹Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Seattle, WA 98109, USA

*Correspondence: ptalbert@fredhutch.org (P.B. Talbert).

In a ground-breaking study in maize [10], Topp *et al.* found that 40–900 nt transcripts from both strands of the centromeric **satellite** CentC and the centromeric retrotransposons (CRMs), but not from noncentromeric repeats, were tightly bound to cenH3-containing chromatin. The authors noted that centromeric transcription might be involved in kinetochore initiation by opening the chromatin to allow cenH3 deposition, and that centromeric transcripts might provide a flexible scaffold for targeting or stabilizing kinetochore proteins.

Centromeric transcripts were later found to be a part of centromeric chromatin in human cells [11], the malarial parasite *Plasmodium falciparum* [12], human neocentromeres [13], mouse cells [14], tammar wallaby cells [15], and fission yeast [16], suggesting that centromeric RNA is an ancient and widespread component of centromeric chromatin. Although transcription of centromeres appears to be general, centromere DNA sequences are not conserved, even between closely related species (e.g., [17]), and therefore centromeric RNAs are likewise not conserved.

Centromeres are transcribed by RNA Polymerase II (Pol II) [16,18–26], often on both strands [10,12,15,16,27–29]. The RNAs are capped [16] and are sensitive to RNase A and other single-stranded RNases, but generally not to DNases, RNaseH, or double-stranded RNase [1,10,11,26], although mouse minor RNA was sensitive to double-stranded RNase [30] and a minor component of RNase A-resistant RNA was observed in tammar wallaby [15]. Typically the RNAs are transcribed as longer precursor RNAs that may be polyadenylated [10] and are processed [15,25,30] to a variety of smaller sizes that are distinct from **small interfering RNAs (siRNAs)** [10,12,15,25,26,30], and can vary between centromeres in the same cell and even at the same centromere between cell lines or developmental stages [26,30]. Transcription and splicing factors, but not new translation, are required for proper assembly of the kinetochore and spindle [25].

Centromeric transcripts are present at a low level, and higher transcription levels are often detrimental. The overexpression of a 120-nt mouse minor satellite centromeric transcript caused mitotic arrest, condensation and cohesion defects, formation of **micronuclei**, and ectopic localization of Aurora-B kinase [30]. Similarly, high-level transcription of centromeric RNA is detrimental in budding yeast [3,19], and centromeric RNA is kept at a low level by the exosome in fission yeast [16]. The abundance of centromeric α -satellite transcripts in human cell lines was estimated to be about 0.5% that of a housekeeping gene [20,31]. In HeLa cells, no effect of overexpression was observed from one or seven monomers of α -satellite [27]. However, in human artificial chromosomes (HACs) containing *tetO* in α -satellite DNA, tetR-fused transcriptional activators and silencers both destabilized kinetochore formation [32]. The strong activator TetR-EYFP-VP16, which increased transcription of α -satellite RNA ~150-fold, caused an increase in Pol II occupancy, caused a loss of CENP-A probably through nucleosome eviction, and blocked new CENP-A assembly [31]. Mouse minor satellite transcripts increased through the cell cycle, from very low in G0/G1 to S to a peak in G2 [14], though in *Plasmodium* [12] and in human HeLa cells [26], centromeric RNA levels were found to be similar throughout the cell cycle.

RNA and CENP-C

In HeLa cells, centromeric α -satellite RNAs were observed in nucleoli, along with the foundational kinetochore protein **CENP-C** and INCENP, a component of the **chromosomal passenger complex (CPC)** [11]. Two regions of human CENP-C were shown to bind to α -satellite RNAs *in vitro*, with partial specificity, since CENP-C bound three different chromosome-specific centromeric RNAs but not tRNA or rRNA [11]. These same two regions were previously shown

Glossary

cenH3, CENP-A: the centromere-specific variant of histone H3 that is the interface between the centromere DNA and the protein complex that forms the kinetochore. It is usually required to assemble other kinetochore components.

CENP-C: a fundamental kinetochore protein that directly binds cenH3 nucleosomes, DNA, and RNA. It serves as an assembly platform for other kinetochore proteins.

CENPC motif: a conserved motif of ~23 amino acid in CENP-C that binds cenH3 nucleosomes. Many vertebrate CENP-Cs have a second variant of this motif (CENPC-v) in their central region, and most non-mammalian KNL2s also have a variant (CENPC-k) that binds cenH3.

Centromere: the chromosomal locus where spindle fibers attach via the kinetochore in order to segregate chromosomes.

Chromosomal passenger

complex (CPC): a mobile complex consisting of the Aurora-B kinase, INCENP, Borealin, and Survivin, which phosphorylates many proteins during mitosis, senses tension at the kinetochore, and corrects inappropriate spindle attachments. It also has roles in spindle formation and cytokinesis.

Dyad symmetry: a palindrome-like region of a DNA molecule with two adjacent or nearly adjacent sequences that are reverse complements, with the potential to form a cruciform structure.

Facilitates Transcription (FACT): a two-subunit complex that is thought to remodel nucleosomes in such a way that RNA or DNA polymerases can traverse them and they can still be retained on DNA.

Inner centromere: the region of two sister chromatids that is between their oppositely oriented kinetochores. It is usually made up of H3 nucleosomes and is rich in cohesins. The CPC relocates from the kinetochores to the inner centromere.

Kinetochore: the protein machine that attaches to spindle microtubules, organizes them on a metaphase plate, senses when biorientation of chromatids is achieved, and signals the onset of anaphase.

to bind DNA [33,34] and contain the central region (CENPC-v) and **CENPC motifs** that bind to CENP-A [35] (Figure 1A). Mutation of three lysines adjacent to the central CENP-A-binding CENPC-v motif abolished RNA binding in that region [11]. ChIP with anti-CENP-C antibody pulled down chromatin containing α -satellite RNAs, and RNase treatment of single-stranded RNAs during metaphase reduced CENP-C localization and eliminated centromeric localization of INCENP and its binding partner Survivin. The authors proposed a model in which CENP-C and INCENP are preassociated with α -satellite RNA in the nucleolus for effective assembly into the mitotic kinetochore. However, the nucleolar localization of centromere transcripts was not confirmed in other studies using HeLa and other human cells [26,27]. Nucleolar centromeric transcripts were also reported in maize [29], but not nucleolar CENP-C [36], leaving the role of a nucleolar assembly pathway unclear.

CENP-C was also found to bind RNA in maize (Figure 1B), although without sequence specificity [36]. The domains encoded by exons 9–12, just upstream of the exon encoding the CENPC motif, were found to bind both RNA and DNA, and binding of a small RNA enhanced DNA binding. Single-stranded DNA, like RNA, also promoted binding to double-stranded DNA. Exons 9–12 were necessary for CENP-C localization to centromeres *in vivo*, consistent with a model in which centromeric RNA stabilizes CENP-C by increasing its binding to DNA, adjacent to where it binds the cenH3 nucleosome.

This model predicts that disruption of centromeric RNA should destabilize CENP-C. Depletion of centromeric RNAs in HeLa cells reduced CENP-C [26], or not [18], in two studies targeting different specific RNAs and using different assays. Mitotic 14ZBHT cells inhibited for Pol II elongation with α -amanitin showed an increase in lagging chromosomes and reduction of CENP-C levels at the kinetochore, with a greater reduction on lagging chromosomes than segregating chromosomes [37]. In contrast, in *Xenopus laevis* α -amanitin treatment increased CENP-C levels [25], a result the authors attributed to increased residence time of the arrested transcript, while they suggested that α -amanitin-dependent degradation of Pol II [38] over longer incubation periods accounted for the previous conflicting results. Blocking transcription initiation or splicing reduced CENP-C levels in *Xenopus* [25]. Overall, results in three organisms are consistent with a model of RNA stabilizing the binding of CENP-C to DNA and/or chromatin.

