



# Ectopic expression of pericentric HSATII RNA results in nuclear RNA accumulation, MeCP2 recruitment, and cell division defects

Catherine C. Landers<sup>1</sup> · Christina A. Rabeler<sup>2</sup> · Emily K. Ferrari<sup>3</sup> · Lia R. D'Alessandro<sup>2</sup> · Diana D. Kang<sup>4</sup> · Jessica Malisa<sup>5</sup> · Safia M. Bashir<sup>2</sup> · Dawn M. Carone<sup>2</sup> 

Received: 11 May 2020 / Revised: 16 January 2021 / Accepted: 19 January 2021 / Published online: 13 February 2021  
© The Author(s) 2021

## Abstract

Within the pericentric regions of human chromosomes reside large arrays of tandemly repeated satellite sequences. Expression of the human pericentric satellite HSATII is prevented by extensive heterochromatin silencing in normal cells, yet in many cancer cells, HSATII RNA is aberrantly expressed and accumulates in large nuclear foci *in cis*. Expression and aggregation of HSATII RNA in cancer cells is concomitant with recruitment of key chromatin regulatory proteins including methyl-CpG binding protein 2 (MeCP2). While HSATII expression has been observed in a wide variety of cancer cell lines and tissues, the effect of its expression is unknown. We tested the effect of stable expression of HSATII RNA within cells that do not normally express HSATII. Ectopic HSATII expression in HeLa and primary fibroblast cells leads to focal accumulation of HSATII RNA *in cis* and triggers the accumulation of MeCP2 onto nuclear HSATII RNA bodies. Further, long-term expression of HSATII RNA leads to cell division defects including lagging chromosomes, chromatin bridges, and other chromatin defects. Thus, expression of HSATII RNA in normal cells phenocopies its nuclear accumulation in cancer cells and allows for the characterization of the cellular events triggered by aberrant expression of pericentric satellite RNA.

**Keywords** Noncoding RNA · Pericentric heterochromatin · Chromatin instability · Cancer biology · Human

## Introduction

Nearly 50% of the human genome consists of repetitive DNA sequence elements. These include transposable element-based repeats such as long and short interspersed nucleotide elements (LINES and SINES) and simple sequence repeats (tandem DNA) such as ribosomal DNAs and satellite sequences (Richard et al., 2008). Distinct classes of repeat elements are

uniquely localized within the genome. While LINE and SINE elements are heterogeneously interspersed genome-wide, tandemly repeated elements are more positionally defined. Satellite DNAs, defined by tandemly repeating units of DNA, localize primarily to centric or pericentric regions and to telomeres of all chromosomes, and are classified based on their sequence and length of repeating unit. The major classes of human satellite DNA are the following: (1) alphoid (alpha satellite;  $\alpha$ -Sat) DNA resident at all centromeres; (2) telomeric repeats; (3) beta satellite DNA ( $\beta$ -Sat) on acrocentric chromosomes; and (4) satellites HSATI, HSATII, and HSATIII found at the pericentric regions of a subset of chromosomes.

Large blocks of HSATII DNA are located on 11 human chromosomes, with the largest arrays residing within the pericentromeres of human chromosomes 1 and 16. Satellite DNA blocks may span several megabases and consist of tandemly repeated elements ranging from five to several hundred nucleotides (Richard et al., 2008). Length and sequence composition of satellites, including HSATII, varies in the human population and may be used to study human genetic variation (Miga, 2019); however, the full extent and diversity of HSATII sequences is currently unknown due to its location

---

Catherine C. Landers and Christina A. Rabeler contributed equally to this work.

---

✉ Dawn M. Carone  
dcarone1@swarthmore.edu

<sup>1</sup> Department of Nutritional Sciences, University of Connecticut, Storrs, CT, USA

<sup>2</sup> Department of Biology, Swarthmore College, Swarthmore, PA, USA

<sup>3</sup> Children's Hospital of Philadelphia, Philadelphia, PA, USA

<sup>4</sup> Division of Pharmaceutics and Pharmacology College of Pharmacy, Ohio State University, Columbus, OH, USA

<sup>5</sup> Stanford University School of Medicine, Stanford, CA, USA

within large sequence assembly gaps in the human genome (Altemose et al., 2014; Eichler et al., 2004). HSATII is roughly defined as a ~26 bp repeat, but exhibits significant heterogeneity in sequence and has been poorly studied, despite its prominence in the human karyotype (Tagarro et al., 1994) and abundance in the human genome. A computational effort to generate a satellite reference database (Altemose et al., 2014) successfully demonstrates the ability to cluster HSAT sequences into subfamilies, which are likely to reside on individual chromosomes, but a comprehensive HSATII map has yet to be completed. In contrast, significant progress has been made in mapping alpha satellite arrays, consisting of 171-bp monomer repeats, particularly on the X (Miga et al., 2019) and Y (Jain et al., 2018) chromosomes, to generate full telomere-to-telomere linear chromosome sequence.

Due to their localization near centromeres, satellite sequences are likely to be subject to strict regulation to maintain both genetic and epigenetic stability. While expression of pericentric satellites has been observed in many species, including yeast, plants, and mammalian cells, it is clear that when satellites are expressed, their expression is highly regulated (reviewed in (Hall et al., 2012; Perea-Resa and Blower, 2018; Smurova and De Wulf, 2018)). The increased presence of satellite RNA has recently emerged as an indicator of instability, as demonstrated by satellite overexpression in cancer cells and tumor tissues (Hall et al., 2017; Ting et al., 2011; Zhu et al., 2011). Whereas  $\alpha$ -sat is expressed at low levels in normal cells (Hall et al., 2017; McNulty et al., 2017) and increases expression in cancer, HSATII expression is restricted to cancer cells, and thus, the presence of HSATII RNA is a potential biomarker of cancer (Hall et al., 2017; Ting et al., 2011). Overexpression of chromosome-specific pericentric HSATII loci (e.g., Chr 7) occurs within nuclei in which other HSATII chromosomal locations (e.g., Chr 1) are not expressed, due to their accumulation of repressive polycomb proteins, thus indicating that HSATII expression is regulated in a locus-specific manner (Hall et al., 2017). Accumulation of HSATII RNA occurs *in cis*, with the RNA accumulating in cancer-associated satellite transcript (CAST) bodies, adjacent to the HSATII locus from which it is transcribed. Within a given tumor or cancer cell line, expression of HSATII maintains this locus-specific expression pattern in most cells (Hall et al., 2017).

HSATII RNA within CAST bodies recruits and directly binds MeCP2 and its protein-binding partner, Sin3A (Hall et al., 2017). MeCP2 is canonically classified as a DNA methyl-binding protein (Hite et al., 2009); however, more recent evidence indicates that MeCP2 also contains an intrinsically disordered domain, which has the capacity to bind RNA (Castello et al., 2016), and can function as a transcriptional activator (Hite et al., 2009). MeCP2 is mutated in Rett syndrome, where it is known to bind transcripts in the brain and affect alternative splicing, thus suggesting MeCP2 has multifunctional roles in gene regulation and splicing (Young et al.,

2005). MeCP2 is recruited to HSATII RNA accumulations and may be sequestered in CAST bodies in cancer cells, but the dynamics, consequences, and direct role of HSATII RNA in this recruitment are currently unknown.

Inappropriate expression of satellite RNAs can be an indicator of heterochromatic instability, which may be increasingly common in cancers, and has wide-ranging implications (Carone and Lawrence, 2013). Given that maintenance of centric/pericentric heterochromatin is key to normal chromosome segregation, structural changes to the underlying chromatin may result in functional changes, including satellite expression and aberrant cell division. Supporting this, prior studies analyzing the effect of satellite expression have demonstrated that forced expression of  $\alpha$ -sat transcripts leads to cell division defects, chromosomal instability, and aneuploidy in normal cells (Chan et al., 2017; Ichida et al., 2018; Zhu et al., 2018). Paradoxically, some  $\alpha$ -sat expression is observed in normal cells (Hall et al., 2017) and low levels of  $\alpha$ -sat expression are thought to be necessary for centromere function and constitutive heterochromatin maintenance (Johnson et al., 2017; McNulty et al., 2017). Further, some studies have found no phenotype following forced expression of satellite sequences (Ideue et al., 2014; Rošić et al., 2014). Thus, the functional role of satellites has been an enigma, despite their prevalence and structural conservation within pericentric regions. This conserved, tandemly repetitive structure of centric and pericentric satellites is in direct juxtaposition to the demonstrated lack of sequence conservation between species (Henikoff et al., 2001) and the discovery of functional neocentromeres lacking satellite sequences (Voullaire et al., 1993). Pericentric satellites, in particular, have little known function despite their abundance within the pericentric sequence landscape. It has become clear that the mechanisms governing transcriptional silencing of pericentric satellites are complex, with many histone modifications contributing to maintaining repression. Thus, a ubiquitous role for pericentric repeats, the mechanisms governing their regulation, and the satellite sequences that comprise them, has remained poorly understood. Given that pericentric satellite expression is misregulated in a wide range of cancer cell lines and tissues, and the chromatin structure of pericentric satellites is compromised in both cancer and senescent cells (Brückmann et al., 2018; Hédouin et al., 2017; Slee et al., 2012; Swanson et al., 2013; Tasselli et al., 2016), the direct effect of aberrant pericentric satellite expression warrants investigation.

Here, we established cell lines stably expressing HSATII RNA and  $\alpha$ -sat RNA in order to test the long-term effects of expression of these distinct satellite transcripts. We demonstrate that ectopically expressed HSATII RNA accumulates in nuclear bodies reminiscent of CAST bodies in both HeLa and primary human fibroblast cells. The focal accumulation of HSATII transcripts is unique to HSATII, as stable ectopic expression of  $\alpha$ -sat RNA does not result in focal RNA

accumulation in the nucleus. Further, HSATII RNA accumulates *in cis*, immediately adjacent to the genomic integration site, and recruits MeCP2 to these nuclear foci. Expression of both  $\alpha$ -sat and HSATII transcripts leads to cell division defects including chromatin bridges, blebbing, and micronuclei formation, thus indicating that expression of both centromeric and pericentric satellite sequences has the potential to generate chromosomal instability and aberrant cell division. The study of the effect of HSATII expression represents a unique opportunity to shed light on the function of pericentric satellite sequences more generally, in addition to testing the functional consequences of expression of pericentric satellite sequences in cancer.

