



No longer a nuisance: long non-coding RNAs join CENP-A in epigenetic centromere regulation

Silvana Rošić¹ · Sylvia Erhardt²

Received: 15 September 2015/Revised: 8 December 2015/Accepted: 17 December 2015/Published online: 9 January 2016
© Springer International Publishing 2016

Abstract Centromeres represent the basis for kinetochore formation, and are essential for proper chromosome segregation during mitosis. Despite these essential roles, centromeres are not defined by specific DNA sequences, but by epigenetic means. The histone variant CENP-A controls centromere identity epigenetically and is essential for recruiting kinetochore components that attach the chromosomes to the mitotic spindle during mitosis. Recently, a new player in centromere regulation has emerged: long non-coding RNAs transcribed from repetitive regions of centromeric DNA function in regulating centromeres epigenetically. This review summarizes recent findings on the essential roles that transcription, pericentromeric transcripts, and centromere-derived RNAs play in centromere biology.

Keywords Centromere · Kinetochore · Mitosis · Long non-coding RNA · Epigenetics · Transcription

Epigenetic regulation of centromeres as we know it: CENP-A function and regulation

Centromeres are the essential domain on every chromosome that allows spindle microtubules to attach to duplicated chromosomes and faithfully segregate them to two daughter cells. Centromere function is evolutionarily

highly conserved but the underlying centromeric DNA is neither conserved nor required for centromeres to function. Even though species-specific differences in centromere regulation exist, all eukaryotes except the parasitic trypanosomatids and some yeast species regulate centromeres epigenetically [1, 4, 42]. The best understood epigenetic determinant is a histone H3 variant at centromeres. This evolutionarily conserved protein is generally referred to as CENP-A, although other names are commonly used as well [35, 89].

CENP-A is a key factor of centromere identity and has been extensively reviewed elsewhere [4]. Loss of CENP-A prevents chromosomes from segregating correctly in many different organisms underlining the crucial function of centromeres for accurate development and health [18]. On the other hand, CENP-A overexpression, for instance in *Drosophila*, leads to the formation of ectopic centromeres that cause chromosome lagging and breaking due to multiple attachment points on the chromosome [50]. Importantly, in some tumours, a strong upregulation of CENP-A correlates with lower remission rates, making CENP-A a potential biomarker for predicting tumour aggressiveness [47, 61, 73, 91].

A distinct difference: centromeric and pericentromeric chromatin

Centromeres were long mistaken for a transcriptionally inert region of heterochromatin. However, there are several clear features which distinguish centromeric chromatin from its surrounding heterochromatin.

Centromeres are generally embedded in large blocks of constitutive pericentromeric heterochromatin that constitutes more than half of the genome, for instance in humans

✉ Sylvia Erhardt
s.erhardt@zmbh.uni-heidelberg.de

¹ Medical Research Council Clinical Sciences Centre, Imperial College London, London, UK

² ZMBH, DKFZ-ZMBH-Alliance and CellNetworks Excellence Cluster, University of Heidelberg, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany

or mice [69], and harbours typical marks of heterochromatin as discussed below. As already mentioned, the main characterizing feature of centromeric chromatin is the presence of the histone variant CENP-A. CENP-A nucleosomes are structurally distinct from the canonical histone H3-containing nucleosomes: they form a more rigid interface with H4 histones that is reshaped and stabilized by centromere protein C (CENP-C). These CENP-A-containing nucleosomes bind DNA less tightly than canonical nucleosomes and may therefore have a profound affect on chromatin structure and transcriptional activity of centromeric chromatin [38, 48]. An on-going debate about the actual composition of CENP-A-containing nucleosomes has been extensively reviewed elsewhere [13].

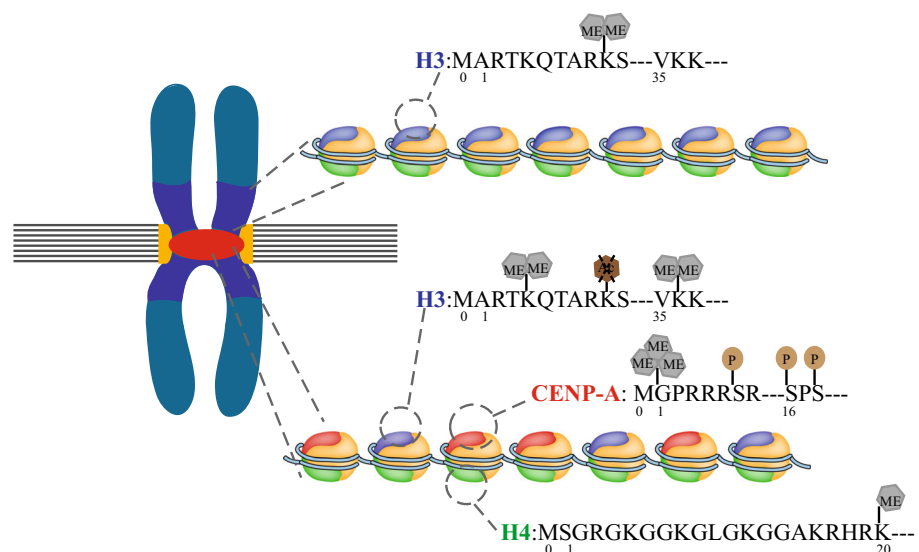
Besides CENP-A, another identifying characteristic of centromeres is the specific set of post-translation modifications within centromeric regions [87]. Centromeric chromatin is distinctly different from its surrounding pericentromeric heterochromatin by canonical histone modifications. Histone H3-containing nucleosomes at centromeric regions are interspersed with CENP-A-containing nucleosomes and contain some marks that are specific for transcriptionally active, open chromatin (e.g. histone H3 Lysine 4 and 36 methylation (H3K36me), H3K4me and H3K36me), other marks of open chromatin such as acetylation as well as heterochromatic marks (e.g. H3K9me) are missing, whereas H4K20me₁ is present (Fig. 1) [9, 52, 87, 95]. Not only the canonical histones bear modifications: CENP-A itself can be modified. Modifications identified so far include trimethylation of Gly1 and phosphorylation of Ser7, Ser16 and 18 in human cells, as well as mono-ubiquitylation in human and fly cells [5, 6, 45, 65]. Phosphorylation of Ser7 is required for correct mitotic progression, mono-ubiquitylation in flies is

required for CENP-A stability, and Arg37 acetylation in yeast regulates the recruitment of kinetochore components to centromeric chromatin [45, 80]. However, most of the growing list of CENP-A modifications still awaits to be functionally characterized.

Centromeric DNA: genetic meets epigenetics

While the epigenetic regulation of centromeres is widely accepted, this does not exclude a role for DNA (reviewed in Ohzeki et al. [67]). The repetitive nature of pericentromeric DNA sequences has been a major challenge to genome-wide sequencing projects in different species. Today, the human and *Drosophila* genome projects still lack a complete continuous sequence of (peri-) centromeric DNA (reviewed in [2]). But it is clear that centromeric DNA is neither conserved between species nor on different chromosomes within one species. Moreover, stable dicentric chromosomes and neocentromeres at non-centromeric sites form in organisms from yeast to human suggesting that centromeric DNA is dispensable for the formation of functional centromeres [42]. Interestingly, neocentromeres often form within or nearby euchromatic regions [7, 99, 100], and in fission yeast near telomeres [54]. Whether any requirements within the primary DNA sequence are pivotal for these preferences or whether the above discussed histone modification landscape at certain sites are more attractive for centromeric chromatin to form needs to be established in the future. One well-studied exception is budding yeast, *Saccharomyces cerevisiae*: here, centromeres are determined by the underlying DNA sequence. Their point centromeres, which consist of only one CENP-

Fig. 1 Histone tail modification associated with centromeres. Pericentromeric and centromeric chromatin are distinct from each other. Centromeric chromatin is characterized by the presence of the histone H3 variant CENP-A that replaces H3 in some of the centromeric nucleosomes. Histone modification pattern on canonical histones also vary between centromeric and pericentromeric chromatin



A containing nucleosome are 125 base pairs long and do not contain satellite repeats [49, 93].

