

REPORT

# Determinants of minor satellite RNA function in chromosome segregation in mouse embryonic stem cells

Yung-Li Chen<sup>1</sup>, Alisha N. Jones<sup>2</sup>, Amy Crawford<sup>3</sup>, Michael Sattler<sup>2,4</sup>, Andreas Ettinger<sup>1</sup>, and Maria-Elena Torres-Padilla<sup>1,5</sup>

The centromere is a fundamental higher-order structure in chromosomes ensuring their faithful segregation upon cell division. Centromeric transcripts have been described in several species and suggested to participate in centromere function. However, low sequence conservation of centromeric repeats appears inconsistent with a role in recruiting highly conserved centromeric proteins. Here, we hypothesized that centromeric transcripts may function through a secondary structure rather than sequence conservation. Using mouse embryonic stem cells (ESCs), we show that an imbalance in the levels of forward or reverse minor satellite (MinSat) transcripts leads to severe chromosome segregation defects. We further show that MinSat RNA adopts a stem-loop secondary structure, which is conserved in human  $\alpha$ -satellite transcripts. We identify an RNA binding region in CENPC and demonstrate that MinSat transcripts function through the structured region of the RNA. Importantly, mutants that disrupt MinSat secondary structure do not cause segregation defects. We propose that the conserved role of centromeric transcripts relies on their secondary RNA structure.

## Introduction

Correct chromosome segregation ensures equivalent distribution of genetic material in mitosis and meiosis (Marston, 2014; Tanaka and Hirota, 2016). The centromere is the primary constriction of mitotic chromosomes and a platform for kinetochore formation and spindle attachment (Westhorpe and Straight, 2014). Defects in centromere formation result in aberrant chromosome segregation, micronuclei formation, and chromosomal defects (Barra and Fachinetti, 2018). However, our understanding of the mechanisms ensuring faithful centromere formation across cell types remains limited.

Genetically, mammalian centromeres are composed of pericentric and centric domains. While the former is heterochromatic and the centric domain is euchromatic, both are transcriptionally active (Bury et al., 2020; Chan et al., 2012). In mammals, RNA polymerase II binds to centromeric DNA during mitosis (Bury et al., 2020; Chan et al., 2012). Unlike pericentric heterochromatin enriched in H3K9me3, core centromeres are characterized by CENPA nucleosomes and H3K4me2 (Barnhart et al., 2011; Black and Cleveland, 2011; Cleveland et al., 2003;

Fachinetti et al., 2013; Bergmann et al., 2011, 2012; Gopalakrishnan et al., 2009; Sullivan and Karpen, 2004). Non-coding centromeric transcripts have been linked to the role of CENPA chaperone, HJURP. In addition, centromeric transcripts have pivotal roles in kinetochore assembly and centromere stability (Chan and Wong, 2012; Djupedal et al., 2005; Ferri et al., 2009).

Transcription and transcripts from the core centromeric regions (centromere transcripts, cenRNA) play a role in recruiting kinetochore components such as chromosomal passenger complex (CPC complex) and inner kinetochore proteins CENPA and CENPC (Blower, 2016; Bobkov et al., 2018; Bouzinba-Segard et al., 2006; Chan et al., 2012; Ferri et al., 2009; Grenfell et al., 2016; Ideue et al., 2014; Jambhekar et al., 2014; McNulty et al., 2017; Quénet and Dalal, 2014; Rošić et al., 2014). Accordingly, a regulatory role of cenRNAs in initiating and/or maintaining functional kinetochore complexes has been proposed. However, how cenRNAs regulate centromere function remains unclear.

In mice, cenRNAs are transcribed from minor satellite (MinSat) repeats, which are stretches of highly similar units of

<sup>1</sup>Institute of Epigenetics and Stem Cells (IES), Helmholtz Munich, München, Germany; <sup>2</sup>Institute of Structural Biology, Molecular Targets and Therapeutics Center, Helmholtz Munich, Neuherberg, Germany; <sup>3</sup>Department of Chemistry, New York University, New York, NY, USA; <sup>4</sup>Department of Bioscience, Bavarian NMR Center, School of Natural Sciences, Technical University of Munich, Garching, Germany; <sup>5</sup>Faculty of Biology, Ludwig-Maximilians Universität, München, Germany.

Correspondence to Andreas Ettinger: [andreas.ettinger@helmholtz-munich.de](mailto:andreas.ettinger@helmholtz-munich.de); Maria-Elena Torres-Padilla: [torres-padilla@helmholtz-muenchen.de](mailto:torres-padilla@helmholtz-muenchen.de)

A.N. Jones's current affiliation is Department of Chemistry, New York University, New York, USA.

© 2024 Chen et al. This article is distributed under the terms of an Attribution–Noncommercial–Share Alike–No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms/>). After six months it is available under a Creative Commons License (Attribution–Noncommercial–Share Alike 4.0 International license, as described at <https://creativecommons.org/licenses/by-nc-sa/4.0/>).

~120 bp (Bouzinba-Segard et al., 2006; Hörz and Altenburger, 1981). Studies in various model organisms have indicated that cenRNAs may guide centromeric proteins to the centromeric region (Bobkov et al., 2018; Ferri et al., 2009; Ideue and Tani, 2020; Jambhekar et al., 2014; Quénet and Dalal, 2014; Rošić et al., 2014; Wong et al., 2007). Remarkably, however, centromeric sequences differ greatly between species and are one of the rapidly evolving elements in the genome (Henikoff et al., 2001). At the DNA level, a CENPB box with a consensus motif for the binding of CENPB is present in multiple species (Alkan et al., 2011; Masumoto et al., 1989). However, at the RNA level, the poor sequence similarity across species is insufficient to explain a shared mechanism of action.

Considering that conserved functions of non-coding RNAs emerge primarily from secondary structures rather than nucleotide sequences (Diederichs, 2014; Johnsson et al., 2014), we hypothesized that a conserved function for cenRNAs may be explained by structural conservation. Here, we investigated the molecular function of cenRNAs in chromosome segregation. We show that both, forward and reverse mouse MinSat transcripts affect chromosome segregation in murine embryonic stem cells (ESCs). MinSat RNA functions through a specific stem-loop secondary structure, which is conserved between human and mouse cenRNAs. Accordingly, human  $\alpha$ -satellite expression in mouse ESCs also results in chromosome segregation defects. Our data indicate that RNA secondary structure, specifically a stem-loop around the CENPB box sequence, is a conserved key feature of cenRNA transcripts. We propose that the conserved role of cenRNA in centromere function relies, at least partly, on its secondary RNA structure.

## Results and discussion

### Expression of mouse minor satellite transcripts leads to chromosome segregation errors

To investigate the basis of cenRNAs in centromere function, we investigated whether MinSat transcripts affect chromosome segregation in mouse ESCs. Since MinSat regions generate cenRNAs from both strands (Blower, 2016; Bury et al., 2020; Rošić et al., 2014), we asked whether forward, reverse, or both MinSat transcripts have functional roles in mitosis. First, we characterized MinSat transcripts using RT-qPCR to quantify non-strand specific (Fig. S1 A) as well as forward- and reverse-strand (Fig. S1 B) transcripts. Absolute threshold cycle (Ct)-values indicate that cenRNAs are expressed in non-synchronized ESCs (Fig. S1 B). To address whether MinSat transcripts are expressed throughout the cell cycle we used an ESC line harboring the Fucci system (Sakaue-Sawano et al., 2008) (Fig. S1 C). RT-qPCR analyses indicated that MinSat transcript levels increase during G2/M (Fig. 1, A and B), in agreement with previous observations in other cell types (Ferri et al., 2009). However, in mouse ESCs steady state MinSat RNA levels are maintained until G1 and decline during S phase (Fig. 1, A and B). This could be due to the known shorter G1 in ESCs (Coronado et al., 2013). We observed the same cell cycle pattern for both forward and reverse MinSat transcripts (Fig. 1, A and B),

but reverse transcripts displayed consistently lower Ct values, suggesting higher abundance than the forward ones.

To examine whether cenRNAs regulate chromosome segregation, we expressed one or five MinSat repeats in the forward or reverse orientation in ESCs (Fig. S1 D). We quantified chromosome bridges and lagging chromosomes in anaphase and telophase after confirming that the expression of both forward and reverse MinSat transcripts was efficient and specific (Fig. S1 D). Ectopic expression of MinSat transcripts led to a 10–20% increase in chromosome segregation defects (Fig. 1, C and D). The forward transcripts induced a higher proportion of errors compared with the reverse transcripts ( $26\% \pm 14$  for one repeat forward and  $13\% \pm 3$  for one repeat reverse). Expression of the forward MinSat repeat also led to increased chromosome segregation errors in mouse embryonic fibroblasts (MEFs) (Fig. S1, E and F). The effect on chromosome segregation was similar for one or five copies of MinSat repeats (Fig. 1 D). We observed a significant reduction of cell proliferation upon expression of reverse transcripts, for both one and five copies (Fig. 1 E). Importantly, we did not observe changes in the cell cycle in any condition (Fig. S1 G). Thus, an increase in MinSat transcript levels leads to chromosome segregation defects. Because the MinSat repeats are expressed from a heterologous promoter, which is unlikely to undergo genomic integration during the time frame of our experiments (Lim et al., 2023), the effects of MinSat transcripts in centromere function can occur, at least partly, in trans.

To investigate the potential mechanism whereby MinSat transcripts affect chromosome segregation, we asked whether ectopic expression affects the chromatin association of kinetochore proteins CENPA and CENPC. We isolated the chromatin fraction of cells arrested in mitosis and quantified chromatin-bound CENPC and CENPA normalized against H3 (Fig. 1, F and G). We did not detect consistent changes in chromatin association of CENPA (Fig. 1, F and G) but chromatin levels of CENPC were significantly reduced upon the expression of five reverse MinSat repeats (Fig. 1 G). Expression of either one repeat reverse or five repeats forward also led to a mild reduction of CENPC chromatin association, which was consistent across experiments but not statistically significant (Fig. 1 G). These observations suggest that CENPC binding to chromatin during mitosis is reduced upon expression of MinSat reverse transcripts. We did not observe changes in CENPC upon MinSat expression in whole cell lysates (Fig. 1 H), suggesting that the effect of MinSat transcripts is specific to the chromatin association of CENPC. Interestingly, we observed reduced total cellular levels of CENPA after overexpression of reverse MinSat transcripts and, to a lesser extent with the one repeat forward (Fig. 1 I). While this reduction was not statistically significant, we observed the strongest effect upon expression of one and five reverse repeats suggesting a potential role of MinSat RNAs in modulating cellular CENPA levels. Thus, we concluded that forward and reverse cenRNAs can function during mitosis in mouse ESCs and that MinSat expression, particularly the reverse transcript, affects CENPC association with mitotic chromatin.

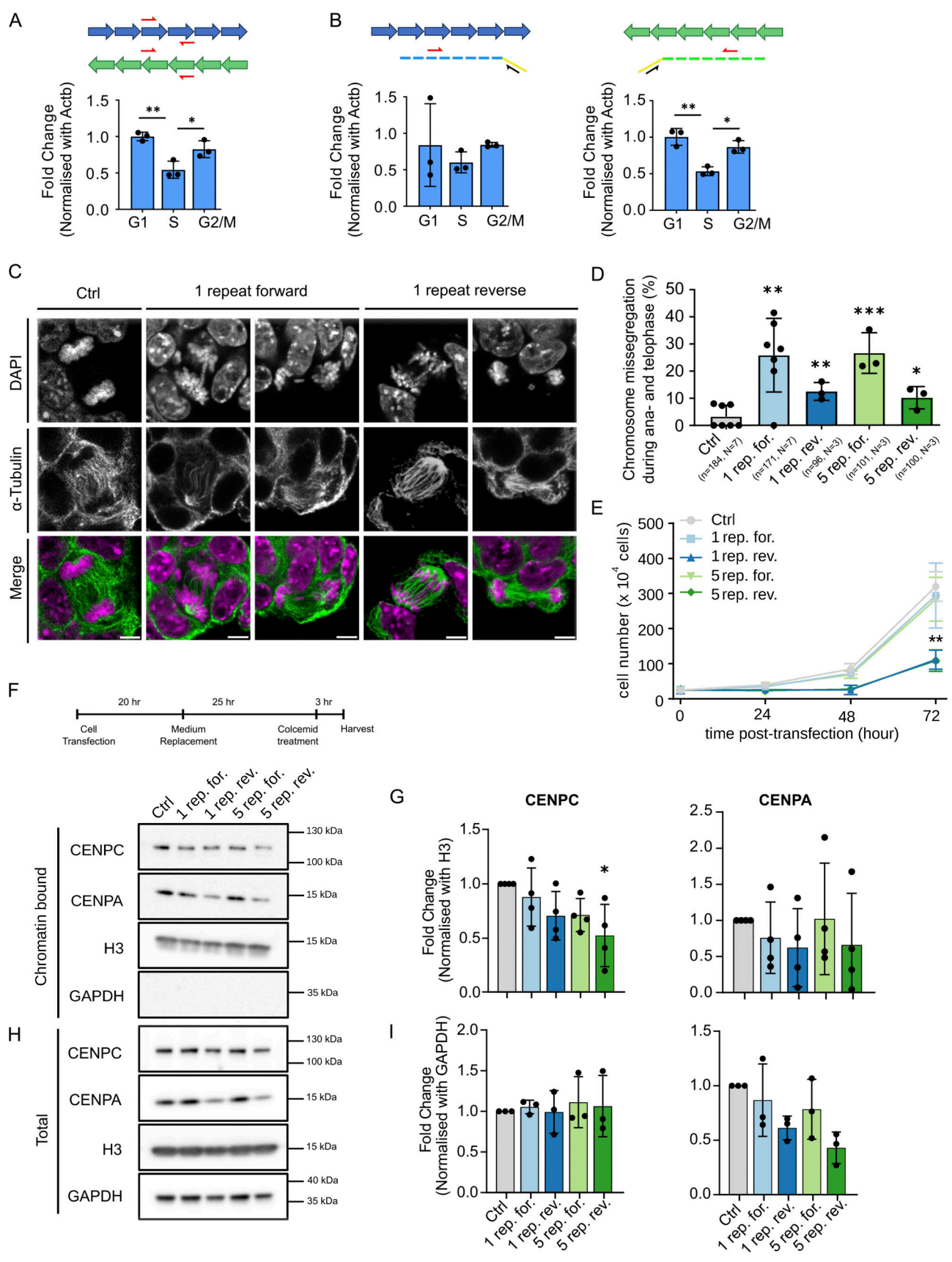


Figure 1. **Ectopic expression of forward and reverse minor satellite transcripts leads to chromosome segregation defects in ESCs. (A and B)** RT-qPCR analysis of total (A) and forward and reverse (B) MinSat transcripts in G1-, S-, and G2/M-phase populations of ESCs harboring the Fucci cell cycle reporter