## Results

### Establishment of a cell culture system to study the effect of HSATII RNA expression

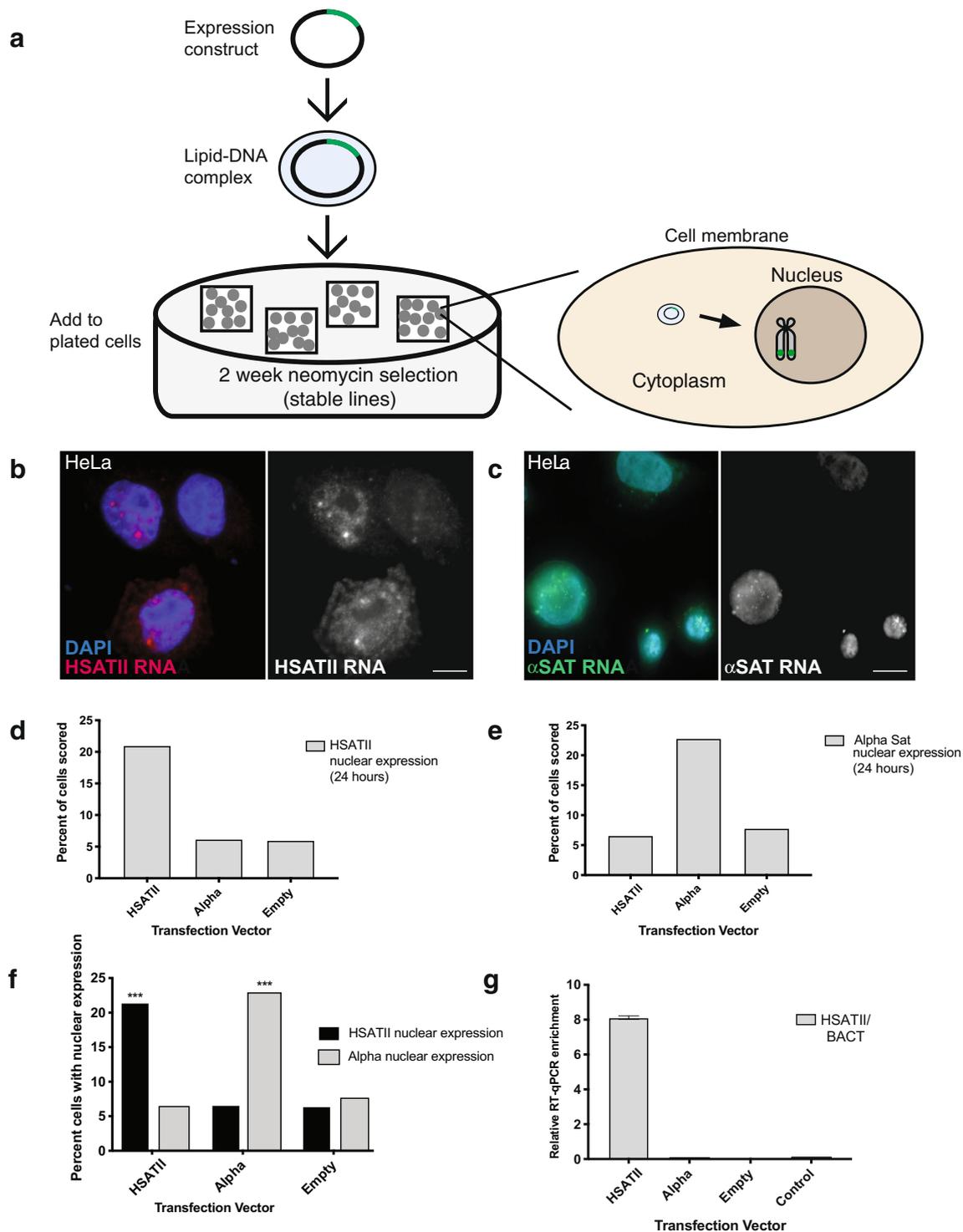
Expression of HSATII RNA in cancer cells has previously been observed in both cancer cells grown in culture and from tissue biopsies (Hall et al., 2017; Ting et al., 2011). While there are distinct loci that harbor HSATII DNA sequences within the pericentric regions of roughly 11 human chromosomes, only a subset of these HSATII sequences are expressed, and the HSATII RNA transcripts expressed from these loci remain *in cis* in CAST bodies. HSATII sequences resident on human chromosomes 7 and 10 are among those sequences displaying a preference for expression in a variety of different cancer cell lines and tissues (Hall et al., 2017). In order to study the direct effect of expression of HSATII RNA, we developed a cell culture model to stably express HSATII sequence derived from Chr 7 in cell lines that do not endogenously express HSATII. To further examine the effect of HSATII expression irrespective of its location of expression, stable cell lines were created in which the Chr 7 HSATII-expression construct had been randomly integrated into the genome.

An HSATII cDNA sequence derived from Chr 7 was cloned into a plasmid designed for mammalian expression and stable integration, containing a CMV promoter and a neomycin selectable marker (Fig. 1a). HeLa cells, while a cancer cell line, do not endogenously express HSATII RNA (Hall et al., 2017); thus, initial transfection experiments were conducted in HeLa cells due to their ease of transfection by lipid-mediated transfection. To test for HSATII-specific effects, control vectors containing an  $\alpha$ -sat sequence derived from Chr 4 and no insert (empty vector) were also used concurrently to transfect HeLa cells (Fig. S2). Cells were assayed for transient satellite expression 24 h after transfection by RNA FISH and RT-qPCR. Twenty-four hours after transfection, approximately 20% of HSATII-transfected HeLa cell nuclei displayed nuclear accumulations of HSATII RNA compared

to less than 5% of  $\alpha$ -sat and empty vector control transfected cells (Fig. 1b, d). Nucleoplasmic and cytoplasmic diffuse expression was also observed at this early timepoint, likely due to high levels of expression driven by the CMV promoter. Cells transfected with  $\alpha$ -sat displayed a similar level of expression, with roughly 23% of cells displaying  $\alpha$ -sat RNA by RNA FISH (Fig. 1c, e). However, a striking difference was observed in the distribution of HSATII and  $\alpha$ -sat RNA in the nucleus. Distinct focal accumulations of HSATII RNA (2–3 per nucleus on average) were observed (Fig. 1b), while  $\alpha$ -sat RNA appeared as a diffuse, primarily nuclear RNA signal (Fig. 1c). Expression of HSATII or  $\alpha$ -sat was dependent on transfection with the respective insert-containing vector, thus demonstrating construct delivery specificity, which was observed upon three independent transient transfections. Further, the percentage of cells expressing the desired sequence insert was significantly different from controls (empty vector) (Fig. 1f). RT-qPCR also confirmed high levels of HSATII expression in HSATII-transfected cell lines compared to  $\alpha$ -sat transfected and controls (Fig. 1g). Since RT-qPCR was performed from total cellular RNA, results here cannot distinguish between nuclear RNA accumulations and diffuse RNA (nuclear or cytoplasmic), thus the greater than eightfold increase in HSATII expression shown for one transfection (Fig. 1g), likely illustrates the total amount of HSATII overexpression compared to  $\alpha$ -sat and control cells.

While HeLa represents an easily transfected cell line that does not express HSATII RNA, they are cancer cells that have been maintained in culture since the 1950s, thus are not representative of “normal” cells. Following successful demonstration of transient HSATII and  $\alpha$ -sat expression in HeLa cells, stably transfected cell lines were generated for both HeLa and a primary (non-transformed) human fibroblast cell line to ensure that the results observed were not simply an effect of expressing satellite RNA in cancer cells, in which the nuclear and chromatin environment can be drastically different from normal cells (Carone and Lawrence, 2013; Zink et al., 2019). Tig-1 cells are female primary fibroblasts maintained at low passage number that retain an inactive X chromosome and a stable karyotype ( $2n=46$ ) (Ohashi et al., 1980), do not express HSATII (Hall et al., 2017), and can be transfected using lipid-mediated transfection. A lipid-mediated transfection protocol was optimized for successful transfection of Tig-1 primary fibroblasts, to ensure high transfection efficiency in addition to high cell viability (Fig. S1).

Following 2 weeks of neomycin selection, mock transfected and control HeLa and Tig-1 cultures displayed complete cell death, while a subset of viable cells (ranging from 10–15%) harboring HSATII,  $\alpha$ -sat, and empty vector integrants were selected and further expanded for a period of several weeks. Cells from three independent stably transfected cell lines were fixed for DNA/RNA FISH and pelleted and harvested for RNA extraction at weekly intervals to assay satellite expression (Fig. 1a). To test for genomic integration



**Fig. 1** Transient transfection of satellite expression results in nuclear satellite RNA accumulation. **(a)** Transfection scheme for transient and stable integration expression. A plasmid harboring HSA TII,  $\alpha$ -sat ( $\alpha$ sat), or no insert (empty vector) is introduced to cultured HeLa or Tig-1 primary fibroblast cells via lipid-mediated transfection and cells are then fixed on coverslips or harvested for RNA isolation. Stable cell lines are further selected with neomycin (G418) for 2 weeks prior to fixation or harvesting. Twenty-four hours after transfection, nuclei are scored for expression of **(b)** HSA TII and **(c)**  $\alpha$ -sat RNA signal by RNA

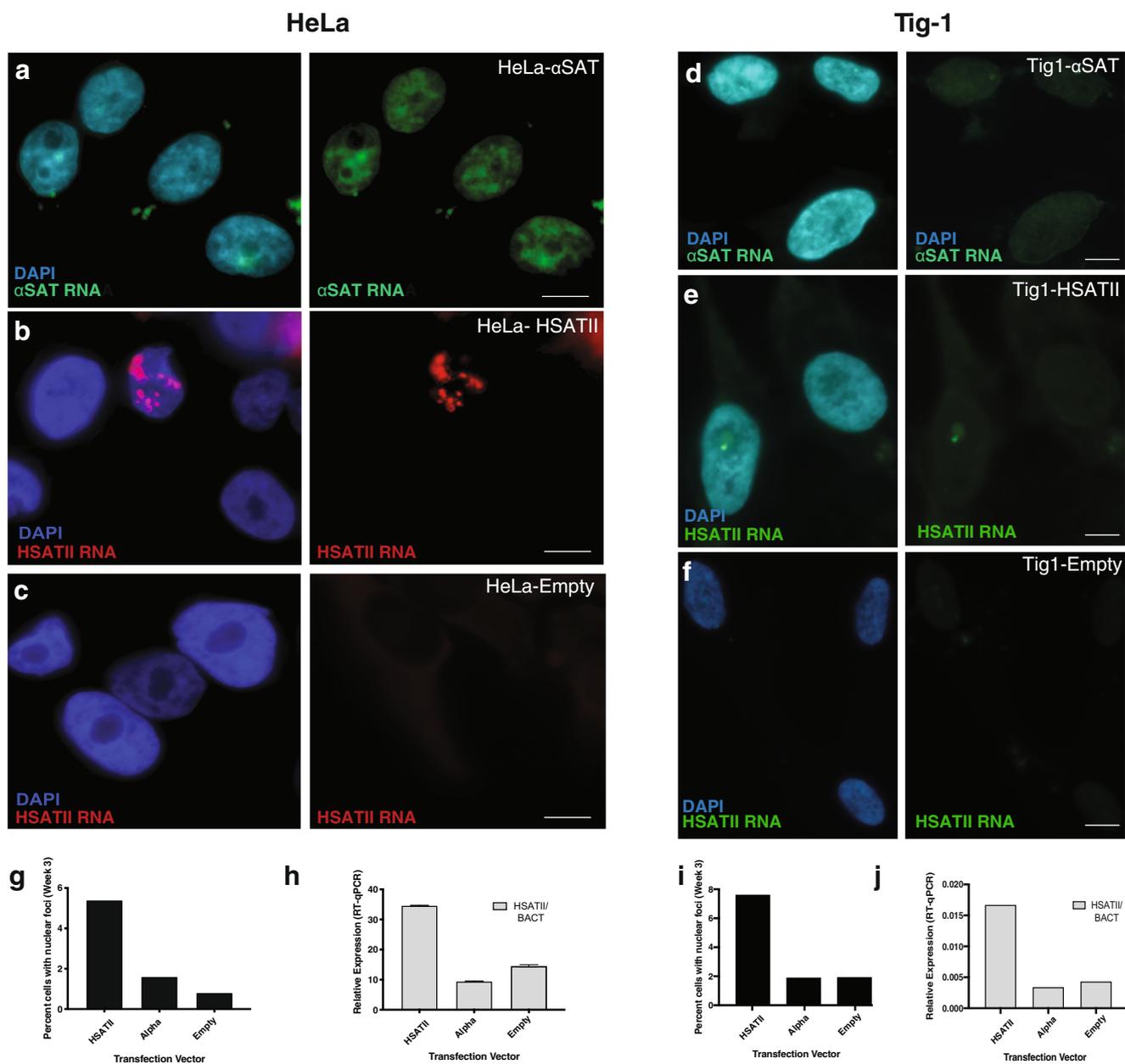
FISH. Percent of cells (out of 500 nuclei) with **(d)** HSA TII RNA nuclear expression and **(e)**  $\alpha$ -sat nuclear expression. **(f)** Nuclear RNA signal detected by RNA FISH is dependent on the sequence harbored within the transfected construct. Asterisks denote significant differences from empty vector transfected cells by Chi-square test,  $p < .0001$ . **(g)** qRT-PCR of HSA TII RNA 24 h after transfection. Expression values shown relative to housekeeping gene expression ( $\beta$ -actin) for each transfection condition. Scale bar, 5  $\mu$ m