Centromeric DNA has certain sequence characteristics, such as an AT-rich composition, and is repetitive in nature [3, 12, 81, 85, 88]. These characteristics may be a prerequisite for the formation of higher order structures at centromeres [62, 105]. The existence and conservation of higher order structures may be necessary for centromere function, but the sequence itself may vary. Interestingly, the length of one repeat unit of satellite DNA might be important for the formation of higher order structures, because different organisms have similar length of repeating units, that approximately correspond to one (178 bp in *Arabidopsis*, 156 bp in maize, 142 bp in beetles, 120 bp in mouse, 171 bp long α -satellite repeat in human) or two (359 bp long SAT III in *Drosophila*, 340 bp repeat in pig) nucleosomal unit lengths [49, 70] (Table 1). Together with the occupancy of CENP-A nucleosomes, data from ChIP-seq analysis and high resolution microscopy currently suggest several possible models for higher order structures that allow clustering of CENP-A nucleosomes on opposite sides of the sister chromatids [4, 14, 42, 75, 88].

There are several reports that showed that repetitive centromeric and pericentromeric DNA is important for cell integrity [71]. The presence of heterochromatin in pericentromeric regions is for instance required to ensure recruitment of cohesin, which in turn promotes bi-orientation of sister chromatids during mitosis whereas cohesion

at centromeric core chromatin induces a mono-orientation [11, 78, 103]. Centromeric DNA is also likely to participate in separating centromeres from surrounding (pericentromeric) regions: In *S. pombe*, centromeres are surrounded by tRNA genes, which physically prevent the expansion of pericentromeric heterochromatin into centromeres [82]. In *Drosophila*, de novo formation of kinetochores upon CENP-A overexpression occurs at euchromatin to heterochromatin transition regions ('boundaries'). These are regions that have a low nucleosome turnover and have been classified as transcriptionally silent, intergenic domains [68]. However, one should keep in mind that the here mentioned examples may require an active transcription of those repetitive sequences in order to fulfil their respective function, which will be discussed in detail in the next chapters.

Transcription of pericentromeric chromatin

Pericentromeric regions flanking centromeric chromatin on both sides are constituted of constitutive heterochromatin, which is characterized by its location at the same genomic regions throughout the development and life span of a cell. This is in contrast to facultative heterochromatin, which adjusts to developmental and external clues. Constitutive heterochromatin consists of tandem repeats that can stretch over mega bases, and, while it is largely devoid of protein-coding genes, transcription has been well-documented [46,

Table 1 Summary of known centromeric transcripts in different species

Species	Centromeric transcript	Repeat unit length/bp	Finding	Proposed interacting proteins	References
Budding yeast	<i>CEN</i> DNA		Certain level of centromeric transcription contributes to centromeric function		[66]
Fission yeast	Central kinetochore domain		Central domain is transcribed, transcripts are subjected to rapid turnover		[26]
Maize	CentC	156	Centromeric transcripts are constituents of the kinetochore complex	CENP-C	[33, 92]
<i>Drosophila</i>	SAT III	359	Centromeric transcripts are vital for kinetochore assembly and functional cell division	CENP-C	[76]
Tammar wallaby	crasiRNA		Overexpression of centromeric RNAs leads to mitotic defects		[17]
Mouse	Minor satellite	120	Centromeric transcripts are components of CPC crucial for Aurora B activity	CENP-A, CPC	[39]
Human	a-satellite	171	Knock-down causes mitotic defects.		[53]
			a-satellite is required for correct localization of CENP-C and INCENP throughout cell cycle	CENP-C	[101]
			Knock-down causes aberrant mitosis	Aurora B, CPC	[53]
			a-satellite is involved in timely incorporation of CENP-A	CENP-A	[74]

As far as known, their size and suggested functions are listed

77]. The most prominent example is pericentromeric heterochromatin formation in fission yeast. Pericentromeric heterochromatin in *S. pombe* is established and maintained with the help of small interfering RNAs (siRNA) generated from pericentromeric 'otr' transcripts [97]. RNA Polymerase II (RNAPII) transcribes 'otr' repeats bidirectionally from a conserved promoter within this region [31, 59]. The transcripts are processed by the RNA interference (RNAi) machinery, and siRNAs generated in this way are incorporated into the 'RNA induced transcriptional silencing' (RITS) complex. Through interaction of siRNAs with Ago1, a component of the RISC complex, the complex is loaded onto pericentromeric chromatin. The RITS complex then recruits the CLRC complex, which contains a methyltransferase responsible for methylation of H3K9. Finally, H3K9 dimethylation recruits fission yeast homolog of heterochromatin protein HP1, Swi6 [60]. In addition, RNAi signals are amplified by Ago1 recruitment of the RNA directed RNA polymerase complex (RDRC) to the siRNA precursor. Due to RDRC interaction with the RNAi machinery, through Dicer protein, more siRNA is produced [29, 86].

Functional importance of pericentromeric transcription and transcripts

In multicellular organisms pericentromeric transcription has been implicated in heterochromatin formation, but also in other processes [36]. In *Drosophila* ovaries, disruption of the RNAi machinery leads to accumulation of pericentromeric transcripts from chromosomes 2 and 3 [94]. Moreover, increased levels of acetylation of H3K9, and a component of the RNAPII transcriptional complex (TAF1) were observed on these pericentromeric repeats, indicating that transcription plays a role in heterochromatin formation via the RNAi machinery [94]. Depletion of Dicer in mouse embryonic stem cells leads to an accumulation of pericentromeric major satellite transcripts. These Dicer-depleted cells are viable and display no genomic instability, although cell differentiation is impaired [58, 64]. Interestingly, long pericentromeric transcripts that do not seem to be processed by the RNAi machinery also play a role in the formation of pericentromeric heterochromatin. Long transcripts that correspond to several repeating units of major pericentromeric satellites have been detected in mouse cells, and heterochromatin protein HP1 preferentially binds to the forward strand of these RNAs. The interaction occurs only after post-translational SUMOylation of HP1 [63]. The forward major RNA remains bound to the site of transcription, and through this interaction, HP1 is recruited to pericentromeric heterochromatin and stabilized by the presence of H3K9me₃. Additional HP1

molecules then accumulate, directly connecting heterochromatin formation with pericentromeric transcription in mouse cells. Major satellite transcripts also have been described to have a function in the formation of heterochromatic chromocenters in early mouse embryos. There, a burst in transcription of forward and reverse strands coincides with this process. Interestingly, disruption of transcription results in a failure of chromocenter formation and a developmental arrest [72]. In human cells, transcription of pericentromeric satellites, that are not α -satellite, has also been detected, and connected with a stress response to heat [37, 57], heavy metals, and oxidative species [96]. Importantly, levels of α -satellite and pericentromeric satellites are elevated in some types of cancer [90].

Transcription of centromeric chromatin: a general view

As reviewed above, centromeric chromatin is defined by its high concentration of CENP-A-containing nucleosomes and by its pattern of post-translational histone modifications that make centromeres permissive for transcription. In fact, centromeric transcription is not only possible, it seems to be required for accurate centromeric function (Fig. 2). In *S. cerevisiae*, the three transcription factors, Cbf1, Ste12 and Dig1 are known to bind to the core centromere, and depletion of these factors leads to disruption of centromeric function and genome instability. Functional centromeres can be restored by inducing centromeric transcripts from an artificial promoter [66]. In *S. pombe*, RNAPII transcribes the forward strand of centromeric repeats during S-phase, which is linked to the loading of RNAi machinery components to centromeres and heterochromatin formation of the surrounding regions [25]. A similar mechanism is found in *Drosophila*, where the ectopically tethered CENP-A loading factor CAL1 recruits RNAPII and the Facilitates Chromatin Transcription (FACT) remodeling complex, which then induces transcription at nearby sites [23]. In *S. pombe*, the situation is less clear: on one hand, incorporation of CENP-A protein in the central domain of centromeres seems to be incompatible with gene expression, since it silences genes inserted into this region [19]. On the other hand, non-coding RNAs from the centromeric central core have been described [26], and the disruption of centromere-bound transcription factor Ams2 leads to reduction of CENP-A levels at centromeres [24]. In conclusion, finely balanced centromeric transcription seems to be required for the correct loading of CENP-A; however the function of these transcripts remains to be determined.