system. The schematics on top indicate the position of the PCR primers used and the strategy for the analysis of strand-specific transcripts. Bars, mean value of  $N = 3$  independent experiments (individual dots), normalized to  $\beta$ -actin mRNA and control sample; error bars, standard deviation; P values \*  $\leq 0.05$ , \*\*  $\leq 0.01$ . We note that reverse transcripts consistently displayed lower Ct values ( $\sim 24.5$ ) than forward transcripts ( $\sim 26$ ). **(C)**  $\alpha$ -Tubulin immunofluorescence and DAPI staining of ESCs expressing one repeat MinSat RNAs. Ctrl, empty pCAG vector control. Scale bars, 5  $\mu\text{m}$ . **(D)** Percentage of chromosome missegregation events after MinSat expression in total anaphase to telophase cells. Bars, mean; error bars, standard deviation; pair-wise comparisons with control (empty pCAG vector) P values \*  $\leq 0.05$ , \*\*  $\leq 0.01$ , \*\*\*  $\leq 0.001$ . The number of biological replicates ( $N$ ) and the total number of mitotic figures analyzed ( $n$ ) are indicated. **(E)** Growth curves of ESCs after MinSat RNA expression. Lines indicate the mean of three independent biological replicates; error bars, standard deviation; P value \*\*  $\leq 0.01$ . **(F)** Timeline diagram for sample collection and representative immunoblot for the chromatin fraction of ESCs after expression of one or five repeats of forward or reverse MinSat RNAs, or empty pCAG vector control (Ctrl). Cells were synchronized by colcemid treatment before harvesting as indicated. The WB was repeated three or four times with independent lysates, corresponding to the datapoints on the quantification (G). **(G)** Immunoblot quantification. Values are fold changes normalized to histone H3 and control group in each individual replicate. Bars, mean; error bars, standard deviation; \* P value  $\leq 0.05$ . Statistical test: ANOVA. **(H)** Representative immunoblot for the total lysate of ESCs after expression of one or five repeats of forward or reverse MinSat RNAs, or empty pCAG vector control (Ctrl). Cells were synchronized by colcemid treatment before harvesting as indicated. The WB was repeated three or four times with independent lysates, corresponding to the datapoints on the quantification (I). **(I)** Immunoblot quantification. Values are fold changes normalized to GAPDH and control group in each individual replicate. Bars, mean; error bars, standard deviation. Statistical test: ANOVA. Source data are available for this figure: SourceData F1.

### Forward and reverse minor satellite transcripts are required for fidelity of chromosome segregation

We next asked if endogenous MinSat transcripts also regulate centromere function. We depleted specifically forward or reverse transcripts, or both, using antisense oligos (ASOs). The reverse ASO strongly reduced global MinSat transcript levels, similar to combined forward and reverse ASOs (Fig. 2 A). However, the forward ASO did not significantly change global MinSat transcript levels (Fig. 2 A), but it did significantly reduce the forward transcript levels (Fig. 2 B). Interestingly, strand-specific targeting of MinSat with ASOs led to a slight increase in transcript levels of the opposite strand (Fig. 2, B and C), particularly in the case of the reverse ASO, which led to a 1.32 ( $\pm 0.3$ )-fold increase in forward MinSat transcripts (Fig. 2 B). These results explain the apparent unchanged global MinSat transcript levels in response to the forward ASO transfection (Fig. 2 A). This points toward a feedback loop between cenRNAs. In addition, since the downregulation effect of the forward ASO on the forward transcripts is masked by the levels of the reverse MinSat, these results suggest that reverse MinSat transcripts are more abundant than the forward transcripts.

Having established conditions for the efficient depletion of endogenous MinSat transcripts, we next investigated effects on chromosome segregation by counting segregation errors in mitotic cells and micronuclei in interphase cells. Downregulation of either forward or reverse MinSat transcripts significantly increased chromosome segregation defects (23%  $\pm$  12 for forward ASO and 31%  $\pm$  16 for reverse ASO compared with 3%  $\pm$  4 in control ASO-treated ESCs; Fig. 2, D and E). We observed a similar phenotype when both forward and reverse ASOs were combined (Fig. 2, D and E). In interphase cells, the percentage of cells with micronuclei increased from 1% to 10–20% in ASO-treated ESCs (10%  $\pm$  3 for forward ASO and 16%  $\pm$  2 for reverse ASO, Fig. 2, F and G). These experiments indicate that endogenous MinSat transcripts are required for correct chromosome segregation.

To address whether depletion of MinSat transcripts leads to a similar molecular phenotype as upon ectopic expression, we analyzed CENPC and CENPA binding to mitotic chromatin. We did not observe consistent changes in either the binding of CENPA to chromatin (Fig. 2, H and I) or in total levels of CENPA (Fig. 2, J and K). However, we detected a consistent, albeit not

statistically significant, reduction of chromatin-bound CENPC when combining both forward and reverse ASOs (Fig. 2, H and I). Global levels of CENPC were also mildly reduced in all ASO conditions (Fig. 2 K), which we confirmed by immunofluorescence (Fig. 2 L). Overall, these data reveal minor but consistent changes in CENPC association with the chromatin, suggesting that MinSat transcripts may contribute to CENPC regulation and chromatin binding. Our data suggest that a feedback mechanism between transcripts from both strands ensures their respective levels throughout the cell cycle. Together with our observations above (Fig. 1), our data suggest that a balanced level of cenRNAs is essential for faithful chromosome segregation in mouse ESCs.

### Minor satellite transcripts are structured and interact with RNA-binding regions in CENPC

CENPC is a core centromere component with structural and regulatory functions during kinetochore assembly (Klare et al., 2015; Kwon et al., 2007; Przewlaka et al., 2011). Depletion of CENPC negatively impacts mitosis progression (Chik et al., 2019; Milks et al., 2009; Moree et al., 2011). Although not detectable across all experimental conditions, the changes of CENPC association to mitotic chromatin upon expression and depletion of MinSat RNAs prompted us to investigate a potential relationship between CENPC and cenRNAs. Indeed, CENPC is a good candidate to mediate interactions at the interphase of centromeric DNA, RNA, and kinetochore effectors.

Human cenRNA,  $\alpha$ -satellite, binds to human CENPC in *in vitro* gel shift assays (Wong et al., 2007), similar to the maize centromeric repeats, CentC (Du et al., 2010), suggesting a potential conserved binding between CENPC and cenRNAs. However, the lack of sequence similarity of centromeric DNA between species against the relatively strong CENPC conservation raises the question of whether such binding is conserved and involved in centromere function. Thus, we first tested if mouse CENPC and MinSat RNAs interact using RNA pull-down. We *in vitro*-transcribed and biotinylated one repeat of MinSat forward or reverse transcripts and *Egfp* RNA as negative control. After incubation with whole ESC lysate, we probed for CENPC with Western blotting. Our data indicate that CENPC binds to the forward transcript and to a lesser extent to the reverse transcript (Fig. 3 A), pointing to a direct interaction of the



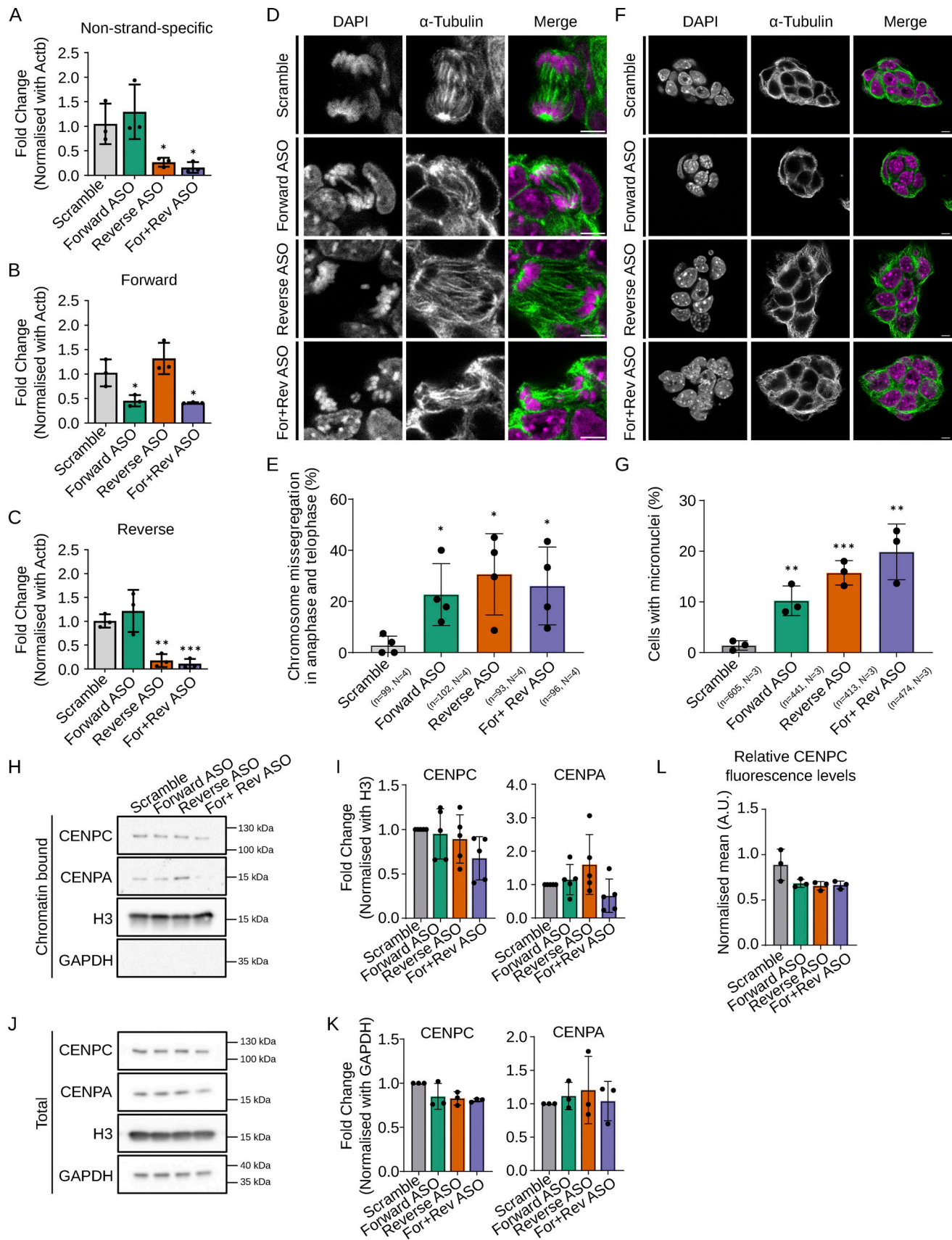


Figure 2. **Depletion of forward or reverse MinSat transcripts impairs chromosome segregation.** (A–C) RT-qPCR analysis of total (A), forward (B), and reverse (C) MinSat transcripts in mouse ESCs transfected with ASO against forward (green), reverse (orange), or both strands (purple), respectively. Scramble