of the expression vector, DNA hybridization was performed using a biotin-labeled probe complementary to the plasmid (pTarget) backbone, which indicated that the majority of cells with detectable signal had one or two sites of integration (Fig. S3). Analysis of chromosome spreads harvested from transfected cell lines confirmed random integration of the HSATII expression vector, with the majority of spreads with hybridization signal harboring one or two integration sites, which were randomly distributed on chromosomes (Fig. S4). Three weeks following transfection, HeLa cells retained the same pattern of satellite expression, with  $\alpha$ -sat RNA being diffuse and nuclear (Fig. 2a). In contrast, ~5% of nuclei in HSATII-transfected HeLa cells displayed nuclear accumulations of HSATII RNA. This pattern was similarly observed in Tig-1 primary fibroblasts, with independent transfections demonstrating 7–28% of HSATII transfected cells harboring nuclear accumulations of HSATII RNA (Fig. 2e, i) and  $\alpha$ -sat RNA displaying a diffuse nuclear signal by RNA FISH (Fig. 2d). RT-qPCR further confirmed higher levels of HSATII RNA in HSATII-transfected HeLa (Fig. 2h) and Tig-1 (Fig. 2j) cells. Of note, while transiently transfected cell lines displayed some cytoplasmic satellite RNA, stably transfected cell lines had very little cytoplasmic satellite RNA ( $\alpha$ -sat or HSATII), suggesting that the vast majority of satellite RNA transcribed ectopically in these cell lines is retained in the nucleus. Taken together, these results indicate that stably expressing HeLa and primary fibroblast cell lines can be used to examine the effect of ectopic satellite RNA (HSATII and  $\alpha$ -sat) expression in cells that do not normally express HSATII. Further, results from transient and stably transfected cell lines indicate that ectopic HSATII RNA accumulates in nuclear foci, while ectopic  $\alpha$ -sat RNA is primarily diffuse in the nucleus, suggesting that when these two satellite RNAs are expressed in both of these cell types, their transcripts may behave in very different ways within the nuclear environment, irrespective of their site of expression within the genome.

### HSATII RNA accumulates adjacent to the location from which it is expressed and recruits MeCP2

In cancer cells that express HSATII RNA, the RNA accumulates *in cis* immediately adjacent to their site of transcription (Hall et al., 2017). Therefore, we asked whether the accumulated HSATII RNA foci in stably transfected cell lines also remain *in cis*. Sequential RNA FISH followed by DNA FISH confirmed HSATII RNA accumulated adjacent to the site into which the expression construct (pTarget backbone) was integrated into the genome in both HeLa (Fig. 3a–b) and Tig-1 cells (Fig. 3e–f). While not all sites of integration had a focal accumulation of RNA, when HSATII RNA was detected, it was adjacent to the location in which the vector had integrated in 53% of nuclei scored. For example, out of 53 Tig-1 HSATII stably transfected nuclei with observed HSATII DNA

integration, 15 of those also had focal HSATII RNA signal, with 8 of those demonstrating an adjacent localization via sequential RNA-DNA FISH experiments. An example of accumulated HSATII RNA without adjacent localization to the pTarget backbone is shown in Figure S5. It is likely that there are genomic locations into which the expression vector has integrated that are not permissible to transcription, that will not tolerate accumulation of the RNA, or for which the CMV promoter may no longer be active that may account for sites that have no detectable HSATII RNA signal. It is also possible that accumulation of HSATII RNA occurs only when proteins are recruited to the RNA to form CAST bodies (Hall et al., 2017), as it is currently unknown whether the transcripts alone accumulate. We observed that focal accumulation of ectopically expressed HSATII RNA is reminiscent of the focal accumulations in cancer cells (CAST bodies) in both their appearance and in the pattern in which they form near their site of transcription. Therefore, we sought to examine whether HSATII RNA foci recruited MeCP2, one of the protein components within HSATII CAST bodies in cancer cells (Hall et al., 2017), using co-immunofluorescence/RNA FISH. In HeLa cells, all HSATII RNA focal accumulations had an overlapping focal signal with MeCP2 (100% of those nuclei scored) (Fig. 3c–d). Colocalization analysis of HSATII RNA and MeCP2 in HeLa cells indicated a Pearson correlation coefficient of  $r \sim 1$ , indicating that these signals are nearly perfectly overlapping, despite being detected in different fluorescent channels (Fig. 3b). This was in contrast to a lack of colocalization of HSATII RNA and pTarget backbone DNA in HeLa cells ( $-0.5 < r < 0.7$ ) (Fig. 3b), as was expected based on their adjacency (Fig. 3a). In Tig-1 cells, some colocalization with MeCP2 was observed for a subset of HSATII RNA accumulations (Fig. 3g) (8% colocalization for one transfected cell line scored). Since not all HSATII RNA accumulations in primary fibroblasts recruited MeCP2, this may suggest that the cellular context may influence the potential for recruitment of MeCP2 into CAST bodies (Fig. S5b). It might also be possible that additional proteins are recruited to CAST bodies independently of MeCP2, or that HSATII RNA, alone, has the ability to condense into focal accumulations *in cis*. Additionally, there were many focal accumulations of MeCP2 in Tig-1 cells that did not overlap HSATII RNA accumulations (Fig. 3g). These MeCP2 foci were also observed in control cells (Fig. S5c) and likely represent normal nuclear accumulations of MeCP2 that do not overlap HSATII transcripts. These data suggest that ectopically expressed HSATII RNA accumulates *in cis* and remains in nuclear foci near, though not overlapping, their site of genomic integration. Further, these HSATII RNA foci recruit MeCP2, a protein known to be in CAST bodies in cancer cells. Differential MeCP2 recruitment in HeLa and primary fibroblast cells to HSATII RNA accumulations suggests MeCP2 colocalization may be dependent on both the cellular



**Fig. 2** Stable expression of satellite RNA results in focal nuclear accumulation of HSATII RNA in a subset of cells. **(a)** Following integration and neomycin selection for the expression construct (HSATII,  $\alpha$ SAT, empty), a subset of HeLa **(a–c)** and Tig-1 **(d–f)** nuclei exhibit satellite RNA expression by RNA FISH. While  $\alpha$ -sat RNA **(a, d)** is primarily diffuse in the nucleus, HSATII RNA accumulates in distinct nuclear foci **(b, e)** in both HeLa and Tig-1 cells. Cells transfected with an

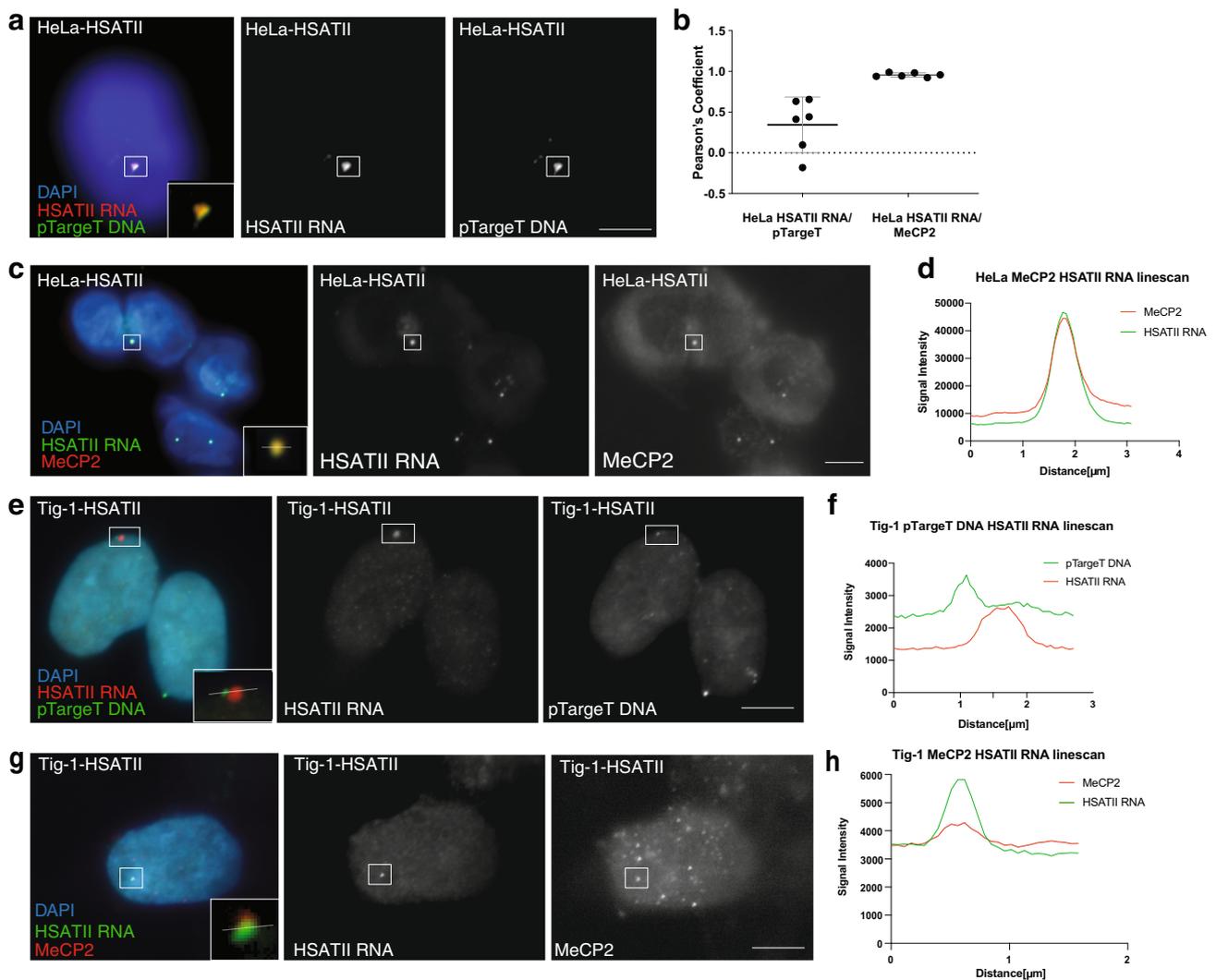
empty vector have no satellite RNA expression detectable by RNA FISH **(c, f)**. **(g, i)** Quantification of cells with HSATII nuclear expression 3 weeks following transfection in HeLa **(g)** and Tig-1 **(i)**. **(h, j)** RT-qPCR of HSATII RNA in transfected cell lines 3 weeks following transfection. Data is shown normalized to  $\beta$ -actin (housekeeping gene). Error bars show 95% confidence interval. Scale bar, 5  $\mu$ m

context and/or chromatin environment throughout the nucleus.