RNA and Aurora-B

The CPC, comprising Aurora-B, INCENP, Survivin, and Borealin, regulates spindle attachment by sensing tension, correcting attachment errors, and activating the spindle assembly checkpoint [39]. It localizes to the kinetochore and **inner centromere** at metaphase, then migrates to the spindle midzone at late anaphase, where it regulates cytokinesis and plays a key role in preventing formation of micronuclei [40]. Affinity precipitation of mouse minor satellite RNA with biotinylated probes recovered CENP-A, Aurora-B, Survivin, and INCENP [14]. Reciprocally, centromeric RNAs of mouse and human cells were pulled down by antibodies to CENP-A, Aurora-B, Survivin, and INCENP [14,27]. Depletion of centromeric transcripts produced micronuclei and growth defects, similar to what was observed from depletion of Aurora-B. Knockdown of α -satellite RNA in HeLa cells [27] or inhibition of transcription with triptolide in *Xenopus* egg extracts [24] reduced the centromeric localization of Aurora-B, and caused improper spindle attachment and unaligned chromosomes. These results suggest that the immediate phenotypic consequences of depleting centromeric RNA are primarily due to improper regulation of Aurora-B and the CPC, with perhaps more minor contributions from destabilizing CENP-C.

Aurora-B from *Xenopus* bound directly and nonspecifically to RNA *in vitro*, but with a slight preference for the centromeric transcript *fcr1* [24,41]. *In vivo* it bound more than 600 RNAs, many

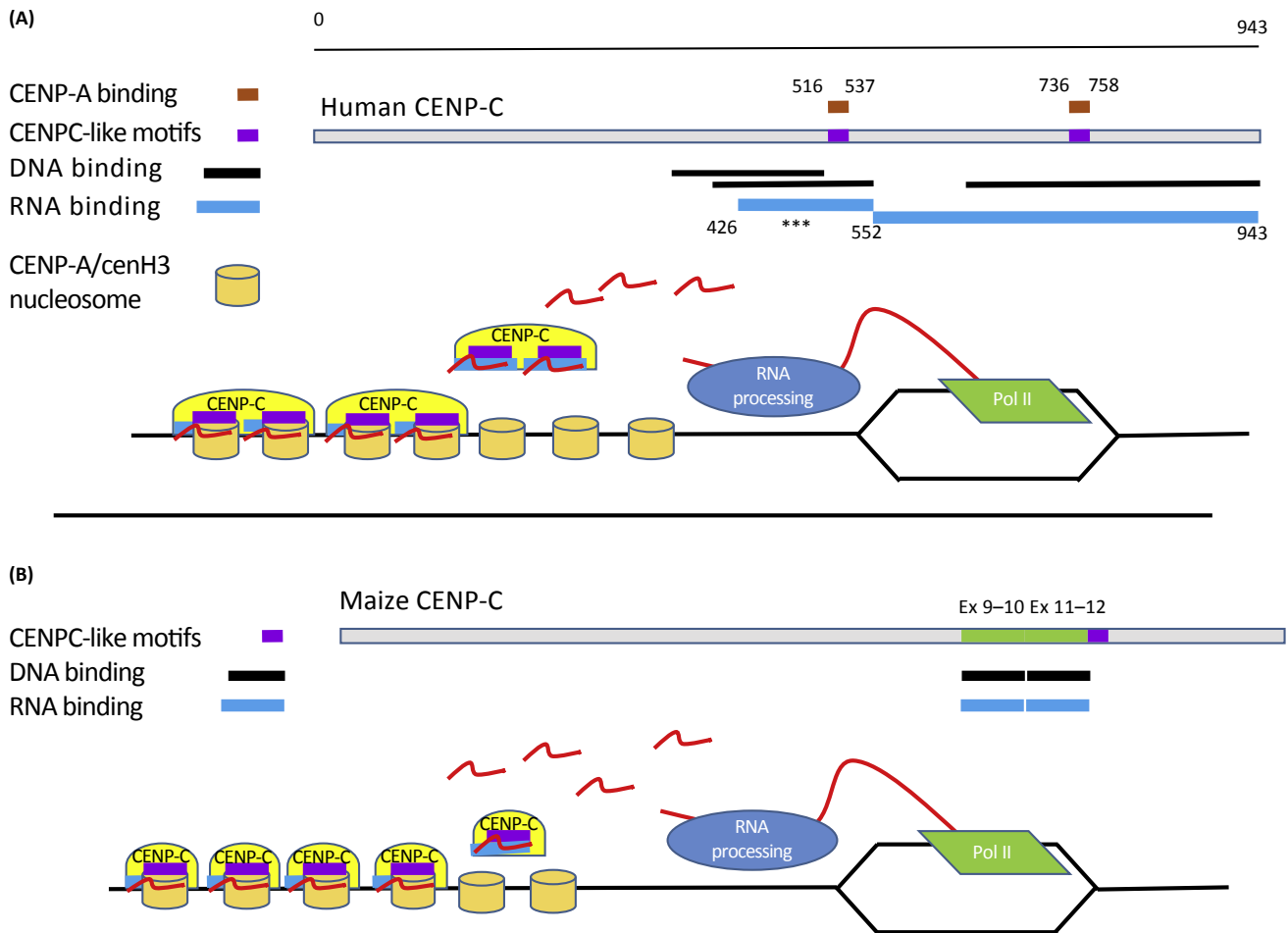
Micronuclei: small defective nuclei that often form around chromosomes lagging at anaphase.

Neocentromere: in humans, a new or recent centromere that forms when CENP-A occupies DNA sequence that was not previously a centromere. They typically lack α -satellite DNA, while old, inactive centromeres may retain some α -satellite but not CENP-A.

Non-B form DNA: any form of DNA which is not the standard B form double helix described by Watson and Crick.

Satellite: highly repetitive sequences arranged in tandem arrays. Centromeres in most plants and animals are composed of satellite DNA, such as α -satellite in humans or minor satellite in mice. Satellite monomers are usually 100–400 bp in length.

Small interfering RNAs (siRNAs): short 20–25 bp double-stranded RNAs that act through the RNA interference pathway to degrade RNAs that share complementary nucleotide sequences.



Trends in Genetics

Figure 1. RNA Binding by CENP-C. Key kinetochore protein CENP-C is characterized by one or two CENPC-like motifs, ~23 amino acid motifs that bind to CENP-A/cenH3 nucleosomes. In both human and maize CENP-C, regions just N terminal to CENPC-like motifs bind both DNA and RNA. (A) Human CENP-C has two CENPC-like motifs, one in the central region that is common to many vertebrates, and a more C terminal motif common to most eukaryotes. CENP-C was found to associate with centromeric RNA, possibly in an assembly pathway for kinetochore formation. *** Indicates three lysine residues critical for RNA binding. Data taken from [11,33–35]. (B) Maize CENP-C has a single CENPC motif. Exons 9–12 encode a duplicated region that binds both DNA and RNA, with small RNA binding facilitating DNA binding. Data from [36].

of which are enriched at the spindle. Borealin (Dasra-A) also bound RNA. Assembly of Aurora-B and Survivin into a complex and the kinase activity of Aurora-B were sensitive to RNase [14,41], and kinase activity could be rescued by adding RNA, perhaps through inducing allosteric changes on binding. Inhibition of transcription or knockdown of centromeric transcripts mislocalized Aurora-B from the inner centromere, but increased its overall activity in mitotic cells [24]. Similarly, in HeLa cells [27] centromeric RNA knockdown partially mislocalized Aurora-B but increased its activity [41], suggesting that centromeric RNA is necessary to localize and activate Aurora-B at the inner centromere and that elsewhere it may be activated by other RNAs.