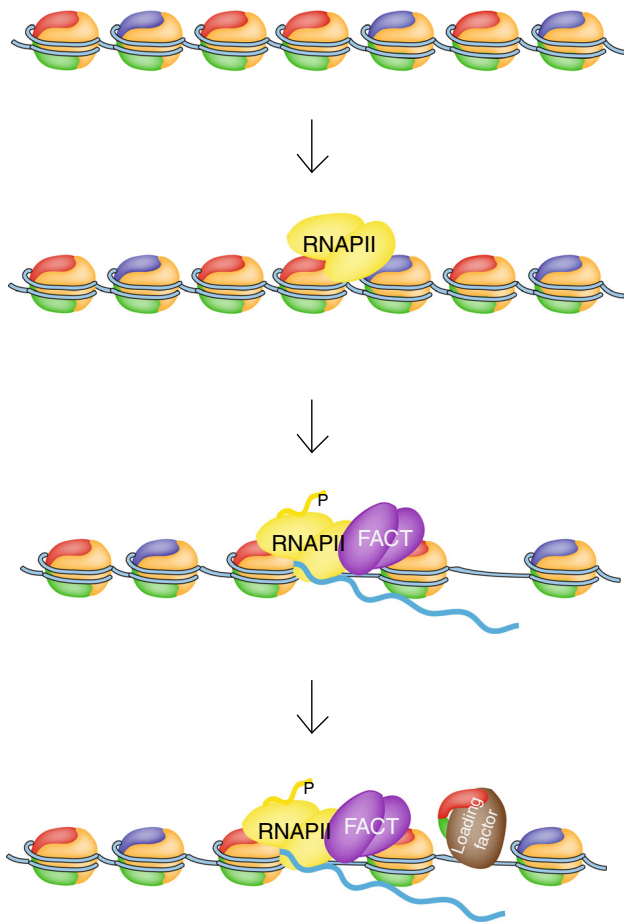


Fig. 2 Transcription from centromeric chromatin influences CENP-A loading in mitosis. The recruitment and activation of RNAPII at centromeres during mitosis may be required to recruit chromatin remodeling complexes such as FACT. FACT can then associate with the loading machinery of CENP-A and facilitates CENP-A loading that may coincide with the transcription of centromeric repeats. Alternatively, the transcription is required to recruit FACT and subsequently the loading machinery of CENP-A to centromeric regions

Transcription of centromeric chromatin: genes

Several publications report the presence of genes in centromeric regions. *S. pombe*, for instance, contains tRNA genes surrounding the centromeric loci. These genes are actively transcribed by RNAPIII, and function to prevent expansion of heterochromatin into centromeres [82]. Rice centromeres also contain active genes within the CENP-A binding region. These genes are expressed in several different tissues of rice, and display histone modifications that are typical for actively transcribed euchromatin (e.g. H3K4me₂). These modifications are, however, not a

ubiquitous component of rice centromeric chromatin, but a hallmark of the transcribed sequences embedded within the centromeric H3 subdomains. Hence, centromeric function and active gene expression are compatible in rice [104], but it is not yet clear whether transcription contributes to centromeric function.

Transcription of centromeric chromatin: repetitive sequences

As summarised in Table 1, centromeric repeats are transcribed in many different species and have many features in common but may also display species-specific differences. For instance in maize, satellite repeats called CentC are transcribed in both directions, and transcripts up to 900 bp long were found to immunoprecipitate with centromeric protein CENP-A [92], indicating a possible role in kinetochore assembly. In *Drosophila*, the active form of RNAPII is present at centromeres during mitosis [76]. Every *Drosophila* chromosome contains a different set of satellite repeats indicating that RNAPII presence is connected with the active centromere, rather than with a specific sequence [94]. Interestingly, the homeobox-containing transcription factor Hth is implicated in transcription of SAT III centromeric repeats in *Drosophila* embryos. Levels of SAT III RNA and localization of CENP-A to centromeres in Hth mutant embryos are strongly impaired, and result in very early developmental arrest [79], implying the importance of centromeric transcription in early *Drosophila* development. RNAPII is also enriched at the central domain of centromeric DNA in *S. pombe*. However, only low levels of transcription is detectable suggesting that in *S. pombe* the presence of stalled RNAPII, rather than a high transcription rate from the central domain, is required for CENP-A loading and centromere maintenance [20].

Minor satellite repeats on mouse centromeres consist of a basic repeating unit of 120 bp in length (Table 1). Under physiological conditions transcripts of up to 4 kb long can be detected. Interestingly, different kinds of stress lead to the accumulation of short transcripts of only one repeating unit by either cleaving longer transcripts or by producing a shorter product. If an exogenous vector is introduced that expresses the 120 bp long transcript, centromeric function is impaired, mitotic chromosomes misalign, and micronuclei accumulate as a result of failed chromosome segregation suggesting a vital role of minor satellite transcripts [15].

Transcription of centromeric chromatin: the act of transcription as regulatory mechanism during mitosis

Centromeric transcription plays an important role in human cells as well. Bergmann et al. [9] noticed a reduction of centromeric transcription when altering or removing H3K4me₂ marks that are associated with transcriptionally active chromatin. These changes lead to decreased loading of CENP-A and CENP-C proteins on human artificial chromosomes. In contrast, increasing acetylation on H3K9 at centromeric region, thereby creating a more open chromatin environment, results in a 150-fold increase of transcription that also impairs loading of CENP-A [10]. A report from Chan et al. [21] stressed the importance of active transcription at centromeres during mitosis in human cells. They detected the active form of RNAPII and its transcription factors at the centromeres of mitotic chromosomes. Addition of RNAPII inhibitors specifically in mitosis led to significant reduction of centromeric α -satellite transcripts and chromosome missegregation. Levels of centromeric CENP-C were also reduced, which may be what was responsible for the observed mitotic defects. Even more evidence for the importance of centromeric transcription comes from the human neocentromere on the model chromosome Mardel (10). This neocentromere lies on a region enriched in transcriptionally active LINE-1 retrotransposons. These LINE-1 transcripts are localized to the centromeric region, and depletion of these transcripts leads to reduced levels of CENP-A at the centromere, and mitotic defects [28].

The above mentioned studies support the hypothesis that RNAs from centromeric repeats are part of an epigenetic regulatory mechanism that participates in centromere function and regulation, whereby the act of transcription alone may be involved in centromere regulation. It is important to stress that active transcription at centromeres takes place during mitosis when overall transcription is mostly shut off by chromosome condensation and changes in chromatin modification [44]. The fact that active transcription at centromeres of different species occurs during mitosis suggests that, in general, transcription coincides with the functional active phase of a chromatin domain. The mitotic transcription also highlights the unique regulation of centromeric chromatin during mitosis. The correlation of active transcription and functional activity may also be extended to single copy genes. For instance, the mitotic regulator cyclin B is reported to be transcribed during mitosis [83]. Generally, transcription has been associated with many chromatin modification and nucleosome remodeling processes [4]. Hence, transcription at centromeres might be coupled to nucleosome exchange in

order to incorporate CENP-A into centromeres. A component of the FACT complex has been connected with CENP-A deposition in yeasts and flies [23, 27, 30] (Fig. 2). FACT destabilizes chromatin for RNAPII to pass through and reassembles it after transcription has taken place [8]. As a general chromatin-remodeling complex, FACT is also able to remove CENP-A from non-centromeric euchromatin and replace it with H3-containing nucleosomes in fission yeast, linking FACT to CENP-A regulation [27]. A more recent finding in budding yeast now shows that FACT, together with the E3 ubiquitin ligase Psh1, targets mislocalized CENP-A in euchromatin for degradation [30]. FACT components also interact with CENP-A nucleosomes in human cells [55], and more recent results showed that human centromeres are not only transcribed during mitosis [21], but also during G1 phase [74], coinciding with CENP-A deposition. The presence of the active form of RNAPII at centromeres at the time of CENP-A loading suggests an evolutionarily conserved mechanism that involves local centromeric transcription and chromatin remodeling for CENP-A loading and centromere organization.