### Satellite RNA expression induces cell division defects and instability

Analysis of cell lines established to stably express HSATII RNA suggests that ectopically expressed HSATII RNA

accumulates in a similar manner to HSATII RNA that is endogenously expressed from pericentric regions. It has been demonstrated that aberrant  $\alpha$ -satellite expression can lead to cell division defects and chromosomal instability (Chan et al., 2017; Ichida et al., 2018; Zhu et al., 2018; Zhu et al., 2011). Therefore, we examined whether cell division defects could be observed in cells that were also forced to express pericentric HSATII, in direct comparison



**Fig. 3** HSATII RNA accumulates adjacent to site of integration and recruits MeCP2. **(a)** RNA hyb for HSATII RNA (red) followed by sequential DNA hyb using a probe to detect the pTargetT backbone (green) demonstrates HSATII RNA accumulates primarily adjacent to sites of pTargetT genome integration in HeLa cells stably expressing HSATII 3 weeks after transfection. All boxes with white border indicate regions of interest with zoomed image at bottom right. Individual channels images shown in black and white. **(b)** Scatter plot of Pearson correlation coefficients for individual nuclei ( $n=10$ ) hybridized for sequential HSATII RNA/pTargetT DNA and HSATII RNA with MeCP2 co-IF (HSATII RNA/MeCP2). Black lines indicate mean; grey error bars indicate 95% confidence interval. **(c)** Co-HSATII RNA hyb and IF for MeCP2 indicate a near perfect colocalization of HSATII RNA and MeCP2 signals within HeLa nuclei 3 weeks following transfection, as

demonstrated by Pearson correlation coefficients plotted in **(b)**. **(d)** Dual channel linescan of inset shown in **(c)** demonstrating colocalization of MeCP2 (red) and HSATII RNA (green). **(e)** RNA hyb for HSATII RNA (red) followed by sequential DNA hyb using a probe to detect the pTargetT backbone (green) demonstrates HSATII RNA accumulates adjacent to, and not overlapping, sites of genome integration in Tig-1 cells stably expressing HSATII 3 weeks after transfection. **(f)** Dual channel linescan of inset in **(e)** demonstrating the HSATII RNA (red) and pTargetT DNA (green) signals are adjacent, and not overlapping. **(g)** Tig-1 fibroblasts stably expressing HSATII RNA for 3 weeks demonstrate overlapping signals for HSATII RNA and MeCP2. Box with white border indicates region of interest with zoomed image at bottom right. **(h)** Dual channel linescan of inset shown in **(g)** demonstrating colocalization of MeCP2 (red) and HSATII RNA (green). Scale bar, 5  $\mu\text{m}$

to cells expressing ectopic  $\alpha$ -sat RNA. Stable cell lines expressing HSATII and  $\alpha$ -sat RNA were compared to cells with stable integration of the pTargetT backbone alone (empty vector) and control (untransfected) cells at identical passage numbers. Two weeks following transfection, some cell division defects were observed in a small percentage of cells; however, this was not significantly different from

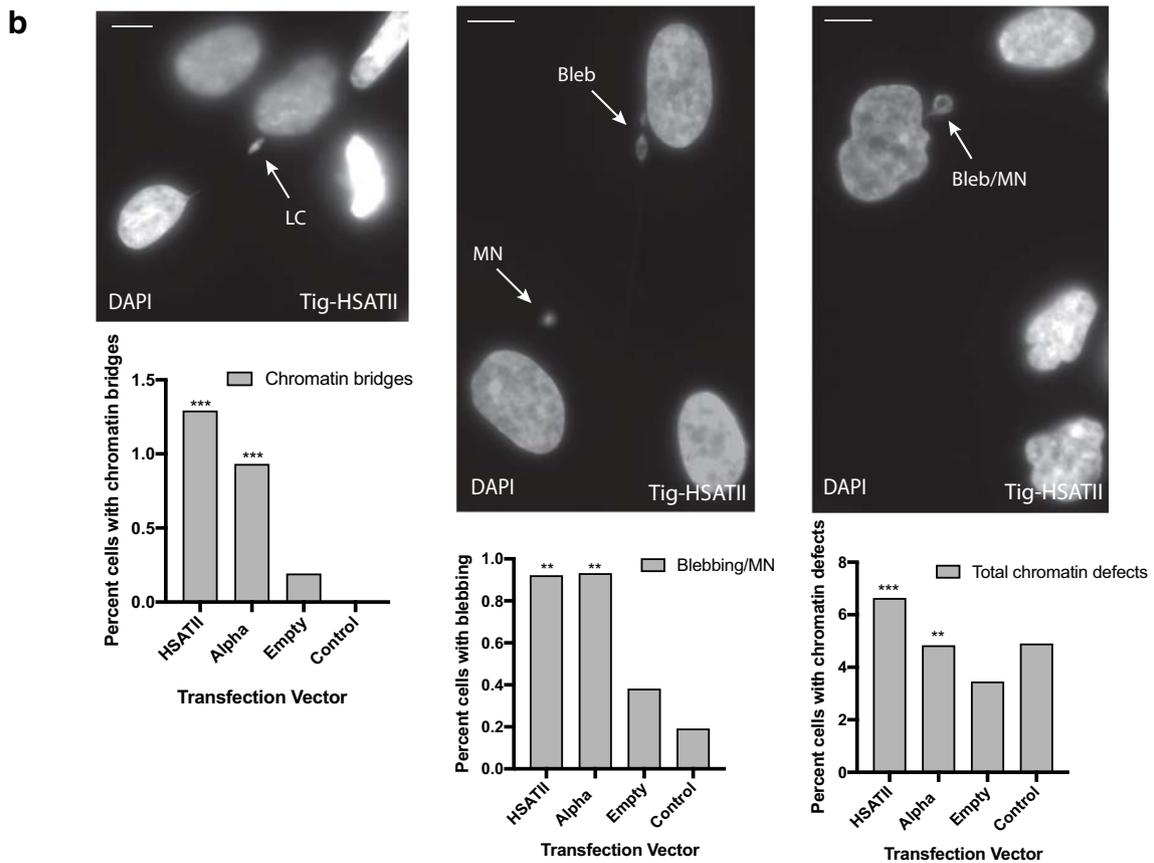
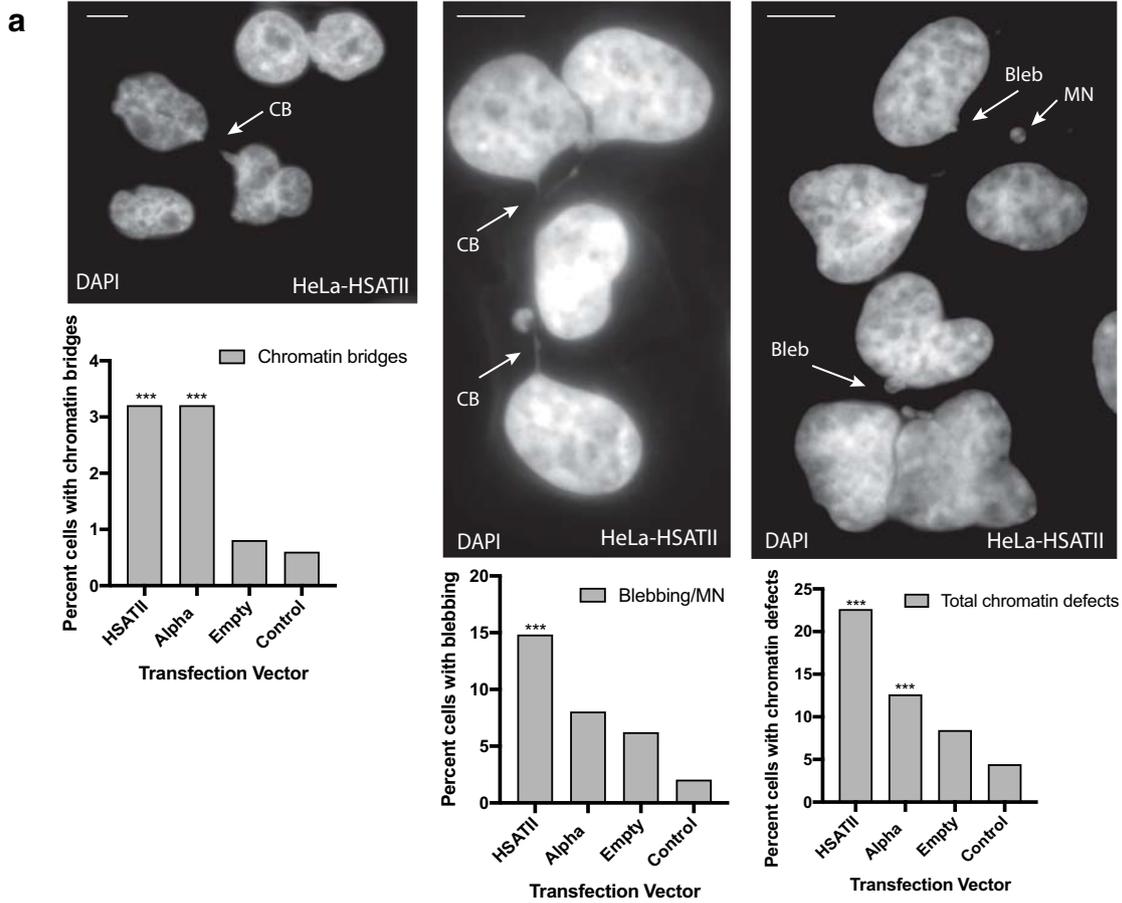
control cells. In contrast, at 21–28 days following transfection, an increase in chromatin and cell division defects was observed, including an increase in the presence of chromatin bridges (CB), nuclear blebbing/micronuclei (MN), lagging chromosomes (LC), and abnormally shaped nuclei (Fig. 4a) in HeLa cells. While a significant increase in chromatin bridges and overall chromatin defects (sum of