Mitotic Transcription and Shugosin1

In animals, Pol II was active at the kinetochore and inner centromere during mitosis [20,23,24,42,43] and early G1 [18,43]. Elongating Pol II was also found at mitotically active neocentromeres, but not at an inactivated α -satellite-containing centromere, suggesting that

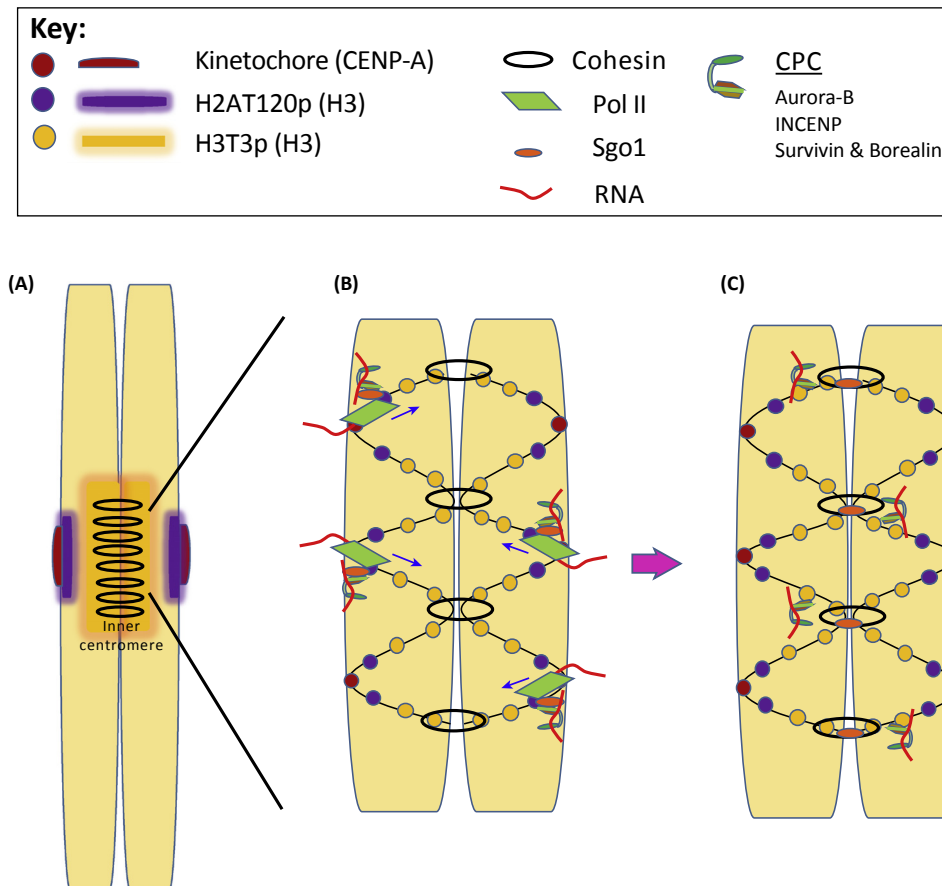
Pol II localizes to active kinetochores [20]. Pol II associated with histone H2A that was phosphorylated on T120 (H2AT120p) by the kinetochore kinase Bub1 [23]. Bub1 activity was required for localization of Pol II and for centromeric transcripts at the mitotic kinetochore of HeLa cells, suggesting that H2AT120p is essential to localize Pol II to kinetochores. However, it is not sufficient, since ectopic GFP-Bub1 on chromosome arms did not accumulate Pol II. CENP-A did not co-immunoprecipitate with H2AT120p in HeLa cells, implying that Pol II is recruited to the centromeric H3 subdomains. Centromeric transcript levels during interphase were unaffected by Bub1 depletion, suggesting that centromeric transcription may occur independent of Bub1 during interphase, but that Pol II is specifically recruited to the kinetochore by Bub1 activity during mitosis. Consistent with this, α -satellite was also transcribed from pericentromeres prior to mitosis, predominantly by Pol I [44].

Shugosin1 (Sgo1) is present at the inner centromere of HeLa cells during metaphase, where it binds to and protects cohesin. Sgo1 localizes initially to the kinetochore, where it binds to H2AT120p [23]. Mutants unable to bind H2AT120p failed to localize to either the kinetochore or the inner centromere. When Pol II transcription elongation during mitosis was inhibited or Pol II subunit Rbp2 was degraded, Sgo1 localized at the kinetochore but did not relocate to the inner centromere. Nonspecific RNA competed with H2AT120p for binding to Sgo1, and Sgo1 interacted with Pol II, leading to a model in which mitotic centromeric transcripts compete with and release Sgo1 from H2AT120p at the kinetochore, allowing it to move with Pol II to the inner centromere (Figure 2). Sgo1 also bound to the CPC subunit Borealin [45], raising the possibility that the CPC 'tags along' with Sgo1 to move from the kinetochore to the inner centromere, where Sgo1 binds to cohesin and Survivin binds H3T3p [46], which is phosphorylated by haspin, a kinase enriched at the inner centromere [39]. While this tag-along model is attractive, the roles of centromeric RNAs and transcription in Sgo1 and Aurora-B localization remain to be clarified.

Transcription and CENP-A Loading

At a human neocentromere lacking α -satellite DNA, a LINE1 transposon-derived transcript within the CENP-A domain of the neocentromere was incorporated into centromeric chromatin. RNA knockdown of the transcript reduced CENP-A incorporation and mitotic stability [13]. In α -satellite-tetO-based HACs [32], there were low levels of α -satellite-tetO transcript, and H3 centromeric subdomains were enriched in H3K4me2 and H3K36me2, chromatin marks associated with active transcription [47]. When the lysine-specific demethylase 1 fused to tetR-EYFP (tetR-EYFP-LSD1) was introduced, H3K4me2 became undetectable at the HAC after 3 days, whereas H3K9me3 and H3K27me3 were unaffected. CENP-A and CENP-C levels were reduced by half, though kinetochores remained functional. Transcripts from the HAC and H3K36me2 levels were rapidly reduced, compared with introducing a catalytically dead mutant tetR-EYFP-LSD1^{K661A}, indicating that loss of H3K4me2 leads to reduced HAC transcription. Incorporation of new CENP-A and the dedicated CENP-A chaperone, Holliday Junction Recognition Protein (HJURP), were significantly reduced. These results suggest that transcription or associated histone marks facilitate HJURP recruitment and CENP-A loading.

In another study, inhibition of Pol II elongation with α -amanitin in HeLa cells during early G1 reduced CENP-A by half and nearly abolished HJURP, although total cellular levels of CENP-A and HJURP were unaffected [18]. Knockdown of centromeric RNAs also greatly reduced CENP-A and HJURP on chromatin fibers, though Pol II, CENP-B, and CENP-C remained unaffected. In ChIP of CENP-A and HJURP from soluble and chromatin fractions of early G1 cells after light MNase digest, CENP-A and HJURP proteins co-purified in both fractions. A 1.3-kb centromeric RNA was co-immunoprecipitated from both fractions with CENP-A, but only from the soluble fraction with HJURP, suggesting that the RNA associates with soluble HJURP