(gray), non-targeting ASO control; bars, mean values of  $N = 3$  independent experiments (dots), normalized to  $\beta$ -actin mRNA and control samples; error bars, standard deviation; P values  $* \leq 0.05$ ,  $** \leq 0.01$ ,  $*** \leq 0.001$ . **(D)** Representative images showing ESCs stained with  $\alpha$ -tubulin and DAPI. Cells were transfected with ASOs against MinSat RNAs. Scale bars, 5  $\mu\text{m}$ . **(E)** Quantification of chromosome missegregation events in mouse ESCs transfected with the indicated ASOs. Bars, mean percentage of aberrant chromosome segregation (dots indicate individual experiments); error bars, standard deviation; P values  $* \leq 0.05$  (pair-wise comparisons to control samples). The number of biological replicates ( $N$ ) and total number of mitotic events analyzed ( $n$ ) is indicated. **(F)** Representative images showing ES cells stained with  $\alpha$ -tubulin and DAPI. Cells were transfected with ASOs against MinSat RNAs. The presence of micronuclei was analysed in interphase cells. Scale bars, 5  $\mu\text{m}$ . **(G)** Quantification of micronuclei in mouse ESCs transfected with the indicated ASOs. Bars indicate the mean percentage of micronuclei per cell (dots indicate individual experiments); error bars, standard deviation; P values  $** \leq 0.01$ ,  $*** \leq 0.001$  (pair-wise comparisons to control samples). The number of biological replicates ( $N$ ) and the total number of cells analyzed ( $n$ ) is indicated. **(H)** Representative immunoblot of the chromatin fraction of ESCs after ASO-mediated knockdown of MinSat RNAs and colcemid synchronization using antibodies to detect the indicated proteins. Scramble, non-targeting ASO control. **(I)** Immunoblot quantification. Values are fold changes normalized to histone H3 and control in each individual replicate. Bars, mean of  $N = 5$  independent biological replicates (dots); error bars, standard deviation. Statistical test: ANOVA. **(J)** Representative immunoblot of the total lysate of ESCs after ASO-mediated knockdown of MinSat RNAs and colcemid synchronization using antibodies to detect the indicated proteins. Scramble, non-targeting ASO control. **(K)** Immunoblot quantification. Values are fold changes normalized to GAPDH and controlled in each individual replicate. Bars, mean of  $N = 3$  independent biological replicates (dots); error bars, standard deviation. Statistical test: ANOVA. **(L)** Immunofluorescent intensity quantification of CENPC. Data points are means of all sampled CENPC intensity normalized with DAPI signal in that specific Z plane. Scramble, non-targeting ASO control. Bars, mean of  $N = 3$  independent biological replicates (dots); error bars, standard deviation. Source data are available for this figure: SourceData F2.

murine CENPC-cenRNA. This interaction is specific since we did not detect CENPC in the *Egfp* RNA pulldown (Fig. 3 A). To identify potential RNA binding region(s) within CENPC we subjected CENPC to in silico RNA binding motif prediction (catRAPID) (Livi et al., 2016). catRAPID identified 5 potential binding regions, which are numbered based on the potential RNA binding ability, with 1 corresponding to the highest (Fig. 3 B). While the catRAPID domains do not fully recapitulate known functional domains of CENPC, they do exhibit an overlap (Fig. 3 C). To test their ability for RNA binding, we prepared recombinant proteins of these regions and incubated each with in vitro transcribed forward MinSat RNA. In line with the ranking of the catRAPID prediction, regions 1, 2, and 3 display RNA binding activity in the micromolar range, whereas no binding is observed in regions 4 and 5 (Fig. 3 D). Incubating increasing protein concentrations indicated that region 1 has the highest RNA binding affinity (Fig. 3 D). This region is uncharacterized but overlaps with the central domain previously characterized in human CENPC (Carroll et al., 2010; Kato et al., 2013) and with the RNA-binding region that interacts with  $\alpha$ -satellite transcripts based on amino acid sequence similarity (Fig. 3 C and Fig. S2 A) (Wong et al., 2007). Thus, we conclude that mouse CENPC interacts with MinSat RNA.

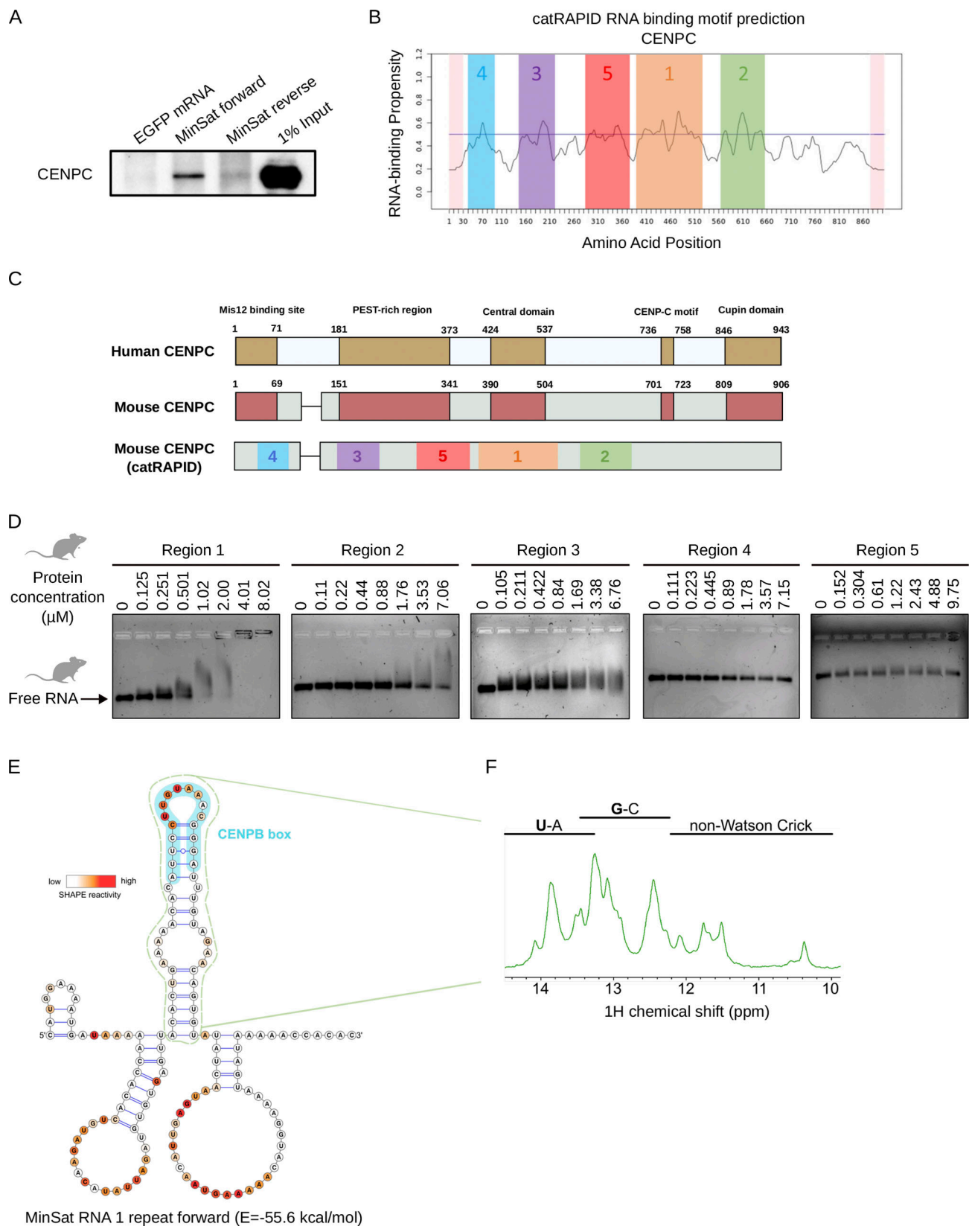
As RNAs can be structured, we next performed selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE) (Merino et al., 2005) on the mouse forward core 120-nucleotide MinSat repeat. From the SHAPE-predicted structure, we identified a stem-loop in the single repeat RNA (Fig. 3 E) even in higher energy predicted structures (Fig. S2 B). This structure occurs repeatedly along the naturally repetitive MinSat sequence (Fig. S2 C). The prominent stem-loop (Fig. 3 E, green dotted) contains the CENPB box motif (Fig. 3 E, light blue), a conserved DNA sequence that provides binding specificity of CENPB to centromeric DNA (Iwahara et al., 1998; Masumoto et al., 1989; Suntronpong et al., 2016), but with no documented roles at the RNA level. Furthermore, the presence of numerous imino signals in 1D nuclear magnetic resonance (NMR) spectra of the RNA motif is consistent with the formation of a well-defined secondary structure (Fig. 3 F). We therefore conclude that MinSat transcripts are structured and that the CENPB box motif adopts a well-defined, stable RNA stem-loop structure.

### Structural conservation rather than sequence motifs underlie segregation effects of MinSat RNAs

Given the presence of a stable stem-loop structure, we hypothesized that such a structural motif might provide a basis for a conserved function of cenRNAs. We first asked whether the MinSat RNA secondary structure is conserved between mouse and human cenRNA. Specifically, we used the  $\alpha$ -satellite sequence from human chromosome 21 (hereafter referred to as ASAT), which displays only 53% of DNA sequence identity to the murine MinSat (Fig. 4 A). SHAPE on the human ASAT revealed that the forward ASAT RNA adopts a secondary structure with a 24 nt stem-loop (Fig. S3 A, green dotted) containing the CENPB box motif (Fig. 4 B and Fig. S3 A, light blue), similar to the mouse MinSat RNA. The CENPB box coincides with a highly probable loop region in the human ASAT forward transcript (Fig. 4 B). Thus, we conclude that despite their low nucleotide sequence similarity, mouse and human forward cenRNAs adopt a similar secondary stem-loop structure, suggesting structural conservation.

We reasoned that if this newly identified structural motif underlies the cenRNA-CENPC interaction, we could formulate three predictions. First, human ASAT RNA should be able to interact with mouse CENPC. We incubated human ASAT RNA with mouse CENPC catRAPID regions 1–5. Similar to mouse MinSat, we observed a clear mobility shift when human ASAT RNA was incubated with the high-affinity binding region 1 (Fig. 4 C). The affinity of the human ASAT RNA for the other catRAPID regions was lower compared with region 1 (Fig. 4 C). These data indicate that human cenRNAs can interact with mouse CENPC protein and that this interaction is mediated by a specific region in CENPC.

Second, ectopic expression of the human ASAT RNA in mouse cells should lead to similar chromosome segregation phenotypes as mouse MinSat. We expressed one repeat of the forward or the reverse ASAT transcripts in mouse ESCs and confirmed efficient expression (Fig. S3 B). In line with our hypothesis, we observed an increase in chromosome segregation errors upon expression of both forward ( $28\% \pm 8$ ) and reverse ( $33\% \pm 11$ ) ASAT RNA compared with empty vector (Fig. 4, D and E). However, our result that the reverse ASAT transcript induces a similar phenotype suggests that, while the secondary structure



**Figure 3. Minor satellite transcripts interact with RNA-binding regions of CENPC and are structured. (A)** Immunoblot for CENPC after RNA pull-down from whole cell lysate of ESCs using biotin-labeled RNAs. The biotinylated RNAs used are indicated on top of the gel image, EGFP was used as negative control. **(B)** Prediction of RNA binding regions in CENPC using CatRapid. The X axis represents amino acid sequence position. The predicted RNA binding regions are highlighted in color and are ranked correspondingly to their predicted RNA binding ability, as indicated by numbers. **(C)** Schematic of the functional domains characterized in human CENPC and their correspondence in mouse CENPC, which we assigned based on the alignment of Fig. S2 A. **(D)** EMSA of the predicted



RNA binding regions in CENPC and 1 repeat of forward MinSat RNA. Protein concentrations are indicated on the top. **(E)** Secondary structure of 1 repeat of forward MinSat RNA obtained using SHAPE. SHAPE reactivity is revealed using color gradient on the top left. The green outlined stem-loop (position 51–96) is predicted as a thermodynamically stable region. The light blue outlined region corresponds to the CENPB box motif (position 65–81). The numbers indicate the nucleotide position relative to the start of the primer extension reaction. **(F)** 1D imino NMR spectrum of the MinSat RNA stem-loop (corresponding to the green dotted-lined region in Fig. 3 D). Source data are available for this figure: SourceData F3.

contributes to the role of cenRNAs in centromere function, it may not be the sole determinant. Alternatively, it is possible that the reverse and/or forward ASAT transcripts compete for interaction with additional potential effectors.