all defects) was observed in both HSATII and  $\alpha$ -sat expressing cell lines compared to empty vector transfected cells, an increase in blebbing alone was significant (Chi-square,  $p < .05$ ) only in HSATII-expressing HeLa cells. As expected for karyotypically normal fibroblast cells, the frequency of total cell division defects in Tig-1 was much lower than in HeLa (Fig. 4b). Observed cell division defects were significantly different for all HSATII or  $\alpha$ -sat expressing Tig-1 cell lines 24 days after transfection compared to empty vector transfected cells, including increases in chromatin bridges and nuclear blebbing/micronuclei, which are only very rarely seen (1 cell per 500 cells) in this primary fibroblast cell line when normally maintained at low passage numbers. Similar levels of cell division defects were observed in  $\alpha$ -sat and HSATII stably expressing cells established from independently transfected and maintained HeLa and Tig-1 cell lines (Fig. S6). Primary fibroblast cells exhibited increased cell division defects irrespective of which satellite RNA (HSATII or  $\alpha$ -sat) was expressed (Fig. 4b and Fig. S6d–f), whereas HeLa cells only exhibited a significant increase in blebbing/MN and abnormally shaped nuclei in HSATII-expressing cells, suggesting that there is an HSATII-expression-dependent effect in HeLa cells that is not present in primary fibroblasts (Fig. 4). This is intriguing in light of previous evidence that HeLa cells exhibit high levels of  $\alpha$ -sat expression when normally maintained in culture (Hall et al., 2017). Thus, it is possible that HeLa cells respond more marginally to increased  $\alpha$ -sat expression, whereas the presence of HSATII RNA induces a more drastic range of phenotypes. In contrast, untransfected Tig-1 cells exhibit less  $\alpha$ -sat expression (Hall et al., 2017), thus may be more sensitive to expression of alpha satellite in the nucleus. Supporting this, karyotypic analysis of stably expressing Tig-1 cells maintained in culture for 60 days after transfection revealed that while the majority of empty vector transfected and HSATII expressing cells had a normal karyotype ( $2n = 46$ ),  $\alpha$ -sat expressing cell lines had an increased number of chromosome spreads exhibiting aneuploidy (Fig. S7). Forced alpha satellite expression has been previously shown to induce cell division defects and generate chromosomal instability (Chan et al., 2017; Ichida et al., 2018; Zhu et al., 2018; Zhu et al., 2011); therefore, similar mechanisms may be induced upon aberrant  $\alpha$ -sat expression in primary fibroblasts. Further investigation will be required to determine the mechanism by which satellite expression impacts cell division and results in the formation of chromatin bridges and micronuclei, but the observed increase in these cell division aberrations in cells stably expressing  $\alpha$ -sat or HSATII suggests that expression of either pericentric or centromeric satellite sequences induces cell division defects.

## Discussion

In order to determine the consequences of the presence of HSATII RNA within cells, we developed a cell culture system to test the effect of HSATII expression by establishing independent cell lines that stably express either HSATII or  $\alpha$ -sat satellite sequences from randomly integrated sites in the genome. We found that centromeric  $\alpha$ -sat and pericentromeric HSATII transcripts behave differently when ectopically expressed, yet the presence of HSATII and  $\alpha$ -sat transcripts both trigger cell division defects. HSATII RNA accumulates in nuclear bodies *in cis*, immediately adjacent to the integration site (Fig. 5), while  $\alpha$ -sat RNA is overall more diffuse and does not accumulate appreciably in nuclear bodies, suggesting that these two distinct tandemly repeated RNAs have different dynamics within the nuclear environment, and supporting that they may have distinct functions. Nuclear accumulations of HSATII RNA recruit MeCP2 into these nuclear bodies, which are reminiscent of CAST bodies in cancer cells (Hall et al., 2017) (Fig. 5). Although the transcripts localize in very different ways, expression of both  $\alpha$ -sat and HSATII induce cell division defects including chromatin bridges, blebbing, and micronuclei formation, thus have the potential to impact cell division and cause further instability.

Previous research has focused on the effect of expression of alpha satellite RNA, which is normally expressed at low levels in all cell types and thought to be involved in centromere protein recruitment (Hall et al., 2017; Johnson et al., 2017; McNulty et al., 2017; Wong et al., 2007). Though a low level of  $\alpha$ -sat expression is normal and likely necessary for centromere protein recruitment, several groups have described that when  $\alpha$ -sat is overexpressed (Chan et al., 2017; Ichida et al., 2018; Zhu et al., 2018) or knocked down (Ideue et al., 2014; Rošić et al., 2014), this has detrimental effects on cell division, suggesting that the levels of  $\alpha$ -sat expression are critical to maintaining proper cell division. In contrast, HSATII RNA is not expressed in normal cells, but is highly overexpressed in cancer cells and tissues (Hall et al., 2017; Ting et al., 2011). HSATII RNA accumulates in large nuclear foci within these cells, where it recruits nuclear regulatory proteins including MeCP2. Thus, previous evidence suggested that  $\alpha$ -sat and HSATII satellite sequences are likely to have unique regulatory mechanisms governing their expression. Our results here suggest that in addition to distinct transcriptional regulatory mechanisms, the transcripts themselves behave in unique ways within the nuclear environment. The accumulation of HSATII RNA and recruitment of MeCP2 is reminiscent of “toxic repeat RNAs,” which function to sequester nuclear regulatory proteins in diseases such as frontotemporal dementia/ALS and myotonic dystrophy (Swinnen et al., 2020). Further work will be required to understand the molecular interactions between HSATII RNA and MeCP2 and the cellular conditions required for MeCP2 recruitment to HSATII RNA foci,

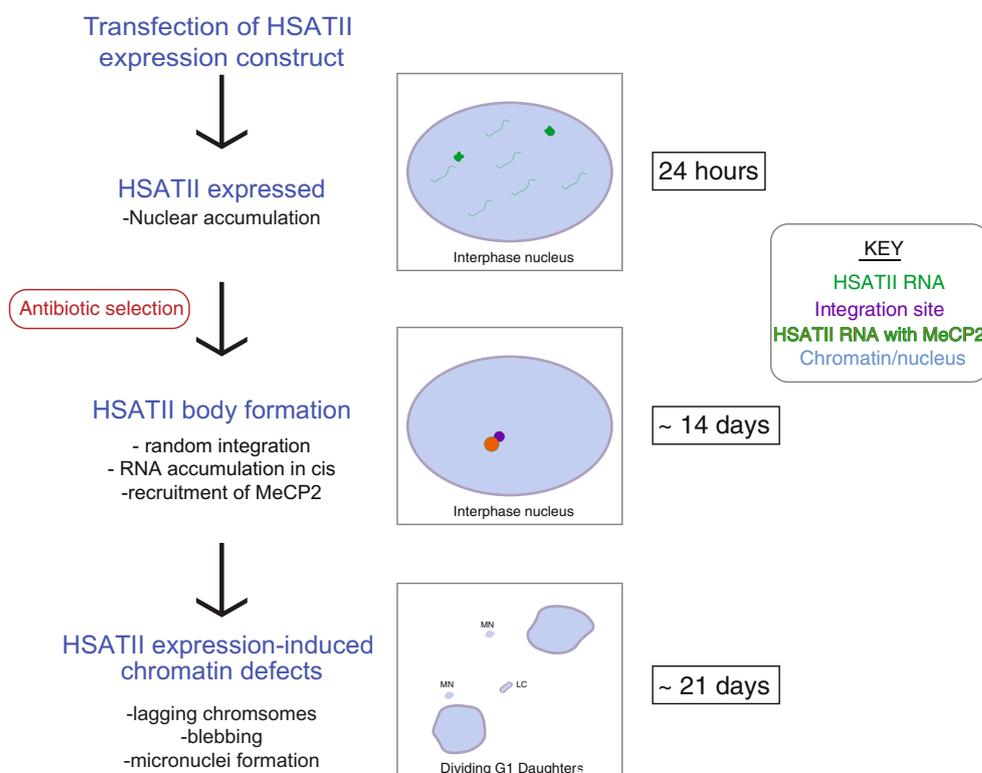


**Fig. 4** Ectopic expression of satellite RNA leads to cell division defects. (a) Representative images of HeLa cells transfected and stably expressing HSATII RNA for 3 weeks exhibiting cell division defects including chromatin bridges (CB), blebbing (Bleb), and the formation of micronuclei (MN), as demonstrated in the images in the top 3 panels (DAPI stained, greyscale). Arrows indicate the type of defect observed. Graphs depict percentages of total cells counted displaying these defects: chromatin bridges (bottom left), blebbing/MN (bottom middle), and total chromatin defects (bottom right) out of 500 total cells for each transfection vector used or control (untransfected) cells. (b) Representative images of Tig-1 normal fibroblast cells transfected and stably expressing HSATII RNA for 3 weeks displaying cell division defects (lagging chromosomes (LC), blebbing, MN) as indicated by arrows. Graphs depict percentages of cells displaying these defects: chromatin bridges (bottom left), blebbing/MN (bottom middle), and total chromatin defects (bottom right) out of 500 total cells for each transfection vector used (X axis). Frequencies of chromatin defects significantly ( $p < .05$ ) different from empty vector transfected cells by Chi-square test are indicated with two asterisks (\*\*). Frequencies of chromatin defects significantly ( $p < .01$ ) different from empty vector transfected cells by Chi-square test are indicated with three asterisks (\*\*\*). Untransfected control cells included for comparison. Scale bar, 5  $\mu$ m

but our results suggest that the recruitment of MeCP2 is likely to be a sequence-dependent mechanism, rather than a location-dependent mechanism since HSATII is expressed here from random integration sites in both HeLa and Tig-1 primary fibroblasts (Figs. S3 and S4).

**Fig. 5** Timeline of ectopic HSATII expression. HSATII RNA (green) is expressed transiently 24 h after transfection in multiple nuclear accumulations and some diffuse nuclear RNA. Following antibiotic selection for 2 weeks (14 days), HSATII RNA bodies form, which recruit MeCP2 (orange), and accumulate adjacent to the genomic integration site of the HSATII expression vector (purple). At this point, HSATII RNA is restricted to these nuclear bodies and little diffuse nuclear RNA is observed. Approximately 3 weeks after transfection (21 days), significant cell division defects are observed, including lagging chromosomes (LC) between dividing daughter cells and the formation of micronuclei (MN)

### Timeline of ectopic HSATII expression



HSATIII RNA, which is transcribed from Chr 9q12 during stress, assembles into nuclear stress bodies (nSBs), recruiting HSF1 and numerous splicing factors to promote intron retention and suppress splicing of mRNAs during recovery from heat shock (Biamonti and Vourc'h, 2010; Jolly et al., 2004; Ninomiya et al., 2020). Intriguingly, HSATII and HSATIII both derive from CATTC pentamer repeats, despite their divergence and distinct chromosomal localizations. HSATII displays a widespread distribution within the pericentric regions of ~11 human chromosomes, while the bulk of HSATIII resides on Chr 9 and Y, with some smaller loci interspersed within pericentric regions of additional chromosomes (Altemose et al., 2014; Tagarro et al., 1994). HSATII and HSATIII RNA appear to recruit different protein-binding partners, yet they both display conserved functional aspects in their ability to accumulate *in cis* and recruit/sequester protein binding partners. It is possible that the tandemly repetitive nature of pentamers embedded within HSATII/HSATIII repeats may facilitate this recruitment given that identical sequences (monomers) are present in high copy number within a linear array. Nuclear stress bodies form on HSATIII non-coding RNA and recruit proteins into nuclear condensates, a theme which is emerging as a conserved feature of RNP granules that are likely to exist within liquid-liquid phase separated domains in the nucleus. Recent work highlights the role of

RNA in modulating the biophysical properties of liquid droplets, and the accumulation of misregulated transcripts may lead to mislocalization of RNA-binding proteins containing intrinsically disordered domains (IDRs) (Maharana et al., 2018). MeCP2, though classically described as a DNA-binding protein, has been more recently identified in an RNA-binding protein (RBP) screen (Castello et al., 2016) and binds mRNAs in the brain, where it affects alternative splicing (Young et al., 2005). MeCP2 also contains an IDR, and is bound to HSATII RNA within CAST bodies, which we note also resemble nuclear condensates.