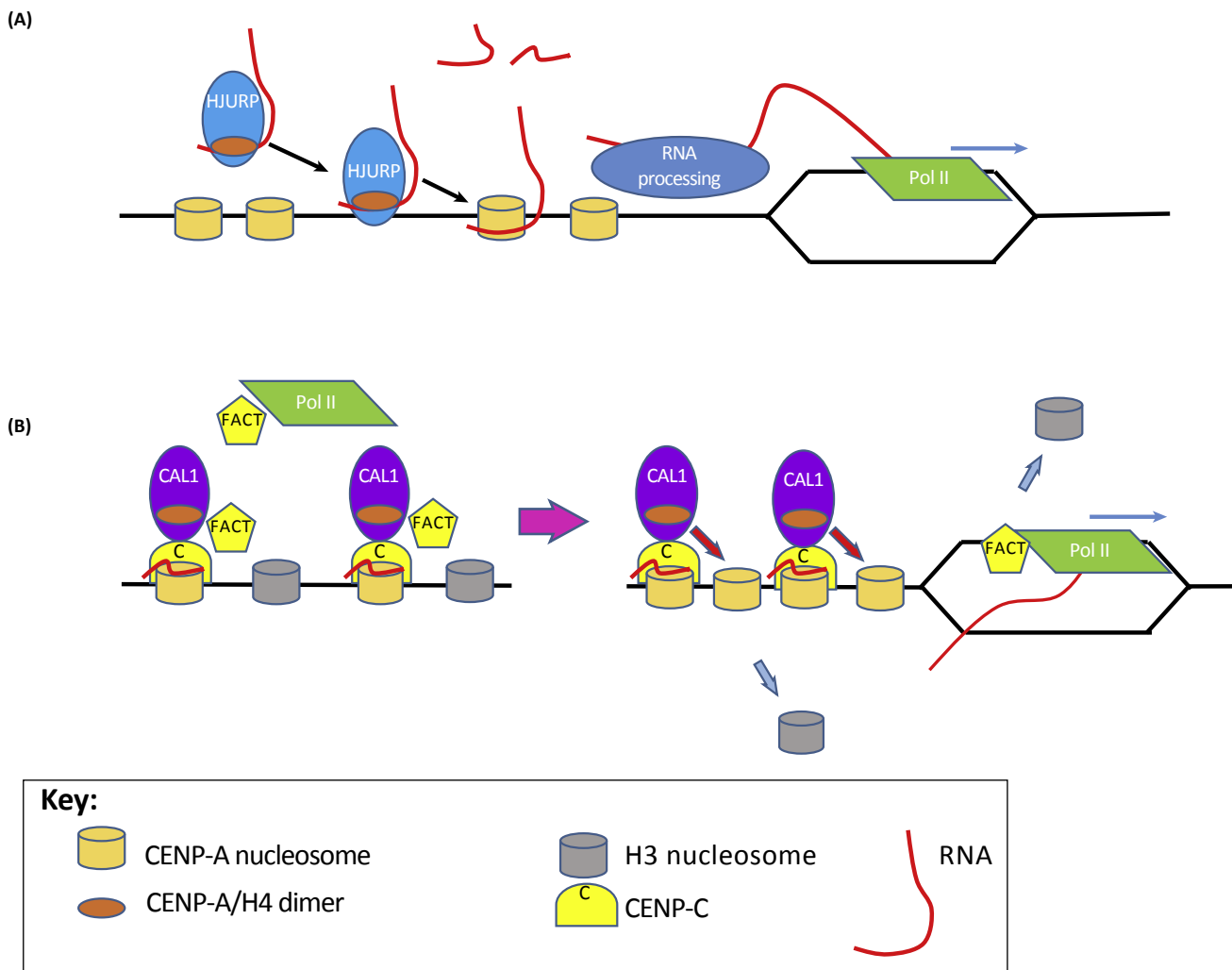


Trends in Genetics

Figure 2. Model of Centromeric Transcription and Localization of Shugosin1 (Sgo1) and the Chromosomal Passenger Complex (CPC) at the Inner Kinetochores. (A) Schematic of mitotic sister chromatids showing the kinetochores containing CENP-A nucleosomes and the kinase Bub1; a kinetochores-proximal region of H3 nucleosomes phosphorylated on H2AT120 by Bub1; and the inner centromere where sisters are held together by cohesin and H3 nucleosomes are phosphorylated on H3T3 by the kinase haspin. (B) Expansion of the centromere region in (A). Both Sgo1 and Pol II are recruited by H2AT120p near the kinetochores. Centromeric RNA binds to Sgo1, displacing it from H2AT120p and allowing it to travel with Pol II toward the inner centromere. Centromeric RNA also helps to assemble the CPC and activate the Aurora-B kinase. The Borealin subunit of the CPC binds to Sgo1, so the CPC can 'tag along' with Sgo1. (C) At the inner centromere, Sgo1 releases from Pol II and binds to and protects cohesin, while Survivin binds H3T3p. Based on [23,24,39,45,46].

and CENP-A in a preassembly complex (Figure 3A). In subsequent studies, these authors [48] and others [26] found centromeric RNAs in sizes ranging from 300 nt to 2450 nt, indicating the 1.3 kb transcript is one among many.

Individual human α -satellite monomers may be only 50–70% identical and are arranged in higher order repeat arrays that are specific for different chromosomes [49]. A recent study found that probes for transcripts from specific arrays in RPE1 cells colocalized only with the corresponding DNA array ('cis-acting'), indicating that the transcripts do not act in trans at other centromeres [26], though trans-acting transcripts have been inferred for *Xenopus* [24] and in *Drosophila* (Box 1). Human α -satellite transcripts were localized to both active centromere arrays and inactive arrays, but were less stable at inactive arrays [26]. Transcripts varied in size between 300 and 2000 nt and



Trends in Genetics

Figure 3. Models of CENP-A Loading and Centromeric Transcription. (A) Transcripts from human α -satellite are found in a soluble complex with CENP-A and its specific chaperone Holliday Junction Recognition Protein (HJURP). Localization of CENP-A and HJURP at centromeres is drastically reduced in the absence of α -satellite RNA, suggesting they are loaded onto centromeres in this complex, dependent on centromeric transcripts. Data from [18]. (B) *Drosophila* CENP-A and its chaperone CAL1 are targeted to centromeres during M-G1 by CENP-C and are loosely associated and extractable in high salt in the absence of transcription, but CENP-A becomes fully incorporated after transcription. New CENP-A loading is dependent on the general nucleosome chaperone Facilitates Transcription (FACT), which remodels chromatin during transcription and may help evict H3 nucleosomes. Data from [22,43].

between cell lines. Immunoprecipitation of CENP-A and CENP-C in HAP1 cells yielded RNAs from active centromere arrays, but not from an inactive array, whereas immunoprecipitation of CENP-B, which is present at both active and inactive arrays, recovered RNAs from both. Depletion of RNA in RPE1 cells with array-specific siRNAs led to reduction of CENP-A and CENP-C only at the targeted array and impaired new CENP-A loading, which implies that centromere RNAs play a causal role in the loading or stabilization of CENP-A.

Taken together these four studies suggest that transcription or transcripts are required for HJURP recruitment and efficient CENP-A loading at human centromeres, neocentromeres, and HACs, and suggest that the primary sequences of the transcripts may be irrelevant.

Box 1. Trans-Acting Noncentromeric Transcripts in Centromeres?

In *Drosophila melanogaster* S2 cells, noncoding transcripts of the 359-bp satellite, SATIII, appear to be important for centromere function [42], even though they are not centromeric [17,70]. Immunoprecipitation of CENP-C co-precipitated SATIII RNA, indicating a complex of SATIII RNA with CENP-C. A SATIII probe detected transcripts on the X chromosome, where 359-bp repeats form an ~11 Mb block of pericentric heterochromatin, and in the pericentric regions of chromosomes 2 and 3, with some overlap with the kinetochore, but not on chromosome 4 [42]. A subsequent study determined that the autosomal signals were from hybridization to DNA of SATIII-related 260-, 353-, and 356-bp repeats found in those locations, and confirmed that both RNA and DNA hybridization signals mark these regions, suggesting signal from cis-transcripts [43]. Regardless of the origin of the SATIII RNAs, knockdown of them resulted in mitotic defects in both S2 cells and embryos, including lagging chromosomes, micronuclei, and partial loss of CENP-A, CENP-C, and the kinetochore protein KNL1 [42], similar to phenotypes of centromeric RNA knockdowns in mammalian cells [18,26,27]. Loading of new CENP-A and CENP-C was also reduced [42]. Conversely, knockdown of CENP-C resulted in significant loss of SATIII transcripts on mitotic chromosomes, though the reason for loss of pericentric RNA that mostly does not overlap CENP-C is not clear. Elongating Pol II was present at *Drosophila* mitotic kinetochores, but was not evident at pericentromeres [42,43], suggesting that the SATIII transcripts are not transcribed during mitosis. That transcription from the centromeres themselves is involved in CENP-A loading is implied by transcription-dependent salt-resistant CENP-A incorporation [43], by CENP-A loading with tethered CAL1-GFP-lacI, and by dependence of new CENP-A loading on FACT at both ectopic and native centromeres [22]. Trans-acting SATIII transcripts seem unlikely to disrupt nucleosomes to aid CENP-A loading directly, though it is possible that they regulate CAL1, CENP-C, or Aurora-B, or have indirect effects by altering pericentric heterochromatin.