Third, mutations that affect the structure but not necessarily the CENPB box, should elicit chromosome segregation defects. We generated three different mutants: (1) a stem-loop without the CENPB box motif (del); (2) a stem-loop in which we replaced the apical loop with a GNRA tetraloop (GNRA). and (3) a stem-loop in which we swapped the CG base pairing in the stem region of the stem-loop (swap) (Fig. 5 A). The mutants are thus as follows: (1) with both the CENPB box motif and the stem-loop structure disrupted (del); (2) with CENPB box intact but the apical loop structure disrupted (GNRA), and (3) with the stem-loop structure maintained but without the CENPB box motif (swap) (Fig. 5 B). We expressed each mutant individually in mouse ESCs with comparable expression levels (Fig. 5 C). From the three stem-loop mutants we tested, only the swap mutant led to chromosome segregation errors to a similar extent as the wild-type transcript (Fig. 5, D and E). The MinSat mutants in which the CENPB box sequence is removed or replaced by an alternative sequence capable of adopting a loop structure behave similarly to control ESCs (Fig. 5, D and E). Thus, only the MinSat RNA which retains the apical loop structure led to increased chromosome missegregation. Of note, the essential nucleotides known to be important for CENPB box function and highly conserved at the DNA level (Suntronpong et al., 2016; Tanaka and Hirota, 2016) are fully replaced in the swap construct (Fig. 5 A). This implies that rather than functioning as a potential consensus RNA sequence, the CENPB box motif contributes to the formation of the apical loop structure through base pairing. Hence, we conclude that the apical loop in the stem-loop that we characterized is important for cenRNA function in chromosome segregation.

cenRNAs have been mainly investigated as co-factor for the activity of enzymes such as Aurora kinase B (Blower, 2016; Ferri et al., 2009; Jambhekar et al., 2014), or as interactor of centromeric proteins (Bobkov et al., 2018; Bouzinba-Segard et al., 2006; Chan and Wong, 2012; Grenfell et al., 2016; Ideue et al., 2014; McNulty et al., 2017; Quénet and Dalal, 2014; Rošić et al., 2014). Depleting centromeric transcripts is sufficient to interfere with mitotic progression. Because centromeric DNA sequences share little similarity (Bensasson et al., 2008; Henikoff et al., 2001; Suzuki et al., 2020), investigating centromeric RNAs may inform on the conserved function. We found both cenRNA strands expressed in mouse ESCs and consistently observed higher expression for the reverse transcript. Transcript overexpression led to slightly different phenotypes: the forward strand transcript resulted in a more severe mitotic phenotype but only the reverse transcript impacted cell

proliferation and reduced levels of chromatin-bound CENPC. Aurora B and the CPC complex interact with the forward, but not with the reverse MinSat transcripts (Ferri et al., 2009). While we cannot exclude additional effects, for instance through modulation of chromatin structure, we favor the interpretation that these observations reflect the contribution of several feedback loops underlying the homeostasis of cenRNAs. In such a “feedback” model the different affinities for the MinSat transcripts could modulate the amount of “free” CENPC. Binding to a lower affinity (e.g., reverse) transcript, could enable CENPC to bind instead to other protein partners (such as the CPC complex). Such feedback loops could be mediated, for example, by processing by DICER (Kanellopoulou et al., 2005), by the activity of transcription factors regulating their expression (Ishikura et al., 2021), and/or by secondary structures such as R-loops. Indeed, R-loop formation at centromeres can activate ATR in human cancer cells to ensure faithful chromosome segregation (Kabeche et al., 2018). It is thus possible that either transcript strand is prone to form R-loops thus blocking the transcription of the opposite strand. We envision such feedback loops can contribute to a correct balance in the proportion of forward versus reverse MinSat transcripts. Indeed, altering such balance leads to chromosome segregation errors albeit CENPC association to chromatin is only reduced when the reverse transcript is overexpressed. Our RNA pull-down experiments indicate that the forward transcript, though comprising the lower MinSat fraction, associates more strongly with CENPC than the reverse transcript. We thus speculate that under overexpression conditions transcripts “compete” for CENPC binding. This is in line with our observations that CENPC binding to chromatin is only affected when both endogenous forward and reverse transcripts are knocked down and suggests that a balance of the forward and reverse transcripts is important for CENPC association with mitotic chromatin. Thus, it is possible that the forward transcripts affect chromosome segregation through the mislocalization of Aurora B, for example. Yet stem cells are known to tolerate aneuploidy and to uncouple apoptosis in spite of spindle check point activation (Mantel et al., 2007) and thus a cell proliferation phenotype would be less severe. Reverse transcripts, on the other hand, could potentially compete with the endogenous forward transcripts for CENPC binding, which could in turn lead to a reduction of CENPC and CENPA in centromeric chromatin, thereby leading to a defect in cell division and to cell death (Kalitsis et al., 1998; Tomkiel et al., 1994), resulting in a greater proliferation phenotype. From a broader perspective, this would imply that the defective binding or misincorporation of CENPC and CENPA leads to a stronger cell proliferation phenotype in ESCs than aneuploidy resulting from potential spindle defects. Overall, our observations substantiate the notion that homeostasis of cenRNAs is critical for chromosome segregation.



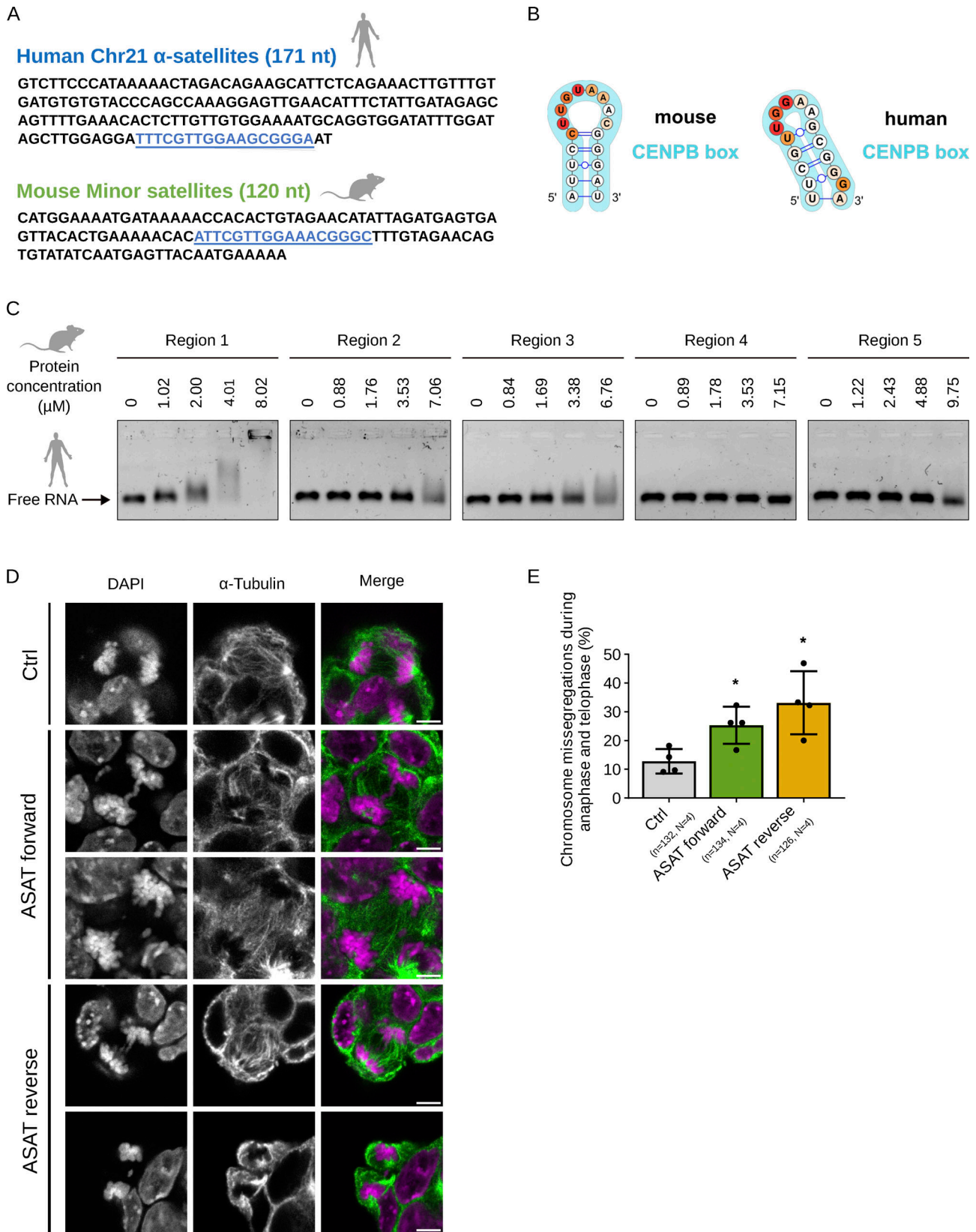


Figure 4. **The CENPB box of both mouse MinSat and human ASAT RNAs folds into highly similar stem-loop structures. (A)** Sequences of 1 repeat mouse MinSat RNA and human ASAT RNA corresponding to the human chromosome 21 (chr21)  $\alpha$ -satellite. The blue underlined region corresponds to the mouse MinSat RNA and human ASAT RNA. (B) Apical loop structures of MinSat RNA and human chr21 ASAT RNA based on SHAPE analysis. Note that in both species the CENPB box motif (blue) are embedded in the apical loop structures. (C) EMSA of the predicted RNA binding regions in CENPC and 1 repeat of forward human chr21 ASAT RNA. Protein concentrations are indicated on the top. (D) Representative images of mouse ESCs stained with  $\alpha$ -tubulin and DAPI under the indicated transfection

conditions. Ctrl, empty vector control. Scale bars, 5  $\mu$ m. **(E)** Quantification of chromosome missegregation events during mitosis in mouse ESCs transfected with a single repeat of human chr21 ASAT RNA. Ctrl, empty vector control. Bars indicate the mean percentages; error bars, standard deviation; \*, P value  $\leq$  0.05 (pair-wise comparison to control). The number of biological replicates (*N*) and the total number of mitotic figures analyzed (*n*) are indicated. Source data are available for this figure: SourceData F4.

We found that the MinSat forward transcript is structured. Remarkably, we identified a similar stem-loop structure in the human ASAT RNA. Taken together, we posit that the stem-loop structure of cenRNA found in the vicinity of the CENPB box is an important feature for the functional evolution of mammalian cenRNAs, potentially by mediating binding to centromeric proteins. In the future, it will be interesting to investigate whether other transcripts capable of recruiting CENPC, such as the Y-chromosome satellites or the long non-coding RNAs PCAT2 and CCTT, which do not contain CENPB motifs, function in a similar way (Arunkumar et al., 2022; Zhang et al., 2022). These considerations could explain how even highly divergent DNA sequences may generate transcripts able to interact with less rapidly evolving protein partners.

## Materials and methods

### Cell culture

Mouse E14 ESCs and Fucci (Nakatani et al., 2022) cells were grown in Dulbecco's modified Eagle's medium (DMEM) with GlutaMAX (31966-021; Invitrogen) containing 15% fetal calf serum (Lot P-10397, Cat P30-3302; PAN-Biotech), 2 $\times$  leukemia inhibitory factor (IGBMC), penicillin-streptomycin (15140122; Invitrogen), 0.1 mM 2-mercaptoethanol (31350010; Invitrogen), 3  $\mu$ M CHIR99021 (13122-25, Cayman Chemical), and 1  $\mu$ M PD0325901 (13034; Cayman Chemical) on gelatin (PO6-20410; PAN-Biotech)-coated plates. MEFs were cultured in DMEM with 10% fetal calf serum and penicillin-streptomycin.

### Plasmid transfection

2.5  $\mu$ g pCAG plasmid DNA containing centromeric sequences was first incubated with 5  $\mu$ l lipofectamine 2000 (11668019; Thermo Fisher Scientific) in 500  $\mu$ l Opti-MEM (31985062; Thermo Fisher Scientific) for 15 min. For each condition, a total of 250,000 ESCs were transfected in suspension in a single well of 6-well plates. Cells were harvested or fixed 48 h after transfection.

### Knockdown of MinSat RNAs

MinSat RNA knockdown was performed with 2'-O-MOE modified antisense oligo (Integrated DNA Technologies). The ASOs are listed in Table S1. Transfection of ASOs was performed using lipofectamine RNAiMAX (13778075; Thermo Fisher Scientific) according to the manufacturer's protocol. A total amount of 200  $\mu$ mol (100  $\mu$ mol each) ASOs were used at a final concentration of 100  $\mu$ M in the 2 ml final mixture. 250,000 ESCs were transfected in a well of a 6-well plate for each condition and assayed after transfection 24 or 48 h.

### Total RNA extraction and Real-Time qPCR

Cells were first lysed in TRIzol reagents (15596018; Invitrogen) following the manufacturer's instructions. Each 10  $\mu$ g of total

RNA was treated with 2 U of TURBO DNase for 1 h. A second round of TRIzol RNA extraction was performed to remove DNase and buffer. 1  $\mu$ g of total RNA was used for cDNA synthesis by applying GoScript Reverse Transcriptase (A5001; Promega). For strand-specific reverse transcription, sequence-specific primers (Table S2) were used to discriminate between strands of transcripts. The expression level was determined by using GoTaq qPCR Master Mix (A6002; Promega) on a LightCycler 96 Real-time PCR system (Roche). The qPCR primers are listed in Table S2.