Non-coding transcripts from centromeric regions: more than random transcription

Even though it appears that transcription itself plays an important role in centromere biology, there is considerable evidence that suggests that the actual nature of centromeric transcripts plays a vital role in centromere regulation. In maize, centromeric transcripts remain bound to the kinetochore after transcription [92], and are thought to participate in stabilization of kinetochore chromatin. Interestingly, addition of single-stranded RNA strongly promotes binding of CENP-C to DNA, regardless of the DNA sequence. However, excess RNA does not compete with DNA binding, and it is unlikely that CENP-C binds simultaneously to RNA and DNA, because the binding module of CENP-C is only 36 amino acids long [33]. The authors propose a transient interaction of CENP-C with RNA, resulting in a conformational change to the CENP-C structure that promotes DNA binding. RNA might therefore play a role similar to a protein chaperone to guide and ease binding and loading of the proteins to their correct locations (Fig. 3). CENP-C does not require its DNA binding domain to localize to the kinetochore, hence its targeting is likely to be dependent on protein interactions. Once at centromeres, CENP-C binding to DNA stabilizes its position, most likely with the help of centromeric RNA. This centromeric RNA probably acts in trans because the RNA that is associated with the kinetochore of a certain

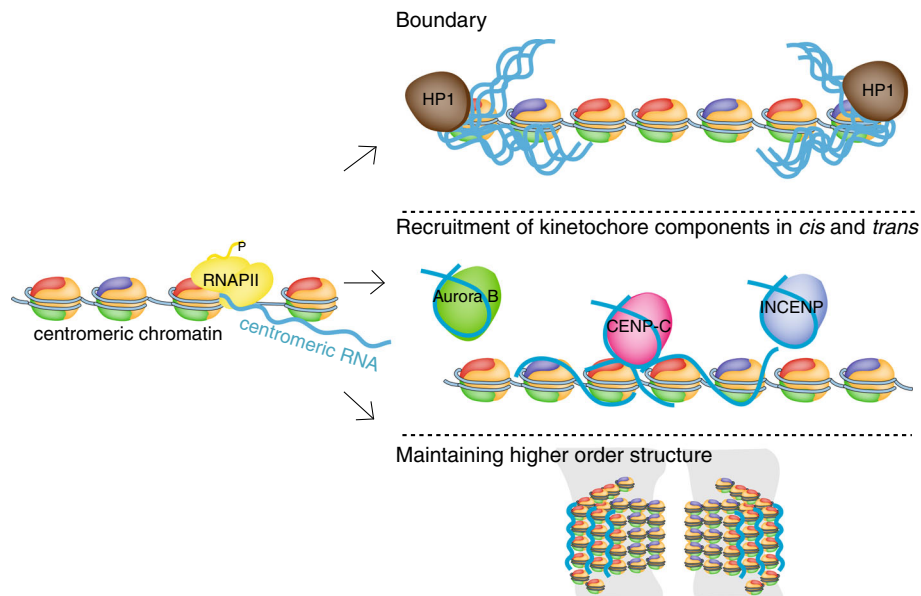


Fig. 3 Different potential functions of centromeric transcripts. Centromeric RNAs may bind to (or originate from) boundary elements to ensure that different chromatin states can coexist adjacent to each other by maintaining different chromatin identities. Different functional chromatin domains may be maintained this way. Centromeric RNAs have also been suggested to bind centromere-associated proteins and guide them to or stabilize them at centromeres. The guidance may occur in *cis* on the chromosome of origin of the RNA or in *trans* on centromeres that do not encode for the centromeric

RNA. Last but not least, centromeric transcripts may be part of a higher order chromatin structure that utilizes RNA as a scaffold to maintain a 3D structure crucial for centromere biology. Depicted is a previously suggested hypothetical structure of mitotic centromeres that folds CENP-C-containing domains into an outer platform for kinetochore components [4]. This structure may be established, stabilized or maintained by non-coding RNA from centromeric repeats

chromosome, is not necessarily encoded by the underlying centromeric DNA of the exact same chromosome [33].

In *Drosophila*, centromeric SAT III repeats are transcribed in sense and anti-sense orientations, and the longest detectable transcripts are approximately 1.3 kb long [76]. SAT III RNA localizes to mitotic centromeres and knockdown of SAT III leads to severe mitotic defects. Mitotic defects are also observed in flies with a deletion of the entire SAT III heterochromatin block on the X chromosome. Analysis of the lagging chromosomes showed that all major *Drosophila* chromosomes are affected by SAT III knock-down, even though this satellite locus is only present on the centromeric region of chromosome X. SAT III RNA, therefore, acts *in trans* on all major chromosomes in *Drosophila melanogaster*, similarly to what has been suggested for maize centromeric RNA described above. Interestingly, levels of centromere and kinetochore proteins are reduced on the lagging chromosomes, as well as the levels of newly deposited CENP-A and CENP-C. Furthermore, SAT III RNA immunoprecipitates with CENP-C, and CENP-C is required for centromeric localization of SAT III RNA. This indicates that SAT III RNA is a structural component of the *Drosophila* kinetochore, and is required for recruitment or stabilization of centromeric proteins [76]. The abundance and localization of SAT III

transcripts on centromeric regions of mitotic chromosomes suggests additional roles of centromeric transcripts in maintaining centromeric chromatin in general, for instance, by establishing or maintaining boundaries to neighboring heterochromatin or to maintain higher order chromatin structure at centromeres (Fig. 3).

Several studies focused on the role of centromeric transcripts in mammalian models. Knock-down of the transcribed murine centromeric minor satellite leads to chromosome segregation defects [53]. These transcripts associate with CENP-A, and with components of the Chromosomal Passenger Complex (CPC): Aurora B, Survivin and INCENP [39]. The CPC is a key regulator of major mitotic events, including chromosome–microtubule attachment error correction and the activation of the spindle assembly checkpoint [16]. Importantly, an RNA component is necessary for the observed interaction of Aurora B with CENP-A, and for the Aurora B kinase activity. Addition of minor satellite RNAs after RNaseA treatment rescues the kinase activity, strongly suggesting an active role of minor satellite RNA in the formation of centromere-associated complexes and mitotic progression [39].

Human α -satellites have also been linked to centromere function: Single-stranded RNA from human α -satellite is necessary for proper localization of the centromeric proteins

CENP-C and INCENP through binding of these RNAs to CENP-C. Interestingly, α -satellite transcripts mediate the localization of CENP-C and INCENP into the nucleolus in interphase cells, possibly for assembly into complexes with other proteins or for storage of these factors outside of mitosis and cytokinesis. CENP-C and INCENP are then relocated to centromeres during mitosis. RNase A treatment disrupts nucleolar localization of these proteins in interphase and their localization to centromeres during mitosis. Therefore, α -satellite RNA seems not only to participate in nucleolar sequestration of centromeric proteins, but also in their timely release to the kinetochore during mitotic. This complex has been termed nucleolar chaperone complex, and may include additional proteins [101].