Two populations of CENP-A chromatin particles have been described in human centromeres that differ in their stability toward salt [50]. MNase-treated centromeric chromatin subjected to native ChIP with antibodies to CENP-A and extracted in low salt yielded smaller fragments that were depleted for the CENP-B box, the binding site of CENP-B, relative to larger fragments from high salt extraction or the pellet. This raised the possibility that the minor fraction (~17%) of low-salt-soluble fragments derives from an incomplete complex, whereas the >80% of centromeric chromatin in the high salt fraction represents the full inner kinetochore complex stabilized by CENP-B. Could centromeric transcription differentiate these two populations? In *Drosophila melanogaster* S2 cells, new tamoxifen-inducible HA-tagged CENP-A localized to the centromere in mitosis and G1 even when transcription was inhibited, but was removed by high salt extraction [43]. Transcription was required for stable salt-resistant incorporation of CENP-A. The dedicated CENP-A chaperone CAL1 was localized to centromeres in low salt but removed by high salt, whereas CENP-C was chromatin-associated in both conditions, suggesting a model in which CAL1 and CENP-A associate with chromatin-bound CENP-C before CENP-A is stably incorporated in the wake of transcription (Figure 3B).

Further evidence for transcription-coupled loading of CENP-A in *Drosophila* S2 cells comes from an inducible CAL1-GFP-lacI fusion that was tethered to a *lacO* array, resulting in transcription of the *lacO* backbone (*lacO^b*) and in loading of CENP-A [22]. Immunoprecipitation of CAL1 revealed an association with Pol II and with the general histone chaperone **Facilitates Transcription (FACT)**, which binds CAL1 *in vitro*. A mutant CAL1-GFP-lacI unable to load CENP-A was nevertheless able to bind FACT and promote transcription of *lacO^b*, indicating that CENP-A loading is unnecessary for CAL1-mediated transcription. Immunostaining of FACT subunits was strongest at the kinetochore on mitotic chromosomes, at the same time that CENP-A normally loads in *Drosophila* [51]. FACT knockdown eliminated CAL1-GFP-lacI-induced transcription of *lacO^b* and reduced loading of CENP-A [22]. Likewise, depletion of FACT strongly reduced new CENP-A loading at endogenous centromeres, but had little effect on retention of existing CENP-A. This supports a model in which CAL1 recruits FACT and Pol II to centromeres, where they facilitate CENP-A loading, perhaps by dislodging H3 nucleosomes and replacing them with CENP-A nucleosomes.

FACT has also been shown to be required for CENP-A loading in chicken DT-40 cells [52], raising the possibility that HJURP acts through a FACT-based mechanism similarly to CAL1.

Interestingly, HJURP-lacI tethered to *lacO* repeats can expand the region of CENP-A loading from 10 kb of *lacO* repeats to 150 kb surrounding *lacO* [53], perhaps through FACT-mediated transcriptional deposition.

In contrast, in fission yeast FACT mutants have little effect on CENP-A loading, but they impair H3 loading, resulting in promiscuous CENP-A incorporation [54]. Similarly, in budding yeast FACT binds to the E3 ubiquitin ligase Psh1 to target misincorporated Cse4 in euchromatin for degradation [55].

Kinetochore Null2 (KNL2, M18BP1 in mammals) promotes cenH3 incorporation at centromeres during G2 in *Arabidopsis thaliana* and is present throughout the cell cycle except from metaphase to late anaphase [56]. The C terminal portion of KNL2 binds both RNA and DNA *in vitro*, with putative DNA-binding domains on either side of the CENPC-k motif that are necessary for centromeric localization [57]. Unlike in maize CENP-C, where small RNA binding increased DNA binding, small RNA had little effect on KNL2 binding but longer RNA competed with DNA for binding KNL2. Possibly centromere transcription prior to metaphase dislodges KNL2 from centromeric DNA at the time when it is replaced by CENP-C to build the kinetochore.

When arrays of *lacO* repeat DNA were injected into *Caenorhabditis elegans* gonads, they became incorporated into one-cell embryos and formed artificial chromosomes (ACs) that were enriched in acetylation of H3K9 and H4 and formed *de novo* centromeres [58]. Tethering GFP::lacI to a histone deacetylase (GFP::lacI::HDA-1), or inhibiting transcription with α -amanitin, reduced levels of H3/H4 acetylation, CENP-A deposition, Pol II initiation, and segregation of new ACs in early embryos. There was less effect on later stage embryos, when acetylation declined in GFP::lacI and in enzymatically dead mutant GFP::lacI::HAD-1 (H145A) control embryos. This suggests that acetylation facilitates transcription and *de novo* CENP-A deposition, though it may be unnecessary to maintain centromeres after establishment.

Promoters in Centromeres

In their pioneering study, Topp *et al.* speculated that centromeric transcripts initiated from CRMs that are abundant in maize centromeres [10]. KERV-1 transposons in tammar wallaby initiated transcription on both strands [15]. In *Plasmodium*, with regional centromeres of 2–3 kb and ~97% AT content, all fragments from centromeres had promoter activity, often bidirectional [12]. The central domains of fission yeast centromeres also contained numerous promoters [21]. Different portions of the central domains were not equivalent in their ability to assemble CENP-A chromatin, and portions with lower transcriptional activity were more effective than those with higher activity. Mutants that are defective in restarting stalled Pol II increased CENP-A loading, suggesting that stalled polymerases make a favorable chromatin environment for CENP-A loading, perhaps through promoting remodeling or eviction of H3 nucleosomes. *In vitro* CENP-A nucleosomes themselves present a greater barrier to transcription than H3 nucleosomes [59] and might promote their own recruitment.

Active transcription requires first denaturing DNA at promoters, and denatured and other **non-B form DNA** regions can be detected by permanganate/S1 nuclease sequencing [60,61]. Such partially denatured non-B form DNA regions were found to be abundant in both human α -satellite and mouse major and minor satellites from activated B cells [62]. Non-B DNA has been frequently predicted to form at satellite centromeres in many eukaryotes [62–68]. A form of non-B DNA that could account for its detection in activated mouse and human B cells is

Box 2. The CENP-B Paradox

CENP-B is a DNA-binding protein that binds to a 17-bp motif, the 'CENP-B box', which is typically present in every other human α -satellite repeat. CENP-B is highly conserved through vertebrate evolution; however, the CENP-B box is not found in all primate α -satellite arrays [71,72], referred to by Earnshaw and coworkers as the 'CENP-B paradox' [73]. Furthermore, CENP-B and the CENP-B Box are required for HAC formation [74,75], yet counterintuitively mouse CENP-B mutants are viable [76]. A possible resolution of the CENP-B paradox arises from the finding that centromeric satellites that lack CENP-B boxes, including neocentromeres from humans and chickens, are more highly enriched in dyad symmetries than are centromeres with CENP-B boxes and are predicted to form more stable DNA secondary structures [62]. In centromeres of great apes and mice, potassium permanganate footprinting revealed enrichment for non-B form DNA despite having few dyad symmetries, suggesting that dyad symmetries and CENP-B binding, which bends DNA 60° , are alternative ways to form non-B form DNA at centromeres. In support of this possibility, budding yeast species with high dyad symmetry enrichment are bound at their centromere determining element I (CDEI) by the helix-loop-helix protein Cbf1, whereas CDEI of budding yeasts with low dyad symmetries have the motif for Reb1, which, like CENP-B, is predicted to bend DNA 60° . These observations led to the speculation [62] that centromeres may be specified by cruciform structures formed by dyad symmetries or induced by DNA-bending proteins, forming a binding substrate for HJURP and its homolog Scm3. Alternatively or in addition, non-B form DNA in centromeres may facilitate transcription by Pol II, enabling CENP-A loading during nucleosome remodeling. CENP-B may be essential to establish new HACs, but dispensable for established centromeres where Pol II can be recruited by H2AT120p [23] and HJURP can be recruited by the KNL2 (M18BP1) complex [77] (Figure I).