### Immunofluorescence

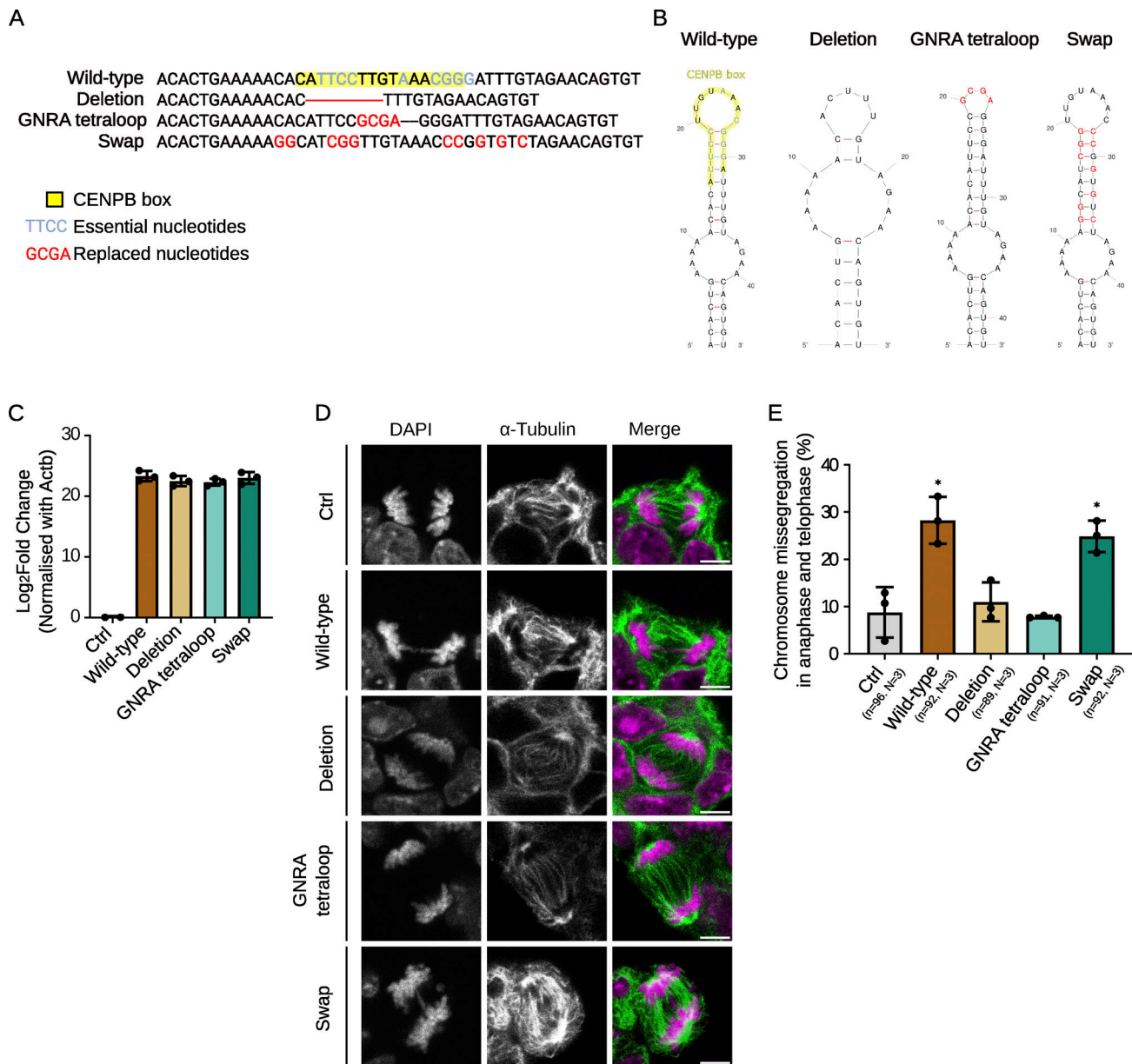
The transfected cells were cultured on gelatin-coated coverslips for 48 h. Cells were fixed with 4% PFA in BRB80 buffer (to preserve tubulin structure) or PBS (for staining of centromeric proteins), respectively, for 10 min at room temperature. After fixation, cells were washed twice with DPBS (14190-144; Gibco) and permeabilized with 0.3% Triton-X100 in DPBS for 10 min at room temperature. Blocking was performed with 3% BSA and 0.1% Triton-X100 in DPBS for 20 min. For antibody hybridization, primary antibodies were incubated overnight at 4°C, washed four times with a wash buffer (1% BSA, 0.1% Triton-X100 in DPBS) followed by incubation for 1 h with secondary antibodies at room temperature. All the antibodies used are listed in Table S3. Then, the samples were washed again four times with wash buffer and three times DPBS. Mowiol was used for sample mounting. Images were acquired on a Leica TCS SP8 confocal microscope with Leica Application Suite X (LASX 3.5.7) with a 63 $\times$  plan-apochromate NA 1.4 oil immersion objective using photomultiplying tubes (PMTs) and hybrid detectors (HyDs). Pixel (px) size was typically set to 70–80 nm/px and optical sections were acquired every 300 nm.

### Image analysis

To measure CENPC intensities in mouse ESCs, we manually selected mitotic cells in micrographs and hand-placed individual points over CREST-labeled foci, which we used as a proxy for kinetochores. Using a Fiji macro, the CENPC mean intensity in a 5-px radius at each selected point was recorded and normalized to the DAPI intensity in the same area to account for decaying fluorescence intensity along the Z axis. The datapoints collected per cell were further aggregated to determine the mean intensity per cell using R, and the mean per experiment was calculated from all mitotic cells.

### Chromatin extraction and Western blot analysis

Cells were first synchronized with 100 ng/ml KaryoMAX Colcemid for 3 h. Next, cells were trypsinized and washed twice with cold DPBS. A million cells were lysed in 100  $\mu$ l Triton Extraction Buffer (PBS containing 0.5% Triton-X100 and protease inhibitor mix (80-6501-23; cytiva) on ice for 10 min. Nuclei were



**Figure 5. Structural conservation rather than sequence motifs determines effects of MinSat RNAs on chromosome segregation.** (A) Sequence comparison of the various MinSat RNA constructs used: wild type: single repeat of consensus forward MinSat RNA sequence; Deletion: deletion of central loop sequence; GNRA tetraloop: replacement of the loop with a smaller stable tetraloop sequence; Swap: swapped CG base pairing in the stem region. The yellow highlighted region indicates the CENPB box motif, blue-colored nucleotides are an integral part of the CENPB box motif and red-colored nucleotides are positions that were mutated within the CENPB box motif in our constructs. (B) Predicted structures of wild-type, deletion, GNRA tetraloop, and swap RNAs using mFold (Zuker, 2003). The yellow highlighted region depicts the CENPB box motif in wild-type sequence and red-colored nucleotides correspond to the mutated positions as depicted in A. (C) Quantification of MinSat RNA levels in ESCs after transfection of the pCAG plasmids containing different mutants of MinSat using RT-qPCR analysis. Bars, mean values of  $N = 3$  independent experiments (dots) normalized to  $\beta$ -actin mRNA and empty pCAG vector control (Ctrl); error bars, standard deviation. (D) Representative images of mouse ESCs transfected with plasmids expressing empty plasmid, wild-type, or different mutants. Cells were immunostained with  $\alpha$ -tubulin and DAPI. Scale bars, 5  $\mu$ m. (E) Quantification of chromosome missegregation events in anaphase and telophase mouse ESCs after overexpression of the structural RNA mutants, wild-type, or empty pCAG vector as control (Ctrl), respectively. Bars indicate the mean percentage of  $N = 3$  independent experiments (dots); error bars, standard deviation; \*,  $P$  value  $\leq 0.05$  (pair-wise comparison to control).

collected by centrifuging at 6,500  $g$  for 10 min at 4°C and washed with 50  $\mu$ l Triton Extraction Buffer. After washing, nuclei were then treated with 20  $\mu$ l 0.2 N HCl overnight. The supernatant was transferred and neutralized with 2 M NaOH at 1/10 of vol of the supernatant the next day. For SDS-PAGE, 1  $\mu$ g of chromatin extract was used for each sample. Protein gel was then

transferred to a PVDF membrane. The following antibodies were applied for immunoblotting: CENPA (2048; Cell Signaling), INCENP (ab12183; Abcam), CENPC (ab193666; Abcam), histone H3 (ab1791; Abcam), and GAPDH (MAB374; Millipore).

After antibody hybridization, membranes were washed four times with TBST buffer and imaged with ChemiDoc Touch



Imaging System (BioRad). Immunoblots were quantified in Fiji (Schindelin et al., 2012) by measuring band intensities within predefined rectangular regions and further processed after subtracting the local background intensity of the immunoblot membranes.

### Cell proliferation assay

Cell proliferation is performed starting with 250,000 cells along with transfection. Cells were trypsinized and hand-counted using a cell counting chamber (8100104; Hirschmann). The cell numbers were documented every 24 h for 3 days.

### Flow cytometry analysis of cell cycle distribution

Cells were first trypsinized and washed with DPBS twice. Cold absolute ethanol was used for cell fixation on ice for 30 min. After fixation, cells were treated with 250 µg/ml (Invitrogen) RNase for 5 min at room temperature. Propidium iodide (PI) was then added to the cell mixture at a final concentration of 50 µg/ml and stained for 30 min on ice. For flow cytometry data analysis, Flowjo 10 software was used. Cell cycle analysis is calculated by fitting the Dean-Jett-Fox model.

### 3' end biotinylated RNA pulldown assay

We in vitro-transcribed RNA using mMMESSAGE mMACHINE T7 Transcription Kit (Invitrogen). Then, equimolar (20 µmol) RNAs were labeled with biotin following the manufacturer's protocol of Pierce RNA 3' End Biotinylation Kit. After RNA clean-up with RNAClean XP (Beckman Coulter) beads, we heated the RNA at 70°C for 10 min and cooled it down gradually until reaching room temperature. 107 cells were harvested and lysed on ice in Pierce IP Lysis Buffer (Thermo Fisher Scientific) containing RNasin Ribonuclease Inhibitor (Promega) and PMSF. Dynabeads MyOne Streptavidin C1 (Invitrogen) were incubated with biotin-labeled RNA. We next centrifuged the cell lysate at 20,000 rcf for 20 min, and the supernatant was precleared with Dynabeads for 1 h at 4°C. The RNA-conjugated beads were then added into the cell lysate and incubated at 4°C on a rotor for 3 h. The RNA-protein complexes were then washed twice with NT2, NT2-middle, NT2-high, and NT2-KSCN buffer. Finally, all samples were boiled in SDS-PAGE sample buffer (Biorad) and processed by immunoblotting as described above.

### RNA transcription and purification

RNA-encoding DNA templates (possessing the T7 promoter sequence) were either prepared using polymerase chain reaction (PCR; primers ordered from Eurofins Genomics) or ordered as single-stranded DNA templates from Eurofins Genomics. PCR reactions were carried out with the Phusion polymerase (New England Biolabs [NEB]) according to manufacturer instructions. RNAs were in vitro-transcribed using in house-prepared T7 polymerase. Briefly, 50 µl transcriptions containing 0.64 µM DNA (supplemented with 0.64 µM T7 top primer for single-stranded DNA templates), 20–80 µM MgCl<sub>2</sub>, 8 mM of each rNTP, 5% PEG 8000, 1× transcription buffer (5 mM Tris, pH 8, 5 mM spermidine, 10 mM DTT), and 0.6 mg of T7 polymerase were incubated at 37°C for 1 h. Following transcription, the RNAs were purified on urea-denaturing polyacrylamide gels, followed

by extraction of the RNA from the gel using the crush and soak method (Mörl and Schmelzer, 1993). Extracted RNAs were equilibrated against water and stored at –20°C until further use.

### Selective 2' hydroxyl acylation analyzed by primer extension (SHAPE)

RNAs possessing 5' (5'-GGAACAACAAGGCCGGAGUACGGCCAA A-3') and 3' (5'-AAAAGCAGCGAGUAGCUGCAACAAAAGAAA CAACAACAACAAC-3') SHAPE cassettes and a universal primer binding site were diluted to 1 µM and SHAPE was carried out as previously described (Wilkinson et al., 2006). Briefly, RNA was snap-cooled (95°C for 3 min, followed by 5 min on ice) in a buffer containing 200 mM NaCl, 100 mM HEPES, pH 8, 0.2 mM EDTA, followed by folding at 37°C (30 min) in a buffer containing 100 mM NaCl, 50 mM HEPES, pH 8, and 16.5 mM MgCl<sub>2</sub>. RNA was then treated with 1.65, 3.2, or 6.4 mM 1M7 for 5 min at 37°C followed by purification using Poly(A) purist magnetic beads (Invitrogen) according to manufacturer instructions. Purified RNAs were reverse transcribed using the Superscript III reverse transcriptase (Invitrogen) and a 5' Fam labeled primer (Eurofins Genomics) according to manufacturer instructions, followed by purification using magnetic beads. Fam-labeled cDNA fragments were dissolved in HiDi formamide and sequenced using a SeqStudio Fragment Analyzer. Data were analyzed using HiTRACE (Yoon et al., 2011).

### NMR

RNAs were prepared and equilibrated in a buffer containing 15 mM NaCl, 25 mM sodium phosphate, pH 6.4. The RNA was snap-cooled (95°C for 3 min, 4°C for 5 min) and 1D and 1H NOESY NMR spectra were recorded at 278 K on a 1.2 GHz NMR spectrometer equipped with a cryogenic TCI probe.

### CatRAPID prediction of RNA binding ability

RNA binding regions of the CenPc protein were identified using CatRAPID (Livi et al., 2016) signature.

### Protein expression and purification

DNA encoding for the CENPC protein (amino acids 1–906) was ordered as a codon-optimized GBlock from Integrated DNA Technologies (IDT). Primers (containing NcoI and KpnI enzyme restriction sites) corresponding to the terminal ends of each CatRAPID fragment were ordered from Eurofins Genomics. PCRs were performed using the Phusion polymerase (NEB) according to manufacturer instructions. CatRAPID PCR fragments were double digested with NcoI (NEB) and KpnI (NEB) according to manufacturer instructions, and ligated into a NcoI, KpnI double-digested pETM-11 cloning vector using the NEB Quick Ligase kit. Mini-prepped DNA (obtained from transformations of ligated plasmids into DH10b *E. coli* competent cells) was sequenced and then transformed into BL21 *E. coli* competent cells. Colonies were grown up in 1 liter of lysogeny broth at 37°C and induced with 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) at 18°C. Cells were lysed using a French press in a buffer containing 500 mM NaCl, 10 mM imidazole, 50 mM Tris, pH 8, and 5 mM β-mercaptoethanol. Protein (possessing an N-terminal His<sub>6</sub>-tag followed by a tobacco enterovirus (TEV) protease

cleavage site) was purified by IMAC against an increasing imidazole concentration gradient. The N-terminal His<sub>6</sub>-tag was cleaved using in house-prepared TEV protease, followed by size exclusion chromatography on a Superdex S75 purification column and eluted in buffer containing 300 mM NaCl, 50 mM Tris pH 7.5, and 5 mM DTT. Protein quality was assessed by SDS PAGE.