Two recent reports describe the role of human centromeric transcripts in centromere regulation and chromosome segregation. Ideue et al. [53] found that knock-down of α -satellite (satellite I) transcripts induces aberrant mitosis and formation of ‘grape-shape’ nuclei. The knock-down of α -satellite RNA leads to an altered kinase activity of the RNA interacting protein Aurora B resulting in increased auto-phosphorylation of Aurora B, as well as an increase of phosphorylation of its substrate H3 at S10. Quenet and Dalal [74] addressed the role of centromeric transcripts in CENP-A loading. They found that a 1.3 kb long α -satellite transcript co-immunoprecipitates with CENP-A in the chromatin fraction and with the soluble CENP-A pre-assembly complex that contains the CENP-A-specific loading factor HJURP [34, 40, 74, 84]. The interaction of the pre-assembly complex with α -satellite transcripts suggests a role of the RNA in the timely incorporation of CENP-A. These studies indicate that not only the act of active transcription, but the actual centromeric transcripts, are crucial for the integrity of centromeric chromatin in human cells (Fig. 3). This, in turn, is in agreement with studies of SAT III transcripts in *Drosophila* cells [76]. However, further studies are required to understand the properties of repetitive transcripts from centromeric regions and their role in chromosome segregation.

Generally, depletion of centromeric transcripts leads to mitotic defects. However, in the tammar wallaby, overexpression of short (~40 nt) centromeric RNAs, called Centromere Repeat-Associated Short Interacting RNAs (crasiRNAs) also leads to mitotic defects, and a decrease in CENP-A signal in telophase/G1 [17]. Bouzinba-Segard et al. [15] observed that forced accumulation of small 120 bp long minor satellite transcripts results in mitotic defects. On the contrary, excessive amount of RNA does not interfere with CENP-C binding to DNA in maize, and overexpression of one repeating unit of SAT III RNA had no obvious effects on mitotic progression in *Drosophila* [33, 76]. Similarly, overexpression of α -satellite in human

cells also had no obvious effect [53]. The role of centromeric RNA overexpression certainly needs to be characterized further. It is possible, for instance, that crasiRNAs play a different role than long centromeric transcripts. Also, the accumulation of shorter (120 bp long) centromeric transcripts in mouse cells that occur under stress conditions may play a specific role in stress response [15]. A possible explanation for mitotic defects in this case comes from Aurora B kinase activity studies. When Aurora B and minor satellite RNA are present in equivalent amounts, the kinase activity is increased twofold. However, a higher amount of minor satellite RNA results in the reduction of the kinase activity [39]. The levels of minor satellite RNA is cell cycle dependent and suggests that a tightly regulated homeostasis of RNA and proteins is important for accurate cell cycle progression.

Some of the above-mentioned studies manipulate the level of centromeric transcripts post-transcriptionally, without disturbing the process of transcription, and therefore provide a clear distinction between a role of centromeric transcription itself and specific centromeric transcripts. Long centromeric transcripts seem to be an integral part of at least two different complexes that are essential for centromeric function. They interact with constitutive centromeric proteins to insure proper loading and stabilization of the position of these proteins at the centromeres [33, 74, 76], and they interact with the CPC complex and influence its localization, assembly, and enzymatic activity [39, 53, 101].

A question that arises is: which centromeric proteins directly interact with centromeric RNAs? To this end two proteins have been implicated: CENP-A and CENP-C. In maize, CENP-C interacts with single-stranded RNA to promote binding of the protein to DNA [33], and in *Drosophila*, SAT III RNA immunoprecipitates with CENP-C and is required for loading of both CENP-C, and CENP-A. Human CENP-C (CENPC1) contains the RNA-binding ‘HP1 hinge-homologous’ sequence that binds α -satellite RNA in vitro [101], a region that *Drosophila* CENP-C does not have a clear conservation to. On the other hand, Quenet and Dalal [74] reported immunoprecipitation of centromeric RNA with CENP-A. Since CENP-A and CENP-C are part of the same essential complex, it is not clear whether this represents a direct interaction, or a CENP-C-mediated interaction.

RNA-dependent mechanism of centromere regulation

Non-coding RNAs are involved in epigenetic regulation of various processes, including gene expression, transposon silencing, dosage compensation of sex chromosomes,

heterochromatin formation, and maintaining higher order chromatin structures [51] (Fig. 3). Single-stranded RNA has a flexibility that allows for the formation of secondary structures, such as stems and loops, which provide surfaces for interaction with other molecules. Proteins are usually folded into relatively rigid structures, with a definite numbers of sites that interact with DNA, RNA or other proteins. This creates a high degree of structural stability, which may be favorable for some cellular functions, for instance for catalytic activities. But in some cases, less rigidity or more structural flexibility may be preferable [102]. Proposed mechanism for RNA–protein interactions is the induction of secondary and tertiary structures: RNAs bind to proteins through one or several binding sites, the remaining RNA may then change its structure according to the bound protein. The binding of viral RNAs to the ribosome assembly is an example of an induced fold mechanism [32, 56]. Centromeric RNAs might fulfill their function by localizing factors that are required for the loading or function of CENP-A, as well as other kinetochore proteins, to centromeres. Alternatively, they may participate in the formation of protein complexes required for proper mitotic progression or maintenance in centromeric chromatin [4, 39, 53, 92, 101].

It is conceivable that centromeric RNAs facilitate *trans*-acting mechanisms. In other words, RNAs transcribed from one centromere may act on all or a subset of chromosomes independent of the primary DNA sequence. This may be important for mitotic progression since active transcription of satellite DNA takes place during mitosis in different species [21, 76]. It is important to point out that there may be a combination of non-coding RNAs from different repetitive regions present during mitosis. These RNAs could serve different functions or have redundant functions, a question that needs to be addressed in the future. Direct evidence for a *trans*-mechanism comes from *Drosophila*, where the knock-down of SAT III RNA, which originates only from chromosome X, equally affects all chromosomes [76]. A similar mechanism was proposed for maize centromeric RNA [33, 102]. Gent and Dawe [43] proposed that normal functioning of the neocentromeres that form on non-repetitive DNA might be possible due to the existence of redundant, *trans*-acting centromeric RNAs.

Conclusions and future perspectives

Centromeres are required for kinetochore formation to achieve attachment of the chromosomes to the mitotic spindle during mitosis. Despite this essential function, centromeres are not defined by the underlying DNA sequences; instead they are controlled by complex epigenetic mechanisms. Today we have a relatively clear picture of the

proteins involved in centromeric regulation, however recent studies propose that non-coding RNAs transcribed from centromeric regions are an additional epigenetic regulator of centromeres. While centromeric transcripts are crucial for centromere function and accurate cell division, the precise mechanisms that operate still remain to be determined.

There are numerous reports on transcription from both pericentromeric and centromeric chromatin, and it is important to make a distinction between these two types of transcripts, as they arise from two different chromatin environments, and appear to have different functions in the cell. The most important role of pericentromeric transcripts is the maintenance of a heterochromatic state of this region, which in turn provides sister chromatid cohesion, and bi-polar orientation during mitosis [78, 103]. The function of the centromeric transcripts seems to be related to stabilization and assembly of the protein complexes required for normal cell division. The function of pericentromeric and centromeric transcripts, and the exact mechanisms by which they act, seem to be complex. What is evident is that they are involved in the basic regulatory mechanism of crucial cellular processes.