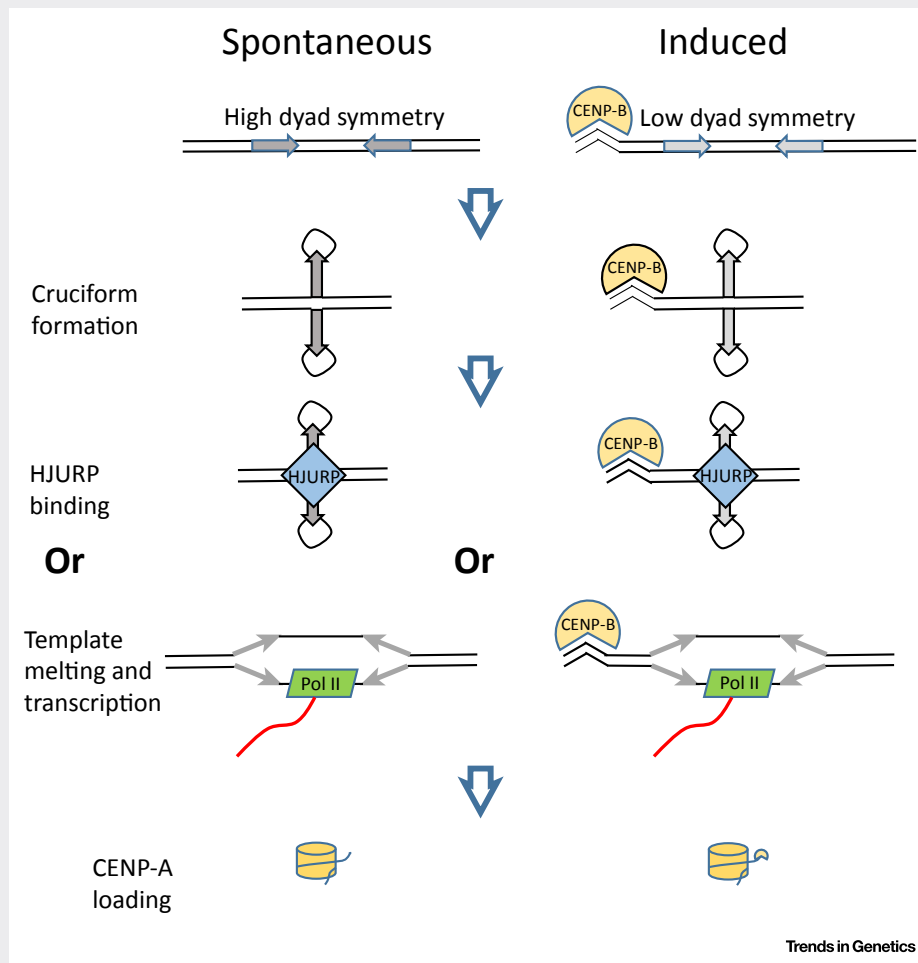


Figure I. Proposals for Non-B form DNA in CENP-A Loading.

cruciform extrusion, promoted by short (<10 bp) **dyad symmetries**. Significant dyad symmetries are widespread at centromeres throughout the eukaryotic domain, including satellite centromeres of primates, mouse, horse, chicken, stickleback, and plants, regional centromeres of fission yeast, and point centromeres of budding yeasts [62]. Interestingly, non-proliferating ('resting') mouse B cells showed reduced levels of centromeric non-B DNA, which is consistent with the possibility that proliferation induces cruciform extrusion for 'seeding' centromeres. Another form of non-B DNA is R loops, which form at transcribing centromeres in RPE cells from RNA–DNA hybrids, and aid Aurora B activation [68].

Although the role of non-B DNA at centromeres is largely unknown, two hypotheses have been proposed [62]. One is that the four-way junctions of cruciforms are bound by Holliday junction binding activity of HJURP and its Scm3 ortholog in yeast, whereupon HJURP would load CENP-A/H4. Alternatively, non-B DNA might result from transcriptional initiation, where melting of DNA is required for engagement of Pol II, and from Pol II elongation, which moves the denaturation bubble forward. These hypotheses are not mutually exclusive, and in both cases the enigmatic CENP-B sequence-specific DNA binding protein likely plays a role (Box 2). Given the enrichment of non-B form DNA at centromeres throughout the eukaryotic domain, it seems likely that this feature of centromeres can provide a basis for centromere specification despite the lack of primary sequence conservation.

Concluding Remarks and Future Perspectives

Centromeric transcription and transcripts have been implicated in stabilizing CENP-C, assembling and localizing the CPC, transporting Sgo1, localizing HJURP and loading CENP-A, and possibly in regulating KNL2. Future experiments need to better distinguish the effects of transcripts and transcription (Box 3), while further exploring and differentiating the roles of RNA with different kinetochore components (see Outstanding Questions). Better characterization of centromeric RNAs and their processing might give insight into how transcripts have their effects. Promoters within centromeres also need to be more fully characterized. Tethering

Box 3. Transcript or Transcription?

As first proposed by Topp *et al.* [10], transcription may facilitate cenH3 deposition [13,18,22,26,43,47,58], or transcripts may serve to target or stabilize kinetochore components [14,24–27,36,41]. Evidence exists for both possibilities, however, the incompleteness of RNA knockdown or transcription inhibition can sometimes complicate distinguishing, for example, reductions in new CENP-A loading from destabilization of new CENP-A [26]. Reports of centromeric RNA complexed with CENP-C in the nucleolus [11], soluble RNA complexes with HJURP and CENP-A [18], and trans-acting RNAs in *Xenopus* [24] and *Drosophila* [42] are evidence in favor of roles for the transcripts not coupled to transcription.

The low abundance of centromeric RNA, often about 1:1 with the DNA that encodes it [10,26], and processing to sizes usually much less than 1 kb [10,11,15,30,36] is in sharp contrast to RNAs that play a structural role in specialized chromatin like *Xist*, *rox1*, and *rox2* [5]. A single passage of Pol II, however, might be sufficient to transport Sgo1 from the kinetochore to the inner centromere [23] or to disrupt chromatin structure and load new cenH3 nucleosomes [22,43]. If low-level transcription favors loading cenH3 over H3, this offers a possible explanation for why most eukaryotes do not load CENP-A during replication where H3 loading is presumably preferred.

Multiple or bidirectional promoters and non-B form DNA in centromeres may provide multiple sites to initiate cenH3 loading, or they may provide a multiplicity of transcripts with indistinguishable purposes. A notable feature of the kinetochore proteins that bind centromeric transcripts is how little specificity they have for their target RNAs. *In vitro* binding assays revealed no sequence specificity for binding of Sgo1 [23], Aurora-B [41], KNL2 [57], or maize CENP-C [36], and partial specificity for mammalian CENP-C. Binding to RNA appears to facilitate Aurora-B activation, but centromeric RNA does not seem to be specifically required, except to insure the correct localization of Aurora-B to the inner centromere [24]. This suggests that proximity to the site of transcription may be the factor that determines which RNAs are bound, consistent with cis-acting effects of human α -satellite RNAs [26]. CENP-C, KNL2, and HJURP all bind directly to cenH3 nucleosomes and to DNA [33–35,57,78–81], suggesting that they may bind to RNA as it is transcribed, which might facilitate their reloading onto DNA after disruption by transcription.