### Binding shift assays

RNAs were diluted to 0.250 μmol and incubated with increasing concentrations of protein as indicated in the figures. Samples were loaded onto a 1% TBE agarose gel supplemented with DNA Stain G (Serva) and run for 40 min at 70 V at room temperature. Gels were imaged on a GelDoku imager.

### Statistical analyses

We used Prism (GraphPad software) and R for statistical analysis. Statistical tests were performed as indicated in the figure legends. For most experiments, we used a paired *t* test for statistical evaluation, unless otherwise stated in the figure legends. For multiple treatments in immunoblotting experiments, we used the “multcomp” package in R on an ANOVA model, evaluating control to treatment conditions. For *t* test, we assumed normal distribution, but this was not formally tested.

### Online supplemental material

**Fig. S1** shows the RT-qPCR analysis of MinSat expression and the controls for the FUCCI sorting and the data upon expression of MinSat in mouse embryonic fibroblasts. **Fig. S2** shows the comparison of human and mouse CENPC and the predicted alternative structures of MinSat repeats by SHAPE. **Fig. S3** shows the SHAPE-derived structure of the human chromosome 21 ASAT forward transcript and the RT-qPCR controls for the expression of human ASAT in ESCs. Table S1 shows includes the sequence of the antisense oligos used for the MinSat knockdown experiments. Table S2 shows the sequence of the primers used for RT-qPCR analysis of MinSat expression and β-actin control. Table S3 shows the list of all the primary and secondary antibodies used in the manuscript.

### Data availability

All data generated are available in the published article and its online supplemental material. The raw data are available from the corresponding author upon reasonable request.

### Acknowledgments

This work has been funded primarily by a grant from the German Research Foundation (DFG) to A. Ettinger (ET-124/2-1) with support from funding to M.-E. Torres-Padilla through the Helmholtz Association and the DFG SFB 1064 (project-ID 213249687 to M.-E. Torres-Padilla) and TRR 267 (project-ID 403584255 to M. Sattler). We acknowledge access to NMR measurement time at the Bavarian NMR Center and Arie Geerlof, Helmholtz Munich protein expression and purification facility, for T7 polymerase and TEV protease.

Author contributions: Y.-L. Chen, A.N. Jones, and A. Ettinger performed and analyzed the experiments; A. Crawford

performed experiments for the revised manuscript; M. Sattler, A. Ettinger, and M.-E. Torres-Padilla conceived, designed, and supervised the study. All authors contributed to manuscript preparation and read, commented on, and approved the manuscript.

Disclosures: All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. M.-E. Torres-Padilla reported consultancy for the Merck Ethics Advisory Board outside the submitted work. No other disclosures were reported.

Submitted: 5 September 2023

Revised: 6 March 2024

Accepted: 29 March 2024

### References

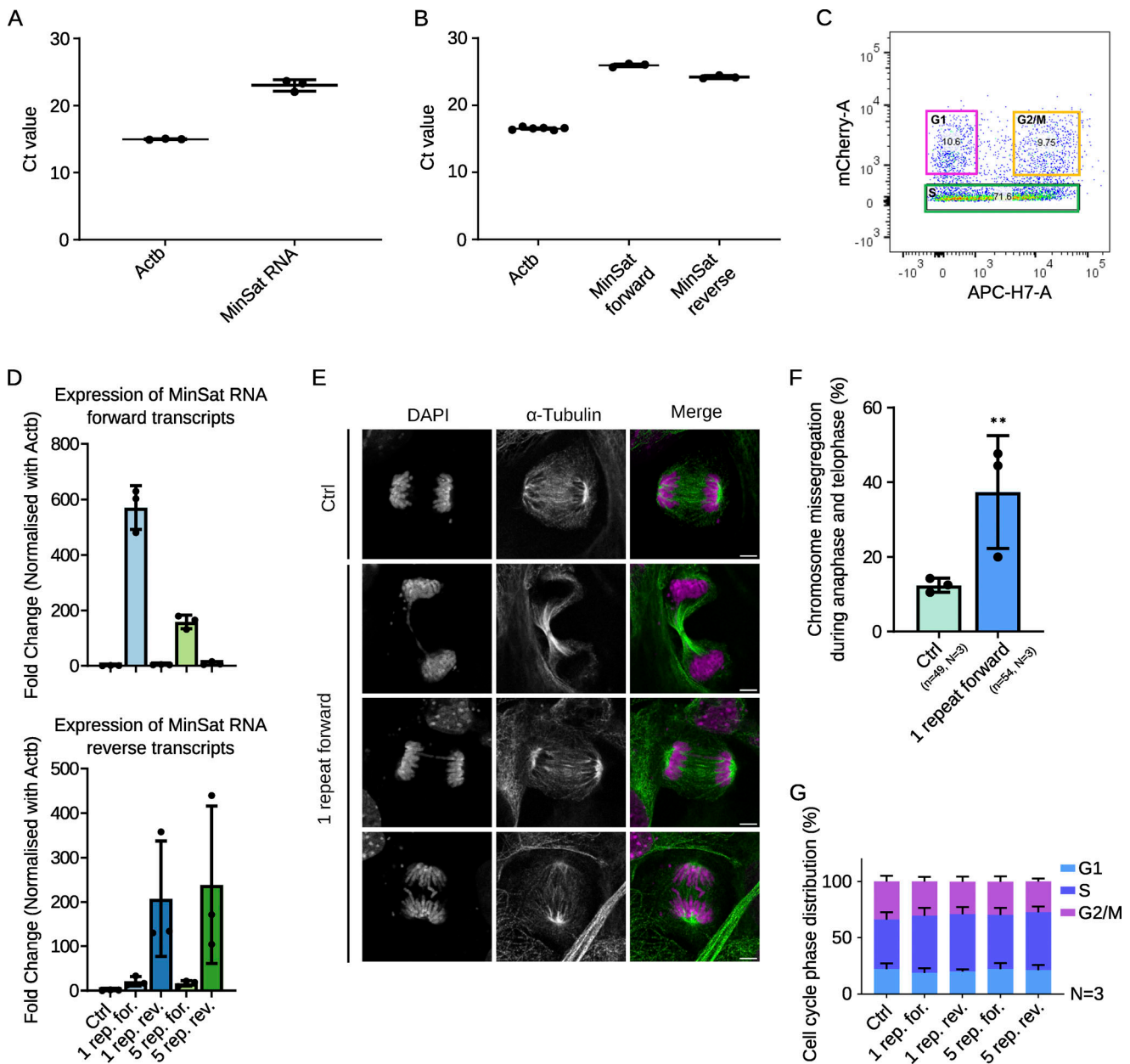
- Alkan, C., M.F. Cardone, C.R. Catacchio, F. Antonacci, S.J. O'Brien, O.A. Ryder, S. Purgato, M. Zoli, G. Della Valle, E.E. Eichler, and M. Ventura. 2011. Genome-wide characterization of centromeric satellites from multiple mammalian genomes. *Genome Res.* 21:137–145. <https://doi.org/10.1101/gr.111278.110>
- Arunkumar, G., S. Baek, D. Sturgill, M. Bui, and Y. Dalal. 2022. Oncogenic lncRNAs alter epigenetic memory at a fragile chromosomal site in human cancer cells. *Sci. Adv.* 8:eabl5621. <https://doi.org/10.1126/sciadv.abl5621>
- Barnhart, M.C., P.H. Kuich, M.E. Stellfox, J.A. Ward, E.A. Bassett, B.E. Black, and D.R. Foltz. 2011. HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J. Cell Biol.* 194: 229–243. <https://doi.org/10.1083/jcb.201012017>
- Barra, V., and D. Fachinetti. 2018. The dark side of centromeres: Types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nat. Commun.* 9:4340. <https://doi.org/10.1038/s41467-018-06545-y>
- Bensasson, D., M. Zarowiecki, A. Burt, and V. Koufopanou. 2008. Rapid evolution of yeast centromeres in the absence of drive. *Genetics.* 178: 2161–2167. <https://doi.org/10.1534/genetics.107.083980>
- Bergmann, J.H., J.N. Jakubsche, N.M. Martins, A. Kagansky, M. Nakano, H. Kimura, D.A. Kelly, B.M. Turner, H. Masumoto, V. Larionov, and W.C. Earnshaw. 2012. Epigenetic engineering: Histone H3K9 acetylation is compatible with kinetochore structure and function. *J. Cell Sci.* 125: 411–421. <https://doi.org/10.1242/jcs.090639>
- Bergmann, J.H., M.G. Rodriguez, N.M. Martins, H. Kimura, D.A. Kelly, H. Masumoto, V. Larionov, L.E. Jansen, and W.C. Earnshaw. 2011. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO J.* 30: 328–340. <https://doi.org/10.1038/emboj.2010.329>
- Black, B.E., and D.W. Cleveland. 2011. Epigenetic centromere propagation and the nature of CENP-a nucleosomes. *Cell.* 144:471–479. <https://doi.org/10.1016/j.cell.2011.02.002>
- Blower, M.D. 2016. Centromeric transcription regulates Aurora-B localization and activation. *Cell Rep.* 15:1624–1633. <https://doi.org/10.1016/j.celrep.2016.04.054>
- Bobkov, G.O.M., N. Gilbert, and P. Heun. 2018. Centromere transcription allows CENP-A to transit from chromatin association to stable incorporation. *J. Cell Biol.* 217:1957–1972. <https://doi.org/10.1083/jcb.201611087>
- Bouzinba-Segard, H., A. Guais, and C. Francastel. 2006. Accumulation of small murine minor satellite transcripts leads to impaired centromeric architecture and function. *Proc. Natl. Acad. Sci. USA.* 103:8709–8714. <https://doi.org/10.1073/pnas.0508006103>
- Bury, L., B. Moodie, J. Ly, L.S. McKay, K.H. Miga, and I.M. Cheeseman. 2020. Alpha-satellite RNA transcripts are repressed by centromere-nucleolus associations. *Elife.* 9:e59770. <https://doi.org/10.7554/eLife.59770>
- Carroll, C.W., K.J. Milks, and A.F. Straight. 2010. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *J. Cell Biol.* 189: 1143–1155. <https://doi.org/10.1083/jcb.201001013>
- Chan, F.L., O.J. Marshall, R. Saffery, B.W. Kim, E. Earle, K.H. Choo, and L.H. Wong. 2012. Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proc. Natl. Acad. Sci. USA.* 109: 1979–1984. <https://doi.org/10.1073/pnas.1108705109>