Since initially found to be expressed in cancer cells, the functional consequence of HSATII transcripts within the nucleus has remained elusive; however, induction and injection of HSATII RNA were found to result in the formation of cDNA intermediates, which can facilitate the expansion of pericentromeric HSATII-containing DNA at endogenous sites in the genome (Bersani et al., 2015). Though the mechanism facilitating this copy number gain is not well understood, it is thought to result from the formation of RNA-DNA hybrids following reverse transcription of HSATII RNA transcripts. In addition to cancer cells, HSATII RNA has recently been found to be induced upon herpesvirus infection, where it is thought to be involved in the viral life cycle (Nogalski et al., 2019) and in human FSHD cell models where it resides in nuclear dsRNA foci that also aggregate proteins (Shadle et al., 2020). HSATII RNA has also been implicated in the growth of cancer cells (Ting et al., 2011), and its expression is further induced in cells grown under 3D culturing conditions (Bersani et al., 2015). Thus, there may be convergent mechanisms by which HSATII RNA can more globally affect both growth and viral replication. Though these effects have been observed, the specific role that HSATII transcripts play in mediating these phenotypes has remained unclear. Our observation that HSATII RNA accumulates and recruits regulatory protein *in cis*, irrespective of the site from which it is expressed, may shed some significant insight into the molecular function of HSATII RNA. The ability of HSATII RNA to bind, and potentially sequester, large amounts of regulatory proteins is likely to have a large effect on transcription and genome regulation more broadly, and may ultimately lead to widespread heterochromatin instability, which is observed in cancer and other diseases (Carone and Lawrence, 2013).

Supporting a more global effect of HSATII accumulations within the nucleus, our results implicate HSATII transcripts in generating cell division defects. When HSATII RNA is expressed ectopically in both HeLa and Tig-1 primary fibroblasts for several weeks, we observed increased frequencies of chromatin bridges, lagging chromosomes, and micronuclei formation (Fig. 4 and Fig. S6). These phenotypes have previously been observed upon introduction of satellite transcripts into both mouse and human cells, where they may induce DNA damage via the formation of RNA-DNA hybrids (Zhu

et al., 2018; Zhu et al., 2011). Overexpression of  $\alpha$ -sat has also been linked to chromosomal instability (CIN) and segregation errors resulting in copy number changes in daughter cells (Ichida et al., 2018). Yet, paradoxically, other studies found no cell division phenotypes associated with overexpression of pericentric satellite sequences (Ideue et al., 2014; Rošić et al., 2014). It is important to note that the cell division defects we observe here are long-term effects of ectopic HSATII expression resulting 3–4 weeks after the initial transfection; thus, the cell division defects we observe here may reflect long-term effects of continual satellite overexpression. Furthermore, differences in the effects of satellite overexpression could be due to the location of the satellite DNA sequences in the genome (i.e., centric/pericentromeric) or the localization and function of the satellite RNA (i.e., *cis* vs *trans*). Results obtained here for HSATII RNA expression support previous reports of centromeric satellite expression-induced cell division defects (Bouzinba-Segard et al., 2006; Chan et al., 2017; Ichida et al., 2018; Slee et al., 2012), and extend this phenomenon to expression of pericentromeric HSATII RNA. While we did observe changes in chromosome number following long-term  $\alpha$ -satellite expression, we did not observe changes in chromosome number in HSATII-expressing cells (Fig. S7); thus, our evidence does not suggest that long-term expression of HSATII RNA leads to CIN.

Despite their abundant presence within pericentromeres and evidence suggesting that low-level expression of satellite RNA may be integral to both centromere protein recruitment and cell division (Johnson et al., 2017; McNulty et al., 2017), the role of tandemly repeated satellite sequences within these regions has been elusive. Recent work suggests that pericentric satellites within fruit fly and mouse heterologous chromosomes may act to tether chromosomes within nuclei via the formation of chromocenters (Jagannathan et al., 2018). Perturbation of proteins within chromocenters results in the formation of micronuclei due to budding from the nucleus, supporting that tethering of pericentric regions to chromocenters is key to nuclear organization and overall nuclear integrity (Jagannathan et al., 2018). Alteration of epigenetic regulatory mechanisms within heterochromatin is also known to cause chromosomal instability (Slee et al., 2012); thus, impaired maintenance of heterochromatin is likely to lead to both satellite expression and chromosomal instability. Our results here add to a growing body of evidence suggesting that the presence of aberrant satellite transcripts, themselves, are likely to induce defects in chromosome segregation. This observation is supported by previous studies of transient  $\alpha$ -sat expression, but the more long-term and specific effect of the presence of satellite RNA has been unclear. By randomly integrating and expressing ectopic HSATII and  $\alpha$ -sat RNA, we demonstrate the ability to observe the specific behavior of these satellite transcripts within the nucleus, irrespective of their site of transcription. Despite displaying strikingly different nuclear

## Materials and methods

Key Resource Table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	DyLight 488 Streptavidin	Vector Laboratories	Vector Laboratories:SA-5488	(1:500)
Antibody	DyLight 594 Streptavidin	Vector Laboratories	Vector Laboratories:SA-5594	(1:500)
Antibody	MeCP2 (D4F3) XP® Rabbit mAb	Cell Signaling Technology	Cell Signaling:3456	(1:250)
Antibody	Goat anti-Mouse IgG (H+L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 594	Invitrogen (Thermo Fisher Scientific)	Thermo Fisher:A-11005	(1:500) or (1:250)
Cell line ( <i>H. sapiens</i> )	HeLa			
Cell line ( <i>H. sapiens</i> )	Tig-1	Coriell	Coriell:AG06173	
Chemical compound, drug	Geneticin™ Selective Antibiotic (G418 Sulfate)	Gibco (Thermo Fisher Scientific)	Thermo Fisher:10131027	
Commercial assay or kit	StrataClone PCR Cloning Kit	Agilent	Agilent:240205	
Commercial assay or kit	pTarget™ Mammalian Expression Vector System	Promega	Promega:A1410	
Commercial assay or kit	iScript™ Reverse Transcription Supermix for RT-qPCR	Bio-Rad	Bio-Rad:1708841	
Commercial assay or kit	Lipofectamine® LTX with PLUS™ Reagent	Invitrogen (Thermo Fisher Scientific)	Thermo Fisher:15338030	
Commercial assay or kit	FuGENE® HD Transfection Reagent	Promega	Promega:E2311	
Sequence-based reagent	Alpha_sat_Chr4_F (used for cloning ASAT)	Integrated DNA Technologies		5'-CTGCACTACCTGAAGAGGAC-3'
Sequence-based reagent	Alpha_sat_Chr4_R (used for cloning ASAT)	Integrated DNA Technologies		5'-GATGGTTCAACTCTTACA-3'
Sequence-based reagent	HSATII_F (used for cloning and qRT-PCR of HSATII)	Integrated DNA Technologies		5'-GGAACCGAATGAATCCTCATTGAATG-3'
Sequence-based reagent	HSATII_R (used for cloning and qRT-PCR of HSATII)	Integrated DNA Technologies		5'-ATTCGATTCCATTGATGATGATTCC-3'
sequence-based reagent	asat_FITC_DNA_probe	Integrated DNA Technologies		5'-/56-FAM/CCTTTTGATAGAGCAGTTTT GAAACACTCTTTTTGTAGAATCTG CAAGTGGATATTTGG-3'
Sequence-based reagent	HSATII_bio_LNA_probe	Qiagen	GG #LCD0156269-BKH, product 339500	5'biotin-ATTCCATTCCAGATTCCATTC GATC-3'
Sequence-based reagent	bACT F QRT (used for qRT-PCR)	Integrated DNA Technologies		5'-CATGTACGTTGCTATCCAGGC-3'
Sequence-based reagent	bACT R QRT (used for qRT-PCR)	Integrated DNA Technologies		5'-CTCCTTAATGTCACGCACGAT-3'
Transfected construct ( <i>H. sapiens</i> )	GFP construct	Other		Gift from P. Jones laboratory
Transfected construct ( <i>H. sapiens</i> )	pTarget-ASAT	This paper		
Transfected construct ( <i>H. sapiens</i> )	pTarget-HSATII	This paper		

localization, the presence of both  $\alpha$ -sat and HSATII RNA leads to chromatin bridge formation and the formation of micronuclei, suggesting convergent roles for these transcripts in the perturbation of cell division.

## Construction of DNA plasmids

The pTarget™ Mammalian Expression Vector System (Promega, Madison, WI), containing a CMV promoter and neomycin resistance, was used for transfection and long-term selection of stably transfected lines. Insert DNA was derived from total RNA extracted from U2OS cells and then reverse transcribed using an iScript™ Reverse Transcription Supermix (Bio-Rad Laboratories, Hercules, CA). A 140-bp region of alpha satellite sequence from chromosome 4 (GenBank M38467.1) and a 349-bp region of HSATII from chromosome 7 were independently cloned into pTarget™ DNA backbone via StrataClone TA cloning (Agilent Technologies, Santa Clara, CA).

## Cell culture and transfection

HeLa human epithelioid cervix carcinoma cells and Tig-1 human primary fibroblast cells (Coriell AG06173) were cultured in the presence of 5% CO<sub>2</sub> prior to transfection in Minimal Essential Medium (Gibco, Thermo Fisher Scientific, Waltham, MA) supplemented with 10% (HeLa) or 15% (Tig-1) fetal bovine serum (HyClone, GE Healthcare Life Sciences, Marlborough, MA; Avantor Seradigm, VWR, Radnor, PA), 100 units/mL penicillin and 100  $\mu$ g/mL streptomycin (Gibco, Thermo Fisher Scientific) and 2 mM L-glutamine (Gibco, Thermo Fisher Scientific). After transfection, cells were maintained in the above media, replacing the penicillin/streptomycin with geneticin (G418 sulfate) (Gibco, Thermo Fisher Scientific) at a concentration of 700  $\mu$ g/mL for HeLa cells and 500  $\mu$ g/mL for Tig-1 cells for selection over a minimum of 14 days.