Addressing the functional relevance of specific centromeric transcripts in comparison to the act of transcription itself will be a very important question to fully answer in order to understand the regulation of centromeres. Transcription of the centromeric regions is important for loading of CENP-A nucleosomes through chromatin opening [41, 55, 98], and the transcripts might provide a flexible scaffold that allows assembly or stabilization of the kinetochore proteins. The existence of RNA scaffolds also gives an explanation for the fast evolution of the centromeric DNA sequence. As long as the secondary structures of RNAs derived from the centromeres is maintained in a way that they can interact with centromeric proteins, the conservation of the exact DNA sequence is not required [22, 92]. Interactions of centromeric transcripts with proteins involved in cell division are just beginning to be elucidated, and this field has the potential to answer many long-standing questions regarding centromere biology, such as high variability of centromeric DNA, and the regulation of centromeric proteins' positioning and stabilization. Centromeric transcripts present an exciting research field, and by uncovering the mechanisms, we may find more surprises along the way.

Acknowledgments We apologize to our colleagues whose work on centromeres and repetitive RNA could not be cited due to space limitations. We gratefully acknowledge members of the Erhardt lab for fruitful discussions and Aubry K. Miller for constructive comments on the manuscript and the Deutsche Forschungsgesellschaft (ER576 and EXC81) and the Boehringer Ingelheim Stiftung for funding our research on centromeres and RNA. SR is an alumna of the Hartmut Hoffmann-Berling International Graduate School of Molecular and Cellular Biology.

References

1. Akiyoshi B, Gull K (2013) Evolutionary cell biology of chromosome segregation: insights from trypanosomes. *Open Biol* 3:130023
2. Aldrup-Macdonald ME, Sullivan BA (2014) The past, present, and future of human centromere genomics. *Genes* 5:33–50
3. Alfenito MR, Birchler JA (1993) Molecular characterization of a maize B chromosome centric sequence. *Genetics* 135:589–597
4. Allshire RC, Karpen GH (2008) Epigenetic regulation of centromeric chromatin: old dogs, new tricks? *Nat Rev Genet* 9:923–937
5. Bade D, Pauleau AL, Wendler A, Erhardt S (2014) The E3 ligase CUL3/RDX controls centromere maintenance by ubiquitinating and stabilizing CENP-A in a CAL1-dependent manner. *Dev Cell* 28:508–519
6. Bailey AO, Panchenko T, Sathyan KM, Petkowski JJ, Pai PJ, Bai DL, Russell DH, Macara IG, Shabanowitz J, Hunt DF, Black BE, Foltz DR (2013) Posttranslational modification of CENP-A influences the conformation of centromeric chromatin. *Proc Natl Acad Sci USA* 110:11827–11832
7. Barry AE, Howman EV, Cancilla MR, Saffery R, Choo KH (1999) Sequence analysis of an 80 kb human neocentromere. *Hum Mol Genet* 8:217–227
8. Belotserkovskaya R, Oh S, Bondarenko VA, Orphanides G, Studitsky VM, Reinberg D (2003) FACT facilitates transcription-dependent nucleosome alteration. *Science* 301:1090–1093
9. Bergmann J, Rodríguez M, Martins N, Kimura H, Kelly D, Masumoto H, Larionov V, Jansen L, Earnshaw W (2011) Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO J* 30:328–340
10. Bergmann JH, Jakubsche JN, Martins NM, Kagansky A, Nakano M, Kimura H, Kelly DA, Turner BM, Masumoto H, Larionov V, Earnshaw WC (2012) Epigenetic engineering: histone H3K9 acetylation is compatible with kinetochore structure and function. *J Cell Sci* 125:411–421
11. Bernard P, Schmidt CK, Vaur S, Dheur S, Drogat J, Genier S, Ekwall K, Uhlmann F, Javerzat JP (2008) Cell-cycle regulation of cohesin stability along fission yeast chromosomes. *EMBO J* 27:111–121
12. Biscotti MA, Olmo E, Heslop-Harrison JS (2015) Repetitive DNA in eukaryotic genomes. *Chromosome Res* 23:415–420
13. Black BE, Cleveland DW (2011) Epigenetic centromere propagation and the nature of CENP-a nucleosomes. *Cell* 144:471–479
14. Blower MD, Sullivan BA, Karpen GH (2002) Conserved organization of centromeric chromatin in flies and humans. *Dev Cell* 2:319–330
15. Bouzinba-Segard H, Guais A, Francastel C (2006) Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc Natl Acad Sci USA* 103:8709–8714
16. Carmena M, Wheelock M, Funabiki H, Earnshaw WC (2012) The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. *Nat Rev Mol Cell Biol* 13:789–803
17. Carone DM, Zhang C, Hall LE, Oberfell C, Carone BR, O'Neill MJ, O'Neill RJ (2013) Hypermorphic expression of centromeric retroelement-encoded small RNAs impairs CENP-A loading. *Chromosome Res* 21:49–62
18. Carroll CW, Straight AF (2006) Centromere formation: from epigenetics to self-assembly. *Trends Cell Biol* 16:70–78
19. Castillo AG, Mellone BG, Partridge JF, Richardson W, Hamilton GL, Allshire RC, Pidoux AL (2007) Plasticity of fission yeast CENP-A chromatin driven by relative levels of histone H3 and H4. *PLoS Genet* 3:e121
20. Catania S, Pidoux AL, Allshire RC (2015) Sequence features and transcriptional stalling within centromere DNA promote establishment of CENP-A chromatin. *PLoS Genet* 11:e1004986
21. Chan FL, Marshall OJ, Saffery R, Kim BW, Earle E, Choo KH, Wong LH (2012) Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proc Natl Acad Sci* 109:1979–1984
22. Chan FL, Wong LH (2012) Transcription in the maintenance of centromere chromatin identity. *Nucleic Acids Res* 40:11178–11188
23. Chen CC, Bowers S, Lipinski Z, Palladino J, Trusiak S, Bettini E, Rosin L, Przewloka MR, Glover DM, O'Neill RJ, Mellone BG (2015) Establishment of centromeric chromatin by the CENP-A assembly factor CAL1 requires FACT-mediated transcription. *Dev Cell* 34:73–84
24. Chen ES, Saitoh S, Yanagida M, Takahashi K (2003) A cell cycle-regulated GATA factor promotes centromeric localization of CENP-A in fission yeast. *Mol Cell* 11:175–187
25. Chen ES, Zhang K, Nicolas E, Cam HP, Zofall M, Grewal SI (2008) Cell cycle control of centromeric repeat transcription and heterochromatin assembly. *Nature* 451:734–737
26. Choi ES, Stralfors A, Castillo AG, Durand-Dubief M, Ekwall K, Allshire RC (2011) Identification of noncoding transcripts from within CENP-A chromatin at fission yeast centromeres. *J Biol Chem* 286:23600–23607
27. Choi ES, Stralfors A, Catania S, Castillo AG, Svensson JP, Pidoux AL, Ekwall K, Allshire RC (2012) Factors that promote H3 chromatin integrity during transcription prevent promiscuous deposition of CENP-A(Cnp1) in fission yeast. *PLoS Genet* 8:e1002985
28. Chueh AC, Northrop EL, Brettingham-Moore KH, Choo KH, Wong LH (2009) LINE retrotransposon RNA is an essential structural and functional epigenetic component of a core neocentromeric chromatin. *PLoS Genet* 5:e1000354
29. Colmenares SU, Buker SM, Buhler M, Dlakic M, Moazed D (2007) Coupling of double-stranded RNA synthesis and siRNA generation in fission yeast RNAi. *Mol Cell* 27:449–461
30. Deyter GM, 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
31. Djupedal I, Portoso M, Spahr H, Bonilla C, Gustafsson CM, Allshire RC, Ekwall K (2005) RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes Dev* 19:2301–2306
32. Doudna JA, Rath VL (2002) Structure and function of the eukaryotic ribosome: the next frontier. *Cell* 109:153–156
33. Du Y, Topp CN, Dawe RK (2010) DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. *PLoS Genet* 6:e1000835
34. Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, Daigo Y, Nakatani Y, Almouzni-Pettinotti G (2009) HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* 137:485–497
35. Earnshaw WC, Allshire RC, Black BE, Bloom K, Brinkley BR, Brown W, Cheeseman IM, Choo KH, Copenhagen GP, DeLuca JG, Desai A, Diekmann S, Erhardt S, Fitzgerald-Hayes M, Foltz D, Fukagawa T, Gassmann R, Gerlich DW, Glover DM, Gorbisky GJ, Harrison SC, Heun P, Hirota T, Jansen LE, Karpen G, Kops GJ, Lampson MA, Lens SM, Losada A, Luger K, Maiato H, Maddox PS, Margolis RL, Masumoto H, McAnish AD, Mellone BG, Meraldi P, Musacchio A, Oegema K, O'Neill RJ, Salmon ED, Scott KC, Straight AF, Stukenberg PT, Sullivan BA, Sullivan KF, Sunkel CE, Swedlow JR, Walczak CE, Warburton PE, Westermann S, Willard HF, Wordeman L, Yanagida M, Yen TJ, Yoda K, Cleveland DW (2013) Esperanto