- Chan, F.L., and L.H. Wong. 2012. Transcription in the maintenance of centromere chromatin identity. *Nucleic Acids Res.* 40:11178–11188. <https://doi.org/10.1093/nar/gks921>
- Chik, J.K., V. Moiseeva, P.K. Goel, B.A. Meinen, P. Koldewey, S. An, B.G. Mellone, L. Subramanian, and U.S. Cho. 2019. Structures of CENP-C cupin domains at regional centromeres reveal unique patterns of dimerization and recruitment functions for the inner pocket. *J. Biol. Chem.* 294:14119–14134. <https://doi.org/10.1074/jbc.RA119.008464>
- Cleveland, D.W., Y. Mao, and K.F. Sullivan. 2003. Centromeres and kinetochores: From epigenetics to mitotic checkpoint signaling. *Cell.* 112:407–421. [https://doi.org/10.1016/S0092-8674\(03\)00115-6](https://doi.org/10.1016/S0092-8674(03)00115-6)
- Coronado, D., M. Godet, P.Y. Bourillot, Y. Taponnier, A. Bernat, M. Petit, M. Afanassieff, S. Markossian, A. Malashicheva, R. Iacone, et al. 2013. A short G1 phase is an intrinsic determinant of naïve embryonic stem cell pluripotency. *Stem Cell Res.* 10:118–131. <https://doi.org/10.1016/j.scr.2012.10.004>
- Diederichs, S. 2014. The four dimensions of noncoding RNA conservation. *Trends Genet.* 30:121–123. <https://doi.org/10.1016/j.tig.2014.01.004>
- Djupedal, I., M. Portoso, H. Spåhr, C. Bonilla, C.M. Gustafsson, R.C. Allshire, and K. Ekwall. 2005. RNA Pol II subunit Rpb7 promotes centromeric transcription and RNAi-directed chromatin silencing. *Genes Dev.* 19:2301–2306. <https://doi.org/10.1101/gad.344205>
- Du, Y., C.N. Topp, and R.K. Dawe. 2010. DNA binding of centromere protein C (CENPC) is stabilized by single-stranded RNA. *PLoS Genet.* 6:e1000835. <https://doi.org/10.1371/journal.pgen.1000835>
- Fachinetti, D., H.D. Folco, Y. Nechemia-Arbely, L.P. Valente, K. Nguyen, A.J. Wong, Q. Zhu, A.J. Holland, A. Desai, L.E. Jansen, and D.W. Cleveland. 2013. A two-step mechanism for epigenetic specification of centromere identity and function. *Nat. Cell Biol.* 15:1056–1066. <https://doi.org/10.1038/ncb2805>
- Ferri, F., H. Bouzinba-Segard, G. Velasco, F. Hubé, and C. Francastel. 2009. Non-coding murine centromeric transcripts associate with and potentiate Aurora B kinase. *Nucleic Acids Res.* 37:5071–5080. <https://doi.org/10.1093/nar/gkp529>
- Gopalakrishnan, S., B.A. Sullivan, S. Trazzi, G. Della Valle, and K.D. Robertson. 2009. DNMT3B interacts with constitutive centromere protein CENP-C to modulate DNA methylation and the histone code at centromeric regions. *Hum. Mol. Genet.* 18:3178–3193. <https://doi.org/10.1093/hmg/ddp256>
- Grenfell, A.W., R. Heald, and M. Strzelecka. 2016. Mitotic noncoding RNA processing promotes kinetochore and spindle assembly in *Xenopus*. *J. Cell Biol.* 214:133–141. <https://doi.org/10.1083/jcb.201604029>
- Henikoff, S., K. Ahmad, and H.S. Malik. 2001. The centromere paradox: Stable inheritance with rapidly evolving DNA. *Science.* 293:1098–1102. <https://doi.org/10.1126/science.1062939>
- Hörz, W., and W. Altenburger. 1981. Sequence specific cleavage of DNA by micrococcal nuclease. *Nucleic Acids Res.* 9:2643–2658. <https://doi.org/10.1093/nar/9.12.2643>
- Ideue, T., Y. Cho, K. Nishimura, and T. Tani. 2014. Involvement of satellite I noncoding RNA in regulation of chromosome segregation. *Genes Cells.* 19:528–538. <https://doi.org/10.1111/gtc.12149>
- Ideue, T., and T. Tani. 2020. Centromeric non-coding RNAs: Conservation and diversity in function. *Noncoding RNA.* 6:4. <https://doi.org/10.3390/ncrna6010004>
- Ishikura, S., K. Yoshida, S. Hashimoto, K. Nakabayashi, T. Tsunoda, and S. Shirasawa. 2021. CENP-B promotes the centromeric localization of ZFAT to control transcription of noncoding RNA. *J. Biol. Chem.* 297:101213. <https://doi.org/10.1016/j.jbc.2021.101213>
- Iwahara, J., T. Kigawa, K. Kitagawa, H. Masumoto, T. Okazaki, and S. Yokoyama. 1998. A helix-turn-helix structure unit in human centromere protein B (CENP-B). *EMBO J.* 17:827–837. <https://doi.org/10.1093/emboj/17.3.827>
- Jambhekar, A., A.B. Emerman, C.T. Schweidenback, and M.D. Blower. 2014. RNA stimulates Aurora B kinase activity during mitosis. *PLoS One.* 9:e100748. <https://doi.org/10.1371/journal.pone.0100748>
- Johnsson, P., L. Lipovich, D. Grandér, and K.V. Morris. 2014. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim. Biophys. Acta.* 1840:1063–1071. <https://doi.org/10.1016/j.bbagen.2013.10.035>
- Kabeche, L., H.D. Nguyen, R. Buisson, and L. Zou. 2018. A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science.* 359:108–114. <https://doi.org/10.1126/science.aan6490>
- Kalitsis, P., K.J. Fowler, E. Earle, J. Hill, and K.H. Choo. 1998. Targeted disruption of mouse centromere protein C gene leads to mitotic disarray and early embryo death. *Proc. Natl. Acad. Sci. USA.* 95:1136–1141. <https://doi.org/10.1073/pnas.95.3.1136>
- Kanellopoulou, C., S.A. Muljo, A.L. Kung, S. Ganesan, R. Drapkin, T. Jenkinson, D.M. Livingston, and K. Rajewsky. 2005. Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* 19:489–501. <https://doi.org/10.1101/gad.1248505>
- Kato, H., J. Jiang, B.R. Zhou, M. Rozendaal, H. Feng, R. Ghirlando, T.S. Xiao, A.F. Straight, and Y. Bai. 2013. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science.* 340:1110–1113. <https://doi.org/10.1126/science.1235532>
- Klare, K., J.R. Weir, F. Basilico, T. Zimniak, L. Massimiliano, N. Ludwigs, F. Herzog, and A. Musacchio. 2015. CENP-C is a blueprint for constitutive centromere-associated network assembly within human kinetochores. *J. Cell Biol.* 210:11–22. <https://doi.org/10.1083/jcb.201412028>
- Kwon, M.S., T. Hori, M. Okada, and T. Fukagawa. 2007. CENP-C is involved in chromosome segregation, mitotic checkpoint function, and kinetochore assembly. *Mol. Biol. Cell.* 18:2155–2168. <https://doi.org/10.1091/mbc.e07-01-0045>
- Lim, S., R.R. Yocum, P.A. Silver, and J.C. Way. 2023. High spontaneous integration rates of end-modified linear DNAs upon mammalian cell transfection. *Sci. Rep.* 13:6835. <https://doi.org/10.1038/s41598-023-33862-0>
- Livi, C.M., P. Klus, R. Delli Ponti, and G.G. Tartaglia. 2016. catRAPID signature: identification of ribonucleoproteins and RNA-binding regions. *Bioinformatics.* 32:773–775. <https://doi.org/10.1093/bioinformatics/btv629>
- Mallm, J.-P., and K. Rippe. 2015. Aurora kinase B regulates telomerase activity via a centromeric RNA in stem cells. *Cell Rep.* 11:1667–1678. <https://doi.org/10.1016/j.celrep.2015.05.015>
- Mantel, C., Y. Guo, M.R. Lee, M.K. Kim, M.K. Han, H. Shibayama, S. Fukuda, M.C. Yoder, L.M. Pelus, K.S. Kim, and H.E. Broxmeyer. 2007. Checkpoint-apoptosis uncoupling in human and mouse embryonic stem cells: A source of karyotypic instability. *Blood.* 109:4518–4527. <https://doi.org/10.1182/blood-2006-10-054247>
- Marston, A.L. 2014. Chromosome segregation in budding yeast: Sister chromatid cohesion and related mechanisms. *Genetics.* 196:31–63. <https://doi.org/10.1534/genetics.112.145144>
- Masumoto, H., H. Masukata, Y. Muro, N. Nozaki, and T. Okazaki. 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. <https://doi.org/10.1083/jcb.109.5.1963>
- McNulty, S.M., L.L. Sullivan, and B.A. Sullivan. 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.e6. <https://doi.org/10.1016/j.devcel.2017.07.001>
- Merino, E.J., K.A. Wilkinson, J.L. Coughlan, and K.M. Weeks. 2005. RNA structure analysis at single nucleotide resolution by selective 2'-hydroxyl acylation and primer extension (SHAPE). *J. Am. Chem. Soc.* 127:4223–4231. <https://doi.org/10.1021/ja043822v>
- Milks, K.J., B. Moree, and A.F. Straight. 2009. Dissection of CENP-C-directed centromere and kinetochore assembly. *Mol. Biol. Cell.* 20:4246–4255. <https://doi.org/10.1091/mbc.e09-05-0378>
- Moree, B., C.B. Meyer, C.J. Fuller, and A.F. Straight. 2011. CENP-C recruits M18BP1 to centromeres to promote CENP-A chromatin assembly. *J. Cell Biol.* 194:855–871. <https://doi.org/10.1083/jcb.201106079>
- Mörl, M., and C. Schmelzer. 1993. A simple method for isolation of intact RNA from dried polyacrylamide gels. *Nucleic Acids Res.* 21:2016. <https://doi.org/10.1093/nar/21.8.2016>
- Nakatani, T., J. Lin, F. Ji, A. Ettinger, J. Pontabry, M. Tokoro, L. Altamirano-Pacheco, J. Fiorentino, E. Mahammadov, Y. Hatano, et al. 2022. DNA replication fork speed underlies cell fate changes and promotes reprogramming. *Nat. Genet.* 54:318–327. <https://doi.org/10.1038/s41588-022-01023-0>
- Przewlaka, M.R., Z. Venkei, V.M. Bolanos-Garcia, J. Debski, M. Dadlez, and D.M. Glover. 2011. CENP-C is a structural platform for kinetochore assembly. *Curr. Biol.* 21:399–405. <https://doi.org/10.1016/j.cub.2011.02.005>
- Quénet, D., and Y. Dalal. 2014. A long non-coding RNA is required for targeting centromeric protein A to the human centromere. *Elife.* 3:e03254. <https://doi.org/10.7554/eLife.03254>
- Rošić, S., F. Köhler, and S. Erhardt. 2014. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J. Cell Biol.* 207:335–349. <https://doi.org/10.1083/jcb.201404097>
- Sakaue-Sawano, A., H. Kurokawa, T. Morimura, A. Hanyu, H. Hama, H. Osawa, S. Kashiwagi, K. Fukami, T. Miyata, H. Miyoshi, et al. 2008. Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell.* 132:487–498. <https://doi.org/10.1016/j.cell.2007.12.033>



- Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, et al. 2012. Fiji: An open-source platform for biological-image analysis. *Nat. Methods*. 9: 676–682. <https://doi.org/10.1038/nmeth.2019>
- Sullivan, B.A., and G.H. Karpen. 2004. Centromeric chromatin exhibits a histone modification pattern that is distinct from both euchromatin and heterochromatin. *Nat. Struct. Mol. Biol.* 11:1076–1083. <https://doi.org/10.1038/nsmb845>
- Suntronpong, A., K. Kugou, H. Masumoto, K. Srikulnath, K. Ohshima, H. Hirai, and A. Koga. 2016. CENP-B box, a nucleotide motif involved in centromere formation, occurs in a New World monkey. *Biol. Lett.* 12: 20150817. <https://doi.org/10.1098/rsbl.2015.0817>
- Suzuki, Y., E.W. Myers, and S. Morishita. 2020. Rapid and ongoing evolution of repetitive sequence structures in human centromeres. *Sci. Adv.* 6: eabd9230. <https://doi.org/10.1126/sciadv.abd9230>
- Tanaka, K., and T. Hirota. 2016. Chromosomal instability: A common feature and a therapeutic target of cancer. *Biochim. Biophys. Acta.* 1866:64–75. <https://doi.org/10.1016/j.bbcan.2016.06.002>
- Tomkiel, J., C.A. Cooke, H. Saitoh, R.L. Bernat, and W.C. Earnshaw. 1994. CENP-C is required for maintaining proper kinetochore size and for a timely transition to anaphase. *J. Cell Biol.* 125:531–545. <https://doi.org/10.1083/jcb.125.3.531>
- Westhorpe, F.G., and A.F. Straight. 2014. The centromere: Epigenetic control of chromosome segregation during mitosis. *Cold Spring Harb. Perspect. Biol.* 7:a015818. <https://doi.org/10.1101/cshperspect.a015818>
- Wilkinson, K.A., E.J. Merino, and K.M. Weeks. 2006. Selective 2'-hydroxyl acylation analyzed by primer extension (SHAPE): Quantitative RNA structure analysis at single nucleotide resolution. *Nat. Protoc.* 1: 1610–1616. <https://doi.org/10.1038/nprot.2006.249>
- Wong, L.H., K.H. Brettingham-Moore, L. Chan, J.M. Quach, M.A. Anderson, E.L. Northrop, R. Hannan, R. Saffery, M.L. Shaw, E. Williams, and K.H. Choo. 2007. Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res.* 17: 1146–1160. <https://doi.org/10.1101/gr.6022807>
- Yoon, S., J. Kim, J. Hum, H. Kim, S. Park, W. Kladwang, and R. Das. 2011. HiTRACE: High-throughput robust analysis for capillary electrophoresis. *Bioinformatics.* 27:1798–1805. <https://doi.org/10.1093/bioinformatics/btr277>
- Zhang, C., D. Wang, Y. Hao, S. Wu, J. Luo, Y. Xue, D. Wang, G. Li, L. Liu, C. Shao, et al. 2022. LncRNA CCTT-mediated RNA-DNA and RNA-protein interactions facilitate the recruitment of CENP-C to centromeric DNA during kinetochore assembly. *Mol. Cell.* 82:4018–4032.e9. <https://doi.org/10.1016/j.molcel.2022.09.022>
- Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* 31:3406–3415. <https://doi.org/10.1093/nar/gkg595>

## Supplemental material



**Figure S1. RT-qPCR analysis of MinSat transcripts, flow cytometry of Fucci reporter cells and overexpression of MinSat in mouse embryonic fibroblasts.** (A) Raw Ct values of PCR amplification for  $\beta$ -actin and MinSat after reverse transcription with oligo dT and random hexamers. Solid line, mean of  $N = 3$  independent experiments; error bars, standard deviation. (B) Raw Ct values of PCR amplification after reverse transcription with primers specific for  $\beta$ -actin, MinSat forward, and MinSat reverse RNA, respectively. Solid line, mean of  $N = 3$  (MinSat) or  $N = 6$  ( $\beta$ -actin) experiments; error bars, standard deviation. (C) Representative gating strategy for fluorescence-activated cell sorting (FACS) of ESCs harboring the Fucci reporter. Red, G1 phase; green, S phase; yellow, G2/M phase. (D) Quantification of strand-specific MinSat RNA levels in ESCs after transfection of the pCAG plasmids containing forward or reverse MinSat repeats using RT-qPCR analysis. Bars, mean values of  $N = 3$  independent experiments (dots) normalized to  $\beta$ -actin mRNA and empty pCAG vector control (Ctrl); error bars, standard deviation. Top panel, qPCR for forward strand transcripts; lower panel, qPCR for reverse strand transcripts. (E) Representative images of missegregation events in MEFs after transfection with a single repeat of MinSat forward or empty pCAG vector control (Ctrl), respectively. Cells were immunostained with  $\alpha$ -tubulin (green) and DAPI (magenta). Scale bars, 5  $\mu$ m. (F) Percentage of chromosome missegregation events in total anaphase to telophase cells. Bars, mean; error bars, standard deviation; pair-wise comparisons with control (empty pCAG vector) P values \*\*  $\leq 0.01$ . The number of biological replicates ( $N$ ) and the total number of mitotic figures analyzed ( $n$ ) are indicated. (G) Flow cytometry analysis of cell cycle distribution after MinSat RNA overexpression using PI. Ctrl, empty pCAG vector. Bars indicate mean percentages of the respective cell cycle phase from three independent experiments; error bars, standard deviation.