Transient transfections of the empty pTarget™ vector, pTarget- $\alpha$ SAT, and pTarget-HSATII were performed in parallel in HeLa cells using Lipofectamine® LTX with PLUS™ Reagent (Invitrogen, Thermo Fisher Scientific). Cells were seeded on 22×22mm coverslips or in T-25 flasks such that they were 70–90% confluent at the time of transfection. The plasmid DNA-lipid complexes for transfection were prepared following the manufacturer's protocol. For cells in 6-well plates, 2.5  $\mu$ g DNA was transfected per well using a 2.5  $\mu$ g:22.5  $\mu$ L ratio of plasmid DNA:Lipofectamine® LTX, supplemented with 3.5  $\mu$ L PLUS™ Reagent. Cells in T-25 flasks were transfected with 3  $\mu$ g plasmid DNA per flask, with

312.5  $\mu$ L Lipofectamine® LTX and supplemented with 12.5  $\mu$ L PLUS™ Reagent. Twenty-four hours following transfection, cells on coverslips were fixed for FISH experiments and cells in flasks were pelleted and frozen for RNA extraction and qRT-PCR analysis.

Stably transfected lines of HeLa cells were cultured in T-25 flasks, also using a ratio of 3  $\mu$ g plasmid DNA:315  $\mu$ L Lipofectamine® LTX and 12.5  $\mu$ L PLUS™ Reagent per flask. In parallel with the flasks transfected with the empty pTarget™ vector, pTarget- $\alpha$ SAT, and pTarget-HSATII plasmids, untransfected HeLa cells were cultured in parallel as a control. Cells were transfected 24 h after seeding in flasks, and then maintained in standard growth medium containing 10% FBS for 3 days before adding G418 sulfate-supplemented media. Cells were cultured in selective media for 2 weeks, during which the media was exchanged every 2 days. After this selection period, the stably transfected lines were maintained for an additional several weeks. Cells were then plated on coverslips and fixed for FISH experiments twice a week, and harvested 3 and 4 weeks following the end of selection, for RNA extraction and qRT-PCR analysis.

To optimize lipid-mediated transfection for Tig-1 primary fibroblasts for both expression levels as well as overall health of transfected cells, a GFP expression vector (obtained from P. Jones) was transiently transfected into cells on coverslips using a range of conditions, with 3  $\mu$ g plasmid DNA transfected per well of a 6-well plate. Lipofectamine® LTX was tested at volumes of 9  $\mu$ L and 12  $\mu$ L per well, and FuGENE® HD was tested at a range of concentrations (3:1, 4:1 and 5:1 per  $\mu$ g DNA). After 24 h, media was exchanged and cells maintained in standard media for an additional 24 h. Cells were then visualized in culture prior to fixation and DAPI staining as described below. Transfection efficiency was evaluated by scoring GFP expression and cellular health gauged by scoring the number of cells per field of view.

Tig-1 cells were stably transfected using FuGENE® HD Transfection Reagent (Promega), following the manufacturer's protocol. Cells in T-25 flasks were grown to 60% confluency and then transfected using 8  $\mu$ g plasmid DNA and 44  $\mu$ L FuGENE® HD Transfection Reagent per flask. Three control flasks were cultured in addition to the pTarget lines: a wild-type untransfected flask not treated with selective media, a wild-type untransfected flask selected with G418 sulfate, and a mock flask treated with the transfection reagent and then selected with G418 sulfate. Media containing transfection reagent was removed 24 h post-transfection and replaced with selective media. Selection was complete after 10–14 days, after which the stably transfected cells were maintained for several weeks. During this time, cells were fixed on coverslips weekly and harvested for RNA extraction at multiple time points.

## Preparation of chromosome spreads

Chromosome spreads were prepared as previously described (Howe et al., 2014), with the following modifications. Cells in flasks were treated with colcemid for 1–2 h depending on the cell line and monitored for mitotic cells, at which point the cells were harvested, treated with hypotonic solution (0.075M KCl), and fixed with Carnoy's fixative (3:1 ratio of methanol:glacial acetic acid). Fixed cells were then dropped onto clean glass slides under humid conditions to produce chromosome spreads. Slides were serially dehydrated with ethanol and stored at  $-20^{\circ}\text{C}$ . Thawed slides were dehydrated in 100% ethanol prior to DNA FISH experiments.

## Fluorescence in situ hybridization (FISH) and immunostaining

For all FISH and immunostaining on interphase cells, transfected cell lines were grown on  $22\times 22$  mm coverslips and then fixed as follows: 3–5 min in cytoskeletal buffer (Byron et al., 2013) plus 0.5% Triton X-100 (Sigma-Aldrich, St. Louis, MO) and 10 mM ribonucleoside vanadyl complex (New England Biolabs, Ipswich, MA), 10 min in 4% paraformaldehyde (Ted Pella, Redding, CA) in  $1\times$  PBS, followed by storage in 70% ethanol at  $4^{\circ}\text{C}$ .

RNA FISH was performed using a FITC-labeled DNA oligo probe complementary to alpha satellite (Integrated DNA Technologies, Newark, NJ) or a biotinylated locked nucleic acid (LNA) oligo probe complementary to HSATII (QIAGEN, Hilden, Germany) (see Key Resources Table). For each coverslip, 0.25 pmol of oligonucleotide was denatured for 10 min at  $80^{\circ}\text{C}$  in 30% formamide, and then diluted in hybridization buffer to a final concentration of 15% formamide. Hybridization buffer contained 2 mg/mL bovine serum albumin (BSA) (Roche, Basel, Switzerland), 10% dextran sulfate, and  $2\times$  SSC, supplemented with RNasin Plus RNase Inhibitor (Promega). The probe was then applied to coverslips, incubated overnight in a humid chamber at  $37^{\circ}\text{C}$  and then washed for 15 min in 15% formamide in  $2\times$  SSC at  $37^{\circ}\text{C}$ , 15 min in  $2\times$  SSC at  $37^{\circ}\text{C}$ , 15 min in  $1\times$  SSC at RT, and 5 min in  $4\times$  SSC at RT. HSATII RNA was detected with either 1:500 DyLight 488 streptavidin or DyLight 594 streptavidin (Vector Laboratories, Burlingame, CA) in  $4\times$  SSC+1% BSA. Secondary-detected coverslips were then washed at room temperature for 10 min in  $4\times$  SSC, 10 min in  $4\times$  SSC+0.1% Triton X-100, and 10 min in  $4\times$  SSC. All coverslips were stained with 2  $\mu\text{g}/\text{mL}$  DAPI, mounted with VECTASHIELD® Antifade Mounting Medium (Vector Laboratories), and sealed with fingernail polish.

A pTarget™ vector was labeled by nick translation with digoxigenin-11-dUTP (Roche) (Byron et al., 2013) to detect integration of the construct via DNA FISH in both interphase cells and mitotic chromosome spreads. To prepare the probe

for hybridization, 50 ng of nick translated probe per coverslip was combined with 12  $\mu\text{g}$  of human Cot-1 DNA (Roche), 10  $\mu\text{g}$  salmon sperm ssDNA (Sigma-Aldrich), and 20  $\mu\text{g}$  *E. coli* tRNA (Sigma-Aldrich), dried with a Speed Vac, and then resuspended in 10  $\mu\text{L}$  formamide. Coverslips were first treated with 0.2 N NaOH in 70% ethanol to remove RNA, and then dehydrated in 100% ethanol and air dried. Interphase cells or chromosome spreads were denatured in 70% formamide in  $2\times$  SSC, pH 7.0, for 2 min at  $80^{\circ}\text{C}$ , followed by 5 min in cold ( $4^{\circ}\text{C}$ ) 70% ethanol and 5 min in cold 100% ethanol prior to air drying. Probes were diluted in the hybridization buffer used for RNA FISH, to a final concentration of 50% formamide, before applying to coverslips and incubating overnight at  $37^{\circ}\text{C}$  in a humid chamber. Coverslips were then washed as for RNA FISH, with the adjustment of the first wash to 50% formamide/ $2\times$  SSC. The probe was detected with anti-digoxigenin-fluorescein (Roche), diluted 1:500 in  $4\times$  SSC+1% BSA. Secondary incubation, washes, DAPI staining and mounting were carried out as for the RNA oligo hybridization. RNA and DNA co-hybridization was performed by first completing the RNA hybridization as described, followed by fixation in 4% paraformaldehyde for 10 min prior to DNA FISH, with the removal of the NaOH treatment step.

In HeLa cells, MeCP2 and HSATII RNA co-visualization was performed by first hybridizing oligo probes to RNA as described, followed by a 10-minute fixation in 4% paraformaldehyde prior to staining for MeCP2. For Tig-1 cells, MeCP2 was stained first, and then the signal was fixed in 4% paraformaldehyde for 10 min before proceeding with the RNA FISH protocol. The HSATII probe was detected with DyLight 488 streptavidin and the MeCP2 with an Alexa Fluor 594 goat anti-rabbit secondary. To stain for MeCP2, coverslips were rinsed for 10 min in  $1\times$  PBS, then incubated with a 1:250 dilution of MeCP2 rabbit antibody (Cell Signaling Technology, Danvers, MA) in  $1\times$  PBS+1% BSA at  $37^{\circ}\text{C}$  for 1–3 h. Coverslips were washed at room temperature: 10 min in  $1\times$  PBS, 10 min in  $1\times$  PBS+0.1% Triton X-100, and 10 min in  $1\times$  PBS. The secondary antibody (diluted in  $1\times$  PBS+1% BSA, (1:500) for HeLa cells, and (1:250) for Tig-1 cells) was then applied and incubated as for the primary, with the same set of washes. Following both MeCP2 staining and HSATII RNA hybridization, coverslips were DAPI stained and mounted.

## Image acquisition and analysis

Slides were imaged using a ZEISS Axio Observer Z1 epifluorescent microscope equipped with a Hamamatsu ORCA-Flash 4.0 Digital CMOS camera or ZEISS Axiocam 702 mono camera with a  $\times 100$  oil objective. Both single plane images and Z-stacks were captured and analyzed using ZEISS ZEN2 imaging software.

Counts of nuclear body formation or chromatin defects were scored manually, while colocalization was measured using the ZEN2 Colocalization Module, including generation of Pearson's correlation coefficient. Linescans were generated using the Profile tool in ZEN2. DNA integration foci from pTarget DNA hybridizations on interphase cells were scored on a subjective scale for relative intensity as determined by observation through the microscope oculars. This scale categorized relative foci brightness into the following five subjective classes: 0 (no foci), + (dim), ++ (easily visible), +++ (very bright), or ++++ (extremely bright). DNA integration foci from chromosome spreads hybridized with the pTarget probe were categorized as being on acrocentric or submetacentric chromosomes, and further scored based on signal location relative to the centromere as well as chromosome size. Chromatin defects were scored by imaging DAPI stained slides and evaluating nuclei for abnormal phenotypes. This scoring was completed by multiple individuals in a blind manner (with the labels removed from slides) to ensure robustness and reproducibility in data scoring. Categories for these defects included the following: chromatin bridges between two nuclei (including lagging chromosomes), micronuclei which were fully separated from the primary nucleus, abnormally shaped nuclei (including nuclei in the process of blebbing), and other defects (including nuclei with burst chromatin and holes or indentations in the chromatin). All scoring data were analyzed and graphed using Prism8 (GraphPad), including statistical analyses using Chi-square test (as shown in Figs. 1 and 4, and Figure S6).