Outstanding Questions

Does the primary sequence of centromeric RNA matter?

Do double-stranded RNA or RNA–DNA hybrids (R loops) play a role in centromeres?

How are centromeric RNAs processed?

Is there a nucleolar or soluble pathway of RNA incorporation into chromatin, or is RNA bound as it is being transcribed?

What roles of cis-acting centromeric RNAs differ from those of trans-acting RNAs?

Do trans-acting transcripts compete with cis-acting transcripts?

Are kinetochore proteins dislodged by centromeric transcription? If not, how are they retained? If so, how are they reassembled? Is RNA involved?

How does RNA stabilize DNA binding by CENP-C?

How does RNA affect the assembly of the CPC and the activity of Aurora-B?

Does Pol II carry Sgo1 and/or the CPC from the kinetochore to the inner centromere?

Is transcription required for new CENP-A loading? Or is it an alternative pathway of CENP-A loading? Does this depend on the organism or developmental stage?

Do all centromeres act as promoters? Are transcription factors required?

Does transcription initiate at nucleosome gaps? Does non-B form DNA create nucleosome gaps?

Does non-B form DNA facilitate transcription? Does it stall transcription?

Does HJURP bind to RNA? To non-B form DNA?

Does RNA regulate KNL2 localization?

experiments at ectopic centromeres show promise for teasing apart cotranscriptional mechanisms, which then need verification at native centromeres and assessment of the extent to which these mechanisms are conserved across organisms. It is of interest to note that whereas many kinetochore proteins have not been identified outside of vertebrates and ascomycetes, cenH3, CENP-C, and KNL2 are widely conserved in eukaryotes [57,69], suggesting that centromeric transcription and transcripts may have been ancestral in centromere and kinetochore dynamics.

Acknowledgements

We thank members of our laboratory for helpful discussions, and Sabine Hedouin, Sue Biggins, and anonymous reviewers for comments on the manuscript.

References

- Rieder, C.L. (1979) Ribonucleoprotein staining of centrioles and kinetochores in newt lung cell spindles. *J. Cell Biol.* 80, 1–9
- Brasletton, J.P. (1975) Ribonucleoprotein staining of *Allium cepa* kinetochores. *Cytobiologie* 12, 148–151
- Hill, A. and Bloom, K. (1987) Genetic manipulation of centromere function. *Mol. Cell. Biol.* 7, 2397–2405
- Cardone, M.F. et al. (2006) Independent centromere formation in a capricious, gene-free domain of chromosome 13q21 in Old World monkeys and pigs. *Genome Biol.* 7, R91
- Sun, Q. et al. (2018) Nuclear long noncoding RNAs: key regulators of gene expression. *Trends Genet.* 34, 142–157
- Blower, M.D. et al. (2002) Conserved organization of centromeric chromatin in flies and humans. *Dev. Cell* 2, 319–330
- Sullivan, B.A. and Karpen, G.H. (2004) Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.* 11, 1076–1083
- Nagaki, K. et al. (2004) Sequencing of a rice centromere uncovers active genes. *Nat. Genet.* 36, 138–145
- Yan, H. et al. (2008) Intergenic locations of rice centromeric chromatin. *PLoS Biol.* 6, e286
- Topp, C.N. et al. (2004) Centromere-encoded RNAs are integral components of the maize kinetochore. *Proc. Natl. Acad. Sci. U. S. A.* 101, 15986–15991
- Wong, L.H. et al. (2007) Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res.* 17, 1146–1160
- Li, F. et al. (2008) Nuclear non-coding RNAs are transcribed from the centromeres of *Plasmodium falciparum* and are associated with centromeric chromatin. *J. Biol. Chem.* 283, 5692–5698
- Chueh, A.C. et al. (2009) LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. *PLoS Genet.* 5, e1000354
- Ferri, F. et al. (2009) Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Res.* 37, 5071–5080
- Carone, D.M. et al. (2009) A new class of retroviral and satellite encoded small RNAs emanates from mammalian centromeres. *Chromosoma* 118, 113–125
- Choi, E.S. et al. (2011) Identification of noncoding transcripts from within CENP-A chromatin at fission yeast centromeres. *J. Biol. Chem.* 286, 23600–23607
- Talbert, P.B. et al. (2018) Simple and complex centromeric satellites in *Drosophila* sibling species. *Genetics* 208, 977–990
- 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
- Ohkuni, K. and Kitagawa, K. (2011) Endogenous transcription at the centromere facilitates centromere activity in budding yeast. *Curr. Biol.* 21, 1695–1703
- Chan, F.L. et al. (2012) Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proc. Natl. Acad. Sci. U. S. A.* 109, 1979–1984
- Catania, S. et al. (2015) Sequence features and transcriptional stalling within centromere DNA promote establishment of CENP-A chromatin. *PLoS Genet.* 11, e1004986
- Chen, C.C. et al. (2015) Establishment of centromeric chromatin by the CENP-A assembly factor CAL1 requires FACT-mediated transcription. *Dev. Cell* 34, 73–84
- Liu, H. et al. (2015) Mitotic transcription installs Sgo1 at centromeres to coordinate chromosome segregation. *Mol. Cell* 59, 426–436
- Blower, M.D. (2016) Centromeric transcription regulates Aurora-B localization and activation. *Cell Rep.* 15, 1624–1633
- Grenfell, A.W. et al. (2016) Mitotic noncoding RNA processing promotes kinetochore and spindle assembly in *Xenopus*. *J. Cell Biol.* 214, 133–141
- McNulty, S.M. et al. (2017) Human centromeres produce chromosome-specific and array-specific alpha satellite transcripts that are complexed with CENP-A and CENP-C. *Dev. Cell* 42, 226–240
- Ideue, T. et al. (2014) Involvement of satellite I noncoding RNA in regulation of chromosome segregation. *Genes Cells* 19, 528–538
- May, B.P. et al. (2005) Differential regulation of strand-specific transcripts from *Arabidopsis* centromeric satellite repeats. *PLoS Genet.* 1, e79
- Koo, D.H. et al. (2016) Chromatin-associated transcripts of tandemly repetitive DNA sequences revealed by RNA-FISH. *Chromosome Res.* 24, 467–480
- Bouzinba-Segard, H. et al. (2006) Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc. Natl. Acad. Sci. U. S. A.* 103, 8709–8714
- Bergmann, J.H. et al. (2012) Epigenetic engineering: histone H3K9 acetylation is compatible with kinetochore structure and function. *J. Cell Sci.* 125, 411–421
- Nakano, M. et al. (2008) Inactivation of a human kinetochore by specific targeting of chromatin modifiers. *Dev. Cell* 14, 507–522
- Sugimoto, K. et al. (1997) Characterization of internal DNA-binding and C-terminal dimerization domains of human centromere/kinetochore autoantigen CENP-C *in vitro*: role of DNA-binding and self-associating activities in kinetochore organization. *Chromosome Res.* 5, 132–141
- Yang, C.H. et al. (1996) Identification of overlapping DNA-binding and centromere-targeting domains in the human kinetochore protein CENP-C. *Mol. Cell. Biol.* 16, 3576–3586
- Kato, H. et al. (2013) A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science* 340, 1110–1113
- Du, Y. et al. (2010) DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. *PLoS Genet.* 6, e1000835