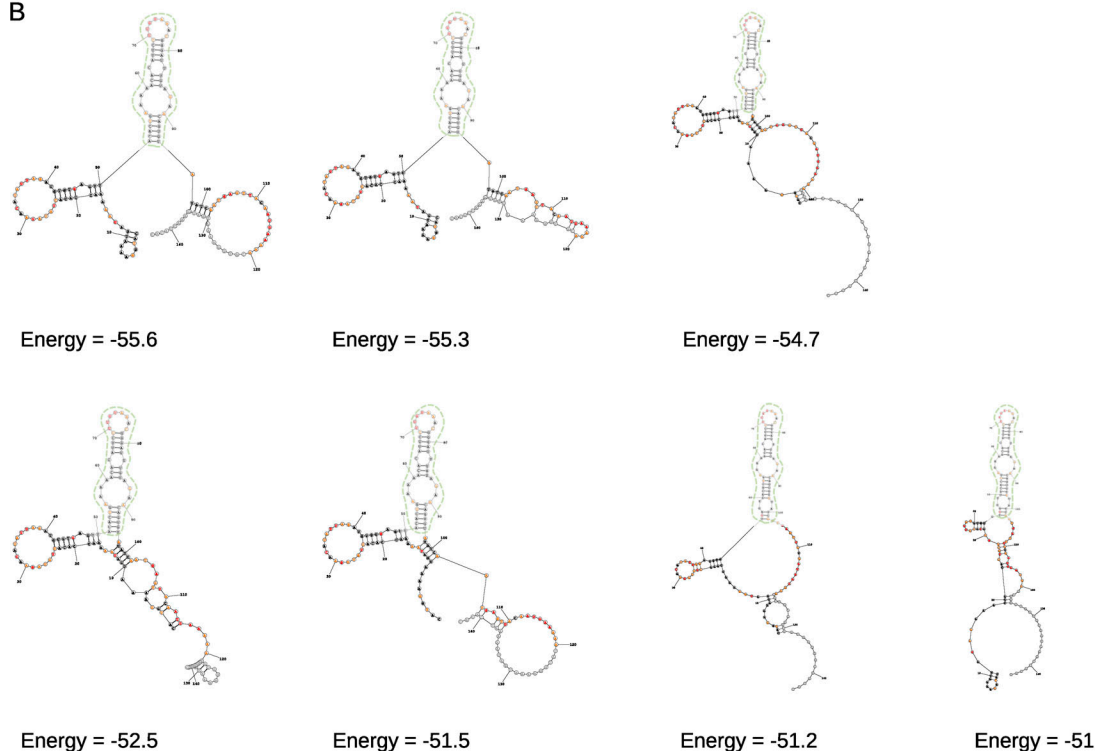


A

HUMAN	1	MAASGLDHLKNGYRRRCRPSRARDINTEGGQNVLEILQDCFEEKSLANDFSTNSTKSVNSTRKIKDTCIQSPSKECQKSHPKSVVPVSSKKEASLQFVVEPSEATNRSVQAHEVHQK	119
MOUSE	1	MASFHLDHLK-NYHRRYCRSSRAPNIHTKKGQNMLEILQDCFEDQSKAS-FLDDFTESLTSSTQKKKANYSSSSKCPESHKPVVSSRTGEASLQASAEPEAAGGSVQANEVHH	116
HUMAN	120	ILATDVSSKNTPDSKKISSRNINDHSEADEEFYLSVGSPSVLLDAKTSVSNQVIPSQAQRETYTFENSVNMLPSSTEVSVKTKKRLNFDKVMKKIEIDNKVSDFEEDKTEGQERK	238
MOUSE	117	-----GASDELDCVGSPPVLLDANVNTLQKAASPAQGRVAVSVRSRSPVDRQASNKNISFKTRKRLNFDKVTLSAETENSVLQVEDNLSKQGEET	208
HUMAN	239	PSGSSQNRIRDSEYEQRQAKKSFSTLFLETYVKKSESSPIVRHAATAPPHSCPPDDTKLIEDEFIIDESDQSFASRSWITIPRKAGSLKQRTISPAESTALLQGRKSREKHHNLPKPT	357
MOUSE	209	SSEITQK-RDDLSSDVQSRSKNFSELFLETYVKKSKSSSVVRHTAAVFPSPPPSDMKLLEDEFIIDRSRDSFSSRLWVMIPSKDRHLSAHKPSP-ENTALLQGGKSREKSHLSAMT	325
HUMAN	358	LANDKHSKHPHVETSQPSDKTVLDTSYALIGETVNNYRSTKYEMYSKNAEKPSRSKRTIKQKRRKFMAKPAEEQLDVQGSKDEIHITSHITQDEFQRNSDRNMEHEEMGNDQVSKK	476
MOUSE	326	FARNTQSDKAHPIEEAQLSVEENPA--TTCTDELENDCRSPENKMQSETAKTPPAWERTIKQSQRRVSKPKAAEELRKQSSWENSNSVNTGQDKLQINSKRNMKDCGEEVRNEPNPKK	441
HUMAN	477	QMPVPVGSKR--SSTRKDKKEESKKRFSSESKNKLVPEEVTSTVTKSRRISRRPSDWWVVKSEESPVVYSNSVSRNELMHHNSRKSSTKKTNQSSKNIKKTIPLKRQKATKGNQRVQK	593
MOUSE	442	QKPALENKKKTNSTQTNKEKSGKFFSGGSKNKFVPKKVTLSRRSCRISQRPSEWVRVKSDESSVDRNPSKFNNSPVVYPNKKKQIKRNHV-SKRAGK--PGSSKRQKTEMSPRVQK	557
HUMAN	594	FLNAEFGGGIV-GHDEISRCSLSEPLESDEADLAKKKNLDCSRSTRSKNEDNIMTAQNVPLKQTSGYTCNIPTESNLDSGEHKTSLVEESGPSRLNNYLSMGKNDVD--DE-EVHG	708
MOUSE	558	SLNVKDSGGTVSGHDDTSR-SQRKPLKIEADP-TQKSLAISRPKRGCKYRNNVMTSPNVHLKSHTEEYTSKTMESASNSEMSKRSVWEESGPSRFKNY-EMPGSSNSEMGDEQDQKS	673
HUMAN	709	SSDDSKQSKVIPKNIHHKLVLPSTPNVRRTRKRLKPLEYWRGERIDYQGRPSGGFVISGVLSPDTISSKRKAKENIGKVNKSNKRIICLDNDERKTNLMVNLGIPLGDPLQPTRV	827
MOUSE	674	LHFTTRSNMVPDKLHHLKLVLPNSPNVRRSNRIRLKPLEYWRGERVDYQESSGQLVLEIISPSVPTKIKAQRLGKVNKKVTKPHTLNSHE-KAKMELPLDMRLGDFQATLA	790
HUMAN	828	KDPETREIILMDLVRPQDTYQFVFKHGELKVKYKTLDTFFSTGKLLGPOEEKGQHVGDQILVYFVNFQDGLLCTLHETPYILSTGDSFYVPSGNYNINKLRNEESVLLFTQIKR	943
MOUSE	791	KDPEIAELVPMDLIRPRDTRFVFEQHLGKVFKTLDTIYFSTGKLVLPQYEEKGQHVGDQILVYFVNFQDGLLCTLHETPYKLTGDSFYVPSGNYNINKLNLNVESSLFTQIKR	906

RNA binding region (Human) (Wong, Brettingham-Moore et al., 2007)  
 RNA binding region (Mouse) (CatRapid predicted domain 1)

B



C

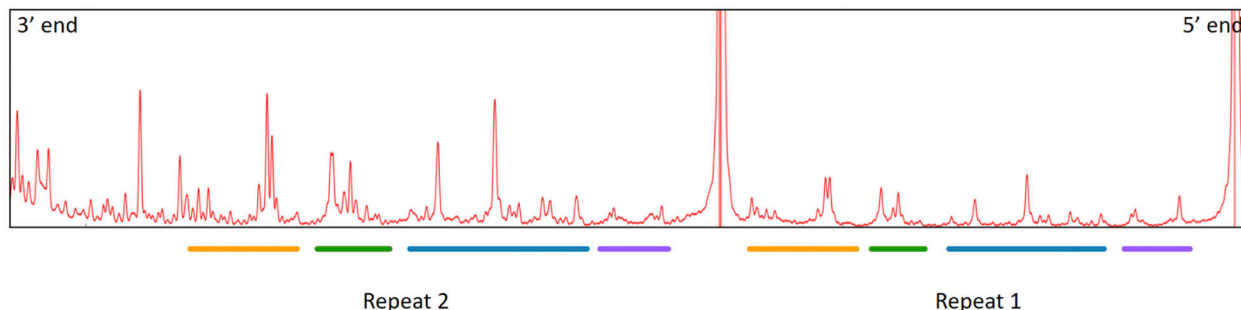


Figure S2. **CENPC amino acid sequences, predicted alternative structures of single repeat MinSat RNA, and raw spectra of 2 repeats MinSat RNA.** (A) Amino acid sequence alignment of human and mouse CENPC proteins. Orange and green highlighted regions are the potential RNA binding regions from human and mouse (CATRAPID 1), respectively. (B) Alternative secondary structures of MinSat RNA predicted based on SHAPE reactivity. Note the reoccurring stem-loop structure in various structure predictions (green dotted-lined). The numbers indicate the nucleotide position relative to the start of the primer extension reaction. (C) Original spectra of SHAPE analysis of 2 repeats of forward MinSat RNA. The colored underlines highlight two repetitive patterns. The Y axis shows SHAPE reactivity and the X axis represents relative nucleotide position from 3' end to 5' end.

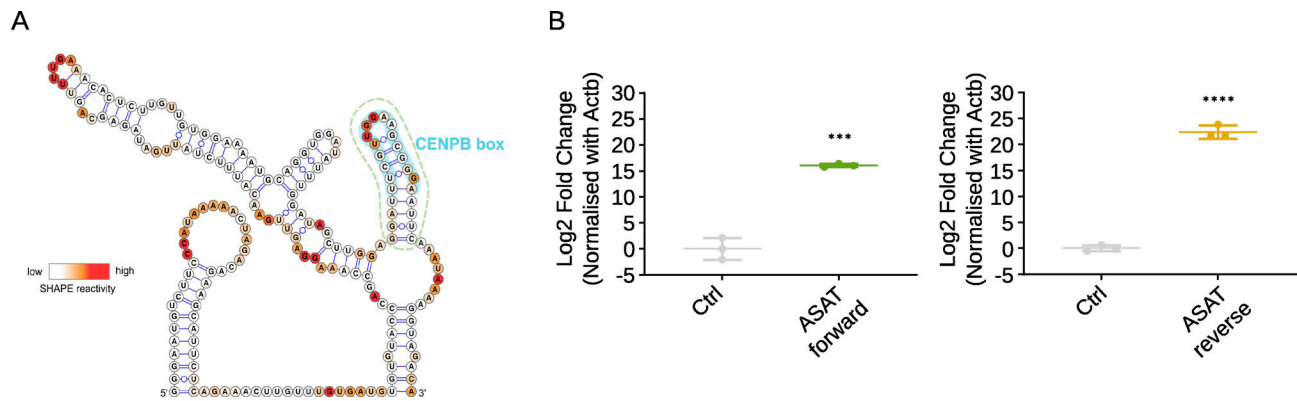


Figure S3. **SHAPE of human ASAT RNA and RT-qPCR of human ASAT overexpression in ESCs.** **(A)** SHAPE-derived secondary structures of 1 repeat human chr21 ASAT forward transcripts. SHAPE reactivity is indicated by the color key as indicated on the left. Note the prominent CENPB box stem-loop structure in the forward transcript, labeled in light blue. The green-dotted line indicates the 24 nt-long stem-loop structure conserved between mouse and human transcripts. The numbers indicate the nucleotide position relative to the start of the primer extension reaction. **(B)** RT-qPCR for human chr21 ASAT RNA in ESCs transfected with forward (left panel) or reverse (right panel) chr21 ASAT expressing plasmids using strand-specific primers. Ctrl, empty pCAG vector control; solid line, mean Ct values of  $N = 3$  independent experiments (dots) normalized to  $\beta$ -actin mRNA and control; error bars, standard deviation, pair-wise comparisons with control (empty pCAG vector) P values \*\*\*  $\leq 0.001$ , \*\*\*\*  $\leq 0.0001$ .

Provided online are three tables. Table S1 shows 2'-MOE antisense oligos for MinSat RNA knockdown. Table S2 shows strand-specific and non-strand-specific qPCR primers applied for MinSat expression level. Table S3 shows antibodies used for immunofluorescence.