- for histones: CENP-A, not CenH3, is the centromeric histone H3 variant. *Chromosome Res* 21:101–106
36. Eymery A, Callanan M, Vourc'h C (2009) The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *Int J Dev Biol* 53:259–268
 37. Eymery A, Horard B, El Atifi-Borel M, Fourel G, Berger F, Vitte AL, Van den Broeck A, Brambilla E, Fournier A, Callanan M, Gazzeri S, Khochbin S, Rousseaux S, Gilson E, Vourc'h C (2009) A transcriptomic analysis of human centromeric and pericentric sequences in normal and tumor cells. *Nucleic Acids Res* 37:6340–6354
 38. Falk SJ, Guo LY, Sekulic N, Smoak EM, Mani T, Logsdon GA, Gupta K, Jansen LE, Van Duyne GD, Vinogradov SA, Lampson MA, Black BE (2015) Chromosomes. CENP-C reshapes and stabilizes CENP-A nucleosomes at the centromere. *Science* 348:699–703
 39. Ferri F, Bouzinba-Segard H, Velasco G, Hube F, Francastel C (2009) Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Res* 37:5071–5080
 40. Foltz DR, Jansen LE, Bailey AO, Yates JR 3rd, Bassett EA, Wood S, Black BE, Cleveland DW (2009) Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* 137:472–484
 41. Foltz DR, Jansen LE, Black E, Bailey AO, Yates JR, Cleveland DW (2006) The human CENP-A centromeric nucleosome-associated complex. *Nat Cell Biol* 8:458–469
 42. Fukagawa T, Earnshaw WC (2014) The centromere: chromatin foundation for the kinetochore machinery. *Dev Cell* 30:496–508
 43. Gent JJ, Dawe RK (2012) RNA as a structural and regulatory component of the centromere. *Annu Rev Genet* 46:443–453
 44. Gottesfeld JM, Forbes DJ (1997) Mitotic repression of the transcriptional machinery. *Trends Biochem Sci* 22:197–202
 45. Goutte-Gattat D, Shuaib M, Ouarrhni K, Gautier T, Skoufias DA, Hamiche A, Dimitrov S (2013) Phosphorylation of the CENP-A amino-terminus in mitotic centromeric chromatin is required for kinetochore function. *Proc Natl Acad Sci USA* 110:8579–8584
 46. Hall LE, Mitchell SE, O'Neill RJ (2012) Pericentric and centromeric transcription: a perfect balance required. *Chromosome Res* 20:535–546
 47. Hanahan D, Weinberg RA (2000) The hallmarks of cancer. *Cell* 100:57–70
 48. Hasson D, Panchenko T, Salimian KJ, Salman MU, Sekulic N, Alonso A, Warburton PE, Black BE (2013) The octamer is the major form of CENP-A nucleosomes at human centromeres. *Nat Struct Mol Biol* 20:687–695
 49. Henikoff S (2013) The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 293:1098–1102
 50. Heun P, Erhardt S, Blower MD, Weiss S, Skora AD, Karpen GH (2006) Mislocalization of the *Drosophila* centromere-specific histone CID promotes formation of functional ectopic kinetochores. *Dev Cell* 10:303–315
 51. Holoch D, Moazed D (2015) RNA-mediated epigenetic regulation of gene expression. *Nat Rev Genet* 16:71–84
 52. Hori T, Shang WH, Toyoda A, Misu S, Monma N, Ikeo K, Molina O, Vargiu G, Fujiyama A, Kimura H, Earnshaw WC, Fukagawa T (2014) Histone H4 Lys 20 monomethylation of the CENP-A nucleosome is essential for kinetochore assembly. *Dev Cell* 29:740–749
 53. Ideue T, Cho Y, Nishimura K, Tani T (2014) Involvement of satellite I noncoding RNA in regulation of chromosome segregation. *Genes Cells* 19:528–538
 54. Ishii K, Ogiyama Y, Chikashige Y, Soejima S, Masuda F, Kakuma T, Hiraoka Y, Takahashi K (2008) Heterochromatin integrity affects chromosome reorganization after centromere dysfunction. *Science* 321:1088–1091
 55. Izuta H, Ikeno M, Suzuki N, Tomonaga T, Nozaki N, Obuse C, Kisu Y, Goshima N, Nomura F, Nomura N, Yoda K (2006) Comprehensive analysis of the ICEN (interphase centromere complex) components enriched in the CENP-A chromatin of human cells. *Genes Cells* 11:673–684
 56. Jain C, Belasco JG (2001) Structural model for the cooperative assembly of HIV-1 Rev multimers on the RRE as deduced from analysis of assembly-defective mutants. *Mol Cell* 7:603–614
 57. Jolly C (2004) Stress-induced transcription of satellite III repeats. *J Cell Biol* 164:25–33
 58. Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev* 19:489–501
 59. Kato H, Goto DB, Martienssen RA, Urano T, Furukawa K, Murakami Y (2005) RNA polymerase II is required for RNAi-dependent heterochromatin assembly. *Science* 309:467–469
 60. Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410:116–120
 61. Lacoste N, Woolfe A, Tachiwana H, Gareau AV, Barth T, Cantaloube S, Kurumizaka H, Imhof A, Almouzni G (2014) Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. *Mol Cell* 53:631–644
 62. Lam AL (2006) Human centromeric chromatin is a dynamic chromosomal domain that can spread over noncentromeric DNA. *Proc Natl Acad Sci* 103:4186–4191
 63. Maison C, Bailly D, Roche D, Oca RM, Probst AV, Vassias I, Dingli F, Lombard B, Loew D, Quivy J, Almouzni G (2011) SUMOylation promotes de novo targeting of HP1 alpha to pericentric heterochromatin. *Nat Genet* 43:220–227
 64. Murchison EP (2005) Characterization of Dicer-deficient murine embryonic stem cells. *Proc Natl Acad Sci* 102:12135–12140
 65. Niikura Y, Kitagawa R, Ogi H, Abdulle R, Pagala V, Kitagawa K (2015) CENP-A K124 ubiquitylation is required for CENP-A deposition at the centromere. *Dev Cell* 32:589–603
 66. Ohkuni K, Kitagawa K (2011) Endogenous transcription at the centromere facilitates centromere activity in budding yeast. *Curr Biol* 21:1695–1703
 67. Ohzeki J, Larionov V, Earnshaw WC, Masumoto H (2015) Genetic and epigenetic regulation of centromeres: a look at HAC formation. *Chromosome Res* 23:87–103
 68. Olszak AM, Essen D, Pereira AJ, Diehl S, Manke T, Maiato H, Saccani S, Heun P (2011) Heterochromatin boundaries are hotspots for de novo kinetochore formation. *Nat Cell Biol* 13:799–808
 69. Perrod S, Gasser SM (2003) Long-range silencing and position effects at telomeres and centromeres: parallels and differences. *Cell Mol Life Sci CMLS* 60:2303–2318
 70. Pezer Z, Ugarković D (2008) RNA Pol II promotes transcription of centromeric satellite DNA in beetles. *PLoS One* 3:e1594
 71. Pidoux AL, Allshire RC (2005) The role of heterochromatin in centromere function. *Philos Trans R Soc B Biol Sci* 360:569–579
 72. Probst AV, Okamoto I, Casanova M, Marjou FE, Marjou F, Baccon PL, Baccon P, Almouzni G (2010) A strand-specific burst in transcription of pericentric satellites is required for chromocenter formation and early mouse development. *Dev Cell* 19:625–638
 73. Qiu JJ, Guo JJ, Lv TJ, Jin HY, Ding JX, Feng WW, Zhang Y, Hua KQ (2013) Prognostic value of centromere protein-A