### Quantitative RT-PCR analysis

RNA was extracted from pelleted cells using either an RNeasy Mini kit and QIAshredder (QIAGEN) or a Direct-zol RNA Miniprep Kit (Zymo Research, Irvine, CA), following manufacturers' protocols for each. Samples were then treated with TURBO™ DNase (Invitrogen, Thermo Fisher Scientific) following manufacturer's guidelines. RNA was purified using Agencourt AMPure XP beads (Beckman Coulter, Brea, CA) and quantified via NanoDrop™ (Thermo Fisher Scientific). Reverse transcription was performed using iScript™ Reverse Transcription Supermix for RT-qPCR (Bio-Rad), including RT-controls for all RNA samples. Quantitative RT-PCR was carried out with PerfeCTa® SYBR® Green FastMix® Reaction Mix (Quantabio, Beverly, MA), using primers included in the Key Resources Table. All reactions were performed in triplicate, using a CFX96 Real-Time C1000 Thermal Cycler (Bio-Rad) or MJ Research PTC-200 thermocycler equipped with a Chromo4 Real-Time PCR Detector (Bio-Rad). Analysis of relative enrichment was calculated using the  $\Delta\Delta C(t)$  method (Pfaffl, 2001), where reported values are normalized to  $\beta$ -actin expression.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00412-021-00753-0>.

**Acknowledgements** We would like to thank S. Akkipeddi for assistance with maintenance of stably transfected cell lines and H. Sen for assistance with chromosome spreads and DNA hybridizations. We are grateful to J. Rubien, S. Akkipeddi, and B. Carone for helpful comments and critical reading of the manuscript.

**Author's contribution** All authors were involved in the design, implementation, data collection, and data interpretation. DMC devised the study and wrote the manuscript. All authors read and approved the final manuscript.

**Funding** This work was supported by funding from the Charles E. Kaufman Foundation of the Pittsburgh Foundation (KA2017-91790) and the National Institutes of Health (R15 GM134495) to DMC.

### Declarations

**Competing interests** The authors declare no competing interests.

### References

- Altemose N, Miga KH, Maggioni M, Willard HF (2014) Genomic characterization of large heterochromatic gaps in the human genome assembly. *PLoS Comput Biol* 10:e1003628
- Bersani F, Lee E, Kharchenko PV, Xu AW, Liu M, Xega K, MacKenzie OC, Brannigan BW, Wittner BS, Jung H, Ramaswamy S, Park PJ, Maheswaran S, Ting DT, Haber DA (2015) Pericentromeric satellite repeat expansions through RNA-derived DNA intermediates in cancer. *Proc Natl Acad Sci* 112:15148–15153
- Biamonti G, Vourc'h C (2010) Nuclear stress bodies. *Cold Spring Harb Perspect Biol* 2:a000695
- 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 U S A* 103:8709–8714
- Brückmann NH, Pedersen CB, Ditzel HJ, Gjerstorff MF (2018) Epigenetic Reprogramming of Pericentromeric Satellite DNA in Premalignant and Malignant Lesions. *Mol Cancer Res* 16:417–427
- Byron M, Hall LL, Lawrence JB (2013) A multifaceted FISH approach to study endogenous RNAs and DNAs in native nuclear and cell structures. *Curr Protoc Hum Genet* Chapter 4, Unit 4.15
- Carone DM, Lawrence JB (2013) Heterochromatin instability in cancer: from the Barr body to satellites and the nuclear periphery. *Semin Cancer Biol* 23:99–108
- Castello A, Fischer B, Frese CK, Horos R, Alleaume A-M, Foehr S, Curk T, Krijgsveld J, Hentze MW (2016) Comprehensive identification of RNA-binding domains in human cells. *Mol Cell* 63:696–710
- Chan DY, Moralli D, Khoja S, Monaco ZL (2017) Noncoding centromeric RNA expression impairs chromosome stability in human and murine stem cells. *Dis Markers* 2017:7506976
- Eichler EE, Clark RA, She X (2004) An assessment of the sequence gaps: unfinished business in a finished human genome. *Nat Rev Genet* 5: 345–354
- Hall LE, Mitchell SE, O'Neill RJ (2012) Pericentric and centromeric transcription: a perfect balance required. *Chromosom Res* 20:535–546
- Hall LL, Byron M, Carone DM, Whitfield TW, Pouliot GP, Fischer A, Jones P, Lawrence JB (2017) Demethylated HSATII DNA and HSATII RNA foci sequester PRC1 and MeCP2 into cancer-specific nuclear bodies. *Cell Rep* 18:2943–2956

- Hédouin S, Grillo G, Ivkovic I, Velasco G, Francastel C (2017) CENP-A chromatin disassembly in stressed and senescent murine cells. *Sci Rep* 7:1–14
- Henikoff S, Ahmad K, Malik HS (2001) The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* 293:1098–1102
- Hite KC, Adams VH, Hansen JC (2009) Recent advances in McCP2 structure and function. *Biochem Cell Biol* 87:219–227
- Howe B, Umrigar A, Tsien F (2014) Chromosome preparation from cultured cells. In: *Chromosome preparation from cultured cells*. In *J Vis Exp*
- Ichida K, Suzuki K, Fukui T, Takayama Y, Kakizawa N, Watanabe F, Ishikawa H, Muto Y, Kato T, Saito M, Futsuhara K, Miyakura Y, Noda H, Ohmori T, Konishi F, Rikiyama T (2018) Overexpression of satellite alpha transcripts leads to chromosomal instability via segregation errors at specific chromosomes. *Int J Oncol*
- 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
- Jagannathan M, Cummings R, Yamashita YM (2018) A conserved function for pericentromeric satellite DNA
- Jain M, Olsen HE, Tumer DJ, Stoddart D, Bulazel KV, Paten B, Haussler D, Willard HF, Akeson M, Miga KH (2018) Linear assembly of a human centromere on the Y chromosome. *Nat Biotechnol* 36:321–323
- Johnson WL, Yewdell WT, Bell JC, McNulty SM, Duda Z, O'Neill RJ, Sullivan BA, Straight AF (2017) RNA-dependent stabilization of SUV39H1 at constitutive heterochromatin. *Elife* 6
- Jolly C, Metz A, Govin J, Vigneron M, Turner BM, Khochbin S, Vourc'h C (2004) Stress-induced transcription of satellite III repeats. *J Cell Biol* 164:25–33
- Maharana S, Wang J, Papadopoulos DK, Richter D, Pozniakovskiy A, Poser I, Bickle M, Rizk S, Guillen-Boixet J, Franzmann TM et al (2018) RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* 360:918–921
- McNulty SM, Sullivan LL, Sullivan BA (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
- Miga KH (2019) Centromeric satellite DNAs: hidden sequence variation in the human population. In *Genes (Basel)*, vol. 10
- Miga KH, Koren S, Rhie A, Vollger MR, Gershman A, Bzikadze A, Brooks S, Howe E, Porubsky D, Logsdon GA et al (2019) Telomere-to-telomere assembly of a complete human X chromosome
- Ninomiya K, Adachi S, Natsume T, Iwakiri J, Terai G, Asai K, Hirose T (2020) LncRNA-dependent nuclear stress bodies promote intron retention through SR protein phosphorylation. *EMBO J* 39:e102729
- Nogalski MT, Solovyov A, Kulkarni AS, Desai N, Oberstein A, Levine AJ, Ting DT, Shenk T, Greenbaum BD (2019) A tumor-specific endogenous repetitive element is induced by herpesviruses. *Nat Commun* 10:90
- Ohashi M, Aizawa S, Ooka H, Ohsawa T, Kaji K, Kondo H, Kobayashi T, Noumura T, Matsuo M, Mitsui Y, Murota S, Yamamoto K, Ito H, Shimada H, Utakoji T (1980) A new human diploid cell strain, TIG-1, for the research on cellular aging. *Exp Gerontol* 15:121–133
- Perea-Resa C, Blower MD (2018) Centromere Biology: transcription goes on stage
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:e45
- Richard GF, Kerrest A, Dujon B (2008) Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol Mol Biol Rev* 72:686–727
- Rošić S, Köhler F, Erhardt S (2014) Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J Cell Biol* 207:335–349
- Shadle SC, Bennett SR, Wong C-J, Karreman NA, Campbell AE, van der Maarel SM, Bass BL, Tapscott SJ (2020) DUX4-induced bidirectional HSATII satellite repeat transcripts form intranuclear double-stranded RNA foci in human cell models of FSHD. *Hum Mol Genet*
- Slee RB, Steiner CM, Herbert BS, Vance GH, Hickey RJ, Schwarz T, Christan S, Radovich M, Schneider BP, Schindelhauer D, Grimes BR (2012) Cancer-associated alteration of pericentromeric heterochromatin may contribute to chromosome instability. *Oncogene* 31:3244–3253
- Smurova K, De Wulf P (2018) Centromere and pericentromere transcription: roles and regulation ... in sickness and in health. *Front Genet* 9:674
- Swanson EC, Manning B, Zhang H, Lawrence JB (2013) Higher-order unfolding of satellite heterochromatin is a consistent and early event in cell senescence. *J Cell Biol* 203:929–942
- Swinnen B, Robberecht W, Van Den Bosch L (2020) RNA toxicity in non-coding repeat expansion disorders. *EMBO J* 39:e101112
- Tagarro I, Fernández-Peralta AM, González-Aguilera JJ (1994) Chromosomal localization of human satellites 2 and 3 by a FISH method using oligonucleotides as probes. *Hum Genet* 93:383–388
- Tasselli L, Xi Y, Zheng W, Tennen RI, Odrowaz Z, Simeoni F, Li W, Chua KF (2016) SIRT6 deacetylates H3K18ac at pericentric chromatin to prevent mitotic errors and cellular senescence. *Nat Struct Mol Biol* 23:434–440
- 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
- Voullaire LE, Slater HR, Petrovic V, Choo KH (1993) A functional marker centromere with no detectable alpha-satellite, satellite III, or CENP-B protein: activation of a latent centromere? *Am J Hum Genet* 52:1153–1163
- Wong LH, Brettingham-Moore KH, Chan L, Quach JM, Anderson MA, Northrop EL, Hannan R, Saffery R, Shaw ML, Williams E, Choo KHA (2007) Centromere RNA is a key component for the assembly of nucleoproteins at the nucleolus and centromere. *Genome Res* 17:1146–1160
- Young JI, Hong EP, Castle JC, Crespo-Barreto J, Bowman AB, Rose MF, Kang D, Richman R, Johnson JM, Berget S, Zoghbi HY (2005) Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc Natl Acad Sci U S A* 102:17551–17558
- Zhu Q, Pao GM, Huynh AM, Suh H, Tonnu N, Nederlof PM, Gage FH, Verma IM (2011) BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. *Nature* 477:179–184
- Zhu Q, Hoong N, Aslanian A, Hara T, Benner C, Heinz S, Miga KH, Ke E, Verma S, Soroczynski J et al (2018) Heterochromatin-encoded satellite RNAs induce breast cancer. *Mol Cell* 70:842–853.e7
- Zink D, Fischer AH, Nickerson JA (2019) Nuclear structure in cancer cells. *Nat Rev Cancer* 4:677–687

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.