37. Chan, F.L. and Wong, L.H. (2012) Transcription in the maintenance of centromere chromatin identity. *Nucleic Acids Res.* 40, 11178–11188
38. Nguyen, V.T. *et al.* (1996) *In vivo* degradation of RNA polymerase II largest subunit triggered by alpha-amanitin. *Nucleic Acids Res.* 24, 2924–2929
39. Hindriksen, S. *et al.* (2017) The ins and outs of Aurora B inner centromere localization. *Front. Cell Dev. Biol.* 5, 112
40. Warecki, B. and Sullivan, W. (2018) Aurora B-mediated exclusion of HP1a from late-segregating chromatin prevents formation of micronuclei. *BioRxiv* Published online February 21, 2018. <http://dx.doi.org/10.1101/268912>
41. Jambhekar, A. *et al.* (2014) RNA stimulates Aurora B kinase activity during mitosis. *PLoS One* 9, e100748
42. Rosic, S. *et al.* (2014) Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J. Cell Biol.* 207, 335–349
43. Bobkov, G.O.M. *et al.* (2018) Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation. *J. Cell Biol.* Published online April 8, 2018. <http://dx.doi.org/10.1083/jcb.201611087>
44. Johnson, W.L. *et al.* (2017) RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *Elife* 6, e25299
45. Tsukahara, T. *et al.* (2010) Phosphorylation of the CPC by Cdk1 promotes chromosome bi-orientation. *Nature* 467, 719–723
46. Yamagishi, Y. *et al.* (2010) Two histone marks establish the inner centromere and chromosome bi-orientation. *Science* 330, 239–243
47. Bergmann, J.H. *et al.* (2011) Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO J.* 30, 328–340
48. Quenet, D. *et al.* (2017) CENP-A associated lncRNAs influence chromosome segregation in human cells. *BioRxiv* Published online January 3, 2017. <http://dx.doi.org/10.1101/097956>
49. Aldrup-Macdonald, M.E. and Sullivan, B.A. (2014) The past, present, and future of human centromere genomics. *Genes (Base)* 5, 33–50
50. Thakur, J. and Henikoff, S. (2018) Unexpected conformational variations of the human centromeric chromatin complex. *Genes Dev.* 32, 20–25
51. Mellone, B.G. *et al.* (2011) Assembly of *Drosophila* centromeric chromatin proteins during mitosis. *PLoS Genet.* 7, e1002068
52. Okada, M. *et al.* (2009) CENP-H-containing complex facilitates centromere deposition of CENP-A in cooperation with FACT and CHD1. *Mol. Biol. Cell* 20, 3986–3995
53. Perpelescu, M. *et al.* (2015) HJURP is involved in the expansion of centromeric chromatin. *Mol. Biol. Cell* 26, 2742–2754
54. Choi, E.S. *et al.* (2012) Factors that promote H3 chromatin integrity during transcription prevent promiscuous deposition of CENP-A(Cnp1) in fission yeast. *PLoS Genet.* 8, e1002985
55. Deyter, G.M. and Biggins, S. (2014) The FACT complex interacts with the E3 ubiquitin ligase Psh1 to prevent ectopic localization of CENP-A. *Genes Dev.* 28, 1815–1826
56. Lermontova, I. *et al.* (2013) *Arabidopsis* kinetochore null2 is an upstream component for centromeric histone H3 variant cenH3 deposition at centromeres. *Plant Cell* 25, 3389–3404
57. Sandmann, M. *et al.* (2017) Targeting of *Arabidopsis* KNL2 to centromeres depends on the conserved CENPC-k motif in its C terminus. *Plant Cell* 29, 144–155
58. Zhu, J. *et al.* (2018) Histone H3K9 and H4 acetylations and transcription facilitate the initial CENP-A (HCP-3) deposition and *de novo* centromere establishment in *Caenorhabditis elegans* artificial chromosomes. *Epigenetics Chromatin* 11, 16
59. Shandilya, J. *et al.* (2014) Centromeric histone variant CENP-A represses acetylation-dependent chromatin transcription that is relieved by histone chaperone NPM1. *J. Biochem.* 156, 221–227
60. Kouzine, F. *et al.* (2017) Permanganate/S1 nuclease footprinting reveals non-B DNA structures with regulatory potential across a mammalian genome. *Cell Syst.* 4, 344–356
61. Kouzine, F. *et al.* (2013) Global regulation of promoter melting in naive lymphocytes. *Cell* 153, 988–999
62. Kasinathan, S. and Henikoff, S. (2018) Non-B-form DNA is enriched at centromeres. *Mol. Biol. Evol.* 35, 949–962
63. Chou, S.H. *et al.* (1996) Hairpin loops consisting of single adenine residues closed by sheared A.A. and G.G. pairs formed by the DNA triplets AAA and GAG: solution structure of the d(GTA-CAAAGTAC) hairpin. *J. Mol. Biol.* 264, 981–1001
64. Ohno, M. *et al.* (2002) Triplex-forming DNAs in the human interphase nucleus visualized *in situ* by polypurine/polypyrimidine DNA probes and antitriplex antibodies. *Chromosoma* 111, 201–213
65. Jonstrup, A.T. *et al.* (2008) Hairpin structures formed by alpha satellite DNA of human centromeres are cleaved by human topoisomerase IIalpha. *Nucleic Acids Res.* 36, 6165–6174
66. Garavis, M. *et al.* (2015) Centromeric alpha-satellite DNA adopts dimeric i-motif structures capped by AT Hoogsteen base pairs. *Chemistry* 21, 9816–9824
67. Aze, A. *et al.* (2016) Centromeric DNA replication reconstitution reveals DNA loops and ATR checkpoint suppression. *Nat. Cell Biol.* 18, 684–691
68. Kabeche, L. *et al.* (2018) A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science* 359, 108–114
69. Drinnenberg, I.A. and Akiyoshi, B. (2017) Evolutionary lessons from species with unique kinetochores. *Prog. Mol. Subcell. Biol.* 56, 111–138
70. Tolchkov, E.V. *et al.* (2000) The size and internal structure of a heterochromatic block determine its ability to induce position effect variegation in *Drosophila melanogaster*. *Genetics* 154, 1611–1626
71. Haaf, T. *et al.* (1995) Presence and abundance of CENP-B box sequences in great ape subsets of primate-specific alpha-satellite DNA. *J. Mol. Evol.* 41, 487–491
72. Masumoto, H. *et al.* (1989) A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. *J. Cell Biol.* 109, 1963–1973
73. Goldberg, I.G. *et al.* (1996) Surprising deficiency of CENP-B binding sites in African green monkey alpha-satellite DNA: implications for CENP-B function at centromeres. *Mol. Cell Biol.* 16, 5156–5168
74. Ohzeki, J. *et al.* (2002) CENP-B box is required for *de novo* centromere chromatin assembly on human alphoid DNA. *J. Cell Biol.* 159, 765–775
75. Okada, T. *et al.* (2007) CENP-B controls centromere formation depending on the chromatin context. *Cell* 131, 1287–1300
76. Kapoor, M. *et al.* (1998) The cenpB gene is not essential in mice. *Chromosoma* 107, 570–576
77. Nardi, I.K. *et al.* (2016) Licensing of centromeric chromatin assembly through the Mis18alpha-Mis18beta heterotetramer. *Mol. Cell* 61, 774–787
78. French, B.T. *et al.* (2017) *Xenopus laevis* M18BP1 directly binds existing CENP-A nucleosomes to promote centromeric chromatin assembly. *Dev. Cell* 42, 190–199
79. Hori, T. *et al.* (2017) Association of M18BP1/KNL2 with CENP-A nucleosome is essential for centromere formation in non-mammalian vertebrates. *Dev. Cell* 42, 181–189
80. Foltz, D.R. *et al.* (2009) Centromere-specific assembly of CENP-A nucleosomes is mediated by HJURP. *Cell* 137, 472–484
81. Kato, T. *et al.* (2007) Activation of Holliday junction recognizing protein involved in the chromosomal stability and immortality of cancer cells. *Cancer Res.* 67, 8544–8553