- expression in patients with epithelial ovarian cancer. *Tumour Biol* 34:2971–2975
74. Quenet D, Dalal Y (2014) A long non-coding RNA is required for targeting centromeric protein A to the human centromere. *eLife* 3:e03254
 75. Ribeiro SA, Vagnarelli P, Dong Y, Hori T, McEwen BF, Fukagawa T, Flors C, Earnshaw WC (2010) A super-resolution map of the vertebrate kinetochore. *Proc Natl Acad Sci USA* 107:10484–10489
 76. Rosic S, Kohler F, Erhardt S (2014) Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J Cell Biol* 207:335–349
 77. Saksouk N, Simboeck E, Dejardin J (2015) Constitutive heterochromatin formation and transcription in mammals. *Epigenet Chromatin* 8:3
 78. Sakuno T, Tada K, Watanabe Y (2009) Kinetochore geometry defined by cohesion within the centromere. *Nature* 458:852–858
 79. Salvany L, Aldaz S, Corsetti E, Azpiazu N (2009) A new role for hth in the early pre-blastodermic divisions in *Drosophila*. *Cell Cycle* 8:2748–2755
 80. Samel A, Cuomo A, Bonaldi T, Ehrenhofer-Murray AE (2012) Methylation of CenH3 arginine 37 regulates kinetochore integrity and chromosome segregation. *Proc Natl Acad Sci USA* 109:9029–9034
 81. Schueler MG, Higgins A, Rudd M, Gustashaw K, Willard H (2001) Genomic and genetic definition of a functional human centromere. *Science* 294:109–115
 82. Scott KC, Merrett SL, Willard HF (2006) A heterochromatin barrier partitions the fission yeast centromere into discrete chromatin domains. *Curr Biol* 16:119–129
 83. Sciortino S, Gurtner A, Manni I, Fontemaggi G, Dey A, Sacchi A, Ozato K, Piaggio G (2001) The cyclin B1 gene is actively transcribed during mitosis in HeLa cells. *EMBO Rep* 2:1018–1023
 84. Shuaib M, Ouararhni K, Dimitrov S, Hamiche A (2010) HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. *Proc Natl Acad Sci USA* 107:1349–1354
 85. Steiner N, Hahnenberger K, Clarke L (1993) Centromeres of the fission yeast *Schizosaccharomyces pombe* are highly variable genetic loci. *Mol Cell Biol* 13:4578–4587
 86. Sugiyama T, Cam H, Verdel A, Moazed D, Grewal SI (2005) RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc Natl Acad Sci USA* 102:152–157
 87. Sullivan BA, Karpen GH (2004) Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat Struct Mol Biol* 11:1076–1083
 88. Sun X, Wahlstrom J, Karpen G (1997) Molecular structure of a functional *Drosophila* centromere. *Cell* 91:1007–1019
 89. Talbert PB, Ahmad K, Almouzni G, Ausio J, Berger F, Bhalla PL, Bonner WM, Cande WZ, Chadwick BP, Chan SW, Cross GA, Cui L, Dimitrov SI, Doenecke D, Eirin-Lopez JM, Gorovsky MA, Hake SB, Hamkalo BA, Holec S, Jacobsen SE, Kamieniarz K, Khochbin S, Ladurner AG, Landsman D, Latham JA, Loppin B, Malik HS, Marzluff WF, Pehrson JR, Postberg J, Schneider R, Singh MB, Smith MM, Thompson E, Torres-Padilla ME, Tremethick DJ, Turner BM, Waterborg JH, Wollmann H, Yelagandula R, Zhu B, Henikoff S (2012) A unified phylogeny-based nomenclature for histone variants. *Epigenet Chromatin* 5:7
 90. Ting DT, Lipson D, Paul S, Brannigan BW, Akhavanfard S, Coffman EJ, Contino G, Deshpande V, Iafrate AJ, Letovsky S, Rivera MN, Bardeesy N, Maheswaran S, Haber DA (2011) Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* 331:593–596
 91. Tomonaga T, Matsushita K, Yamaguchi S, Oohashi T, Shimada H, Ochiai T, Yoda K, Nomura F (2003) Overexpression and mistargeting of centromere protein-A in human primary colorectal cancer. *Cancer Res* 63:3511–3516
 92. Topp CN, Zhong CX, Dawe RK (2004) Centromere-encoded RNAs are integral components of the maize kinetochore. *Proc Natl Acad Sci USA* 101:15986–15991
 93. Torras-Llort M, Moreno-Moreno O, Azorin F (2009) Focus on the centre: the role of chromatin on the regulation of centromere identity and function. *EMBO J* 28:2337–2348
 94. Usakin L, Abad J, Vagin VV, Pablos B, Villasante A, Gvozdev AV (2007) Transcription of the 1.688 satellite DNA family is under the control of RNA interference machinery in *Drosophila melanogaster* ovaries. *Genetics* 176:1343–1349
 95. Vakoc CR, Sachdeva MM, Wang H, Blobel GA (2006) Profile of histone lysine methylation across transcribed mammalian chromatin. *Mol Cell Biol* 26(24):9185–9195. doi:10.1128/MCB.01529-06
 96. Valgardsdottir R, Chiodi I, Giordano M, Rossi A, Bazzini S, Ghigna C, Riva S, Biamonti G (2007) Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Res* 36:423–434
 97. Volpe TA, Kidner C, Hall I, Teng G, Grewal S, Martienssen R (2002) Regulation of heterochromatic silencing and histone H3 Lysine-9 methylation by RNAi. *Science* 297:1833–1837
 98. Walfridsson J, Bjelring P, Thalen M, Yoo E, Park SD, Ekwall K (2005) The CHD remodeling factor Hrp1 stimulates CENP-A loading to centromeres. *Nucleic Acids Res* 33:2868–2879
 99. Warburton PE (2004) Chromosomal dynamics of human neocentromere formation. *Chromosome Res* 12(6):617–626
 100. Williams BC, Murphy TD, Goldberg ML, Karpen GH (1998) Neocentromere activity of structurally acentric mini-chromosomes in *Drosophila*. *Nat Genet* 18:30–37
 101. Wong LH, Brettingham-Moore KH, Chan L, Quach JM, Anderson MA, Northrop EL, Hannan R, Saffery R, Shaw ML, Williams E, Choo KH (2007) Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res* 17:1146–1160
 102. Wutz A (2003) RNAs templating chromatin structure for dosage compensation in animals. *BioEssays News Rev Mol Cell Dev Biol* 25:434–442
 103. Yamagishi Y, Sakuno T, Shimura M, Watanabe Y (2008) Heterochromatin links to centromeric protection by recruiting shugoshin. *Nature* 455:251–255
 104. Yan H, Ito H, Nobuta K, Ouyang S, Jin W, Tian S, Lu C, Venu RC, Wang GL, Green PJ, Wing RA, Buell CR, Meyers BC, Jiang J (2006) Genomic and genetic characterization of rice Cen3 reveals extensive transcription and evolutionary implications of a complex centromere. *Plant Cell* 18:2123–2133
 105. Zinkowski RP, Meyne J, Brinkley BR (1991) The centromere-kinetochore complex: a repeat subunit model. *J Cell Biol* 113:1091–1110