



# CCAT1 lncRNA is chromatin-retained and post-transcriptionally spliced

Chaya Bohrer<sup>1</sup> · Eli Varon<sup>1</sup> · Eldar Peretz<sup>1</sup> · Gita Reinitz<sup>1</sup> · Noa Kinor<sup>1</sup> · David Halle<sup>2</sup> · Aviram Nissan<sup>3,4</sup> · Yaron Shav-Tal<sup>1</sup>

Accepted: 22 April 2024 / Published online: 19 May 2024  
© The Author(s) 2024

## Abstract

Super-enhancers are unique gene expression regulators widely involved in cancer development. Spread over large DNA segments, they tend to be found next to oncogenes. The super-enhancer *c-MYC* locus forms long-range chromatin looping with nearby genes, which brings the enhancer and the genes into proximity, to promote gene activation. The colon cancer-associated transcript 1 (*CCAT1*) gene, which is part of the *MYC* locus, transcribes a lncRNA that is overexpressed in colon cancer cells through activation by *MYC*. Comparing different types of cancer cell lines using RNA fluorescence in situ hybridization (RNA FISH), we detected very prominent *CCAT1* expression in HeLa cells, observed as several large *CCAT1* nuclear foci. We found that dozens of *CCAT1* transcripts accumulate on the gene locus, in addition to active transcription occurring from the gene. The accumulating transcripts are released from the chromatin during cell division. Examination of *CCAT1* lncRNA expression patterns on the single-RNA level showed that unspliced *CCAT1* transcripts are released from the gene into the nucleoplasm. Most of these unspliced transcripts were observed in proximity to the active gene but were not associated with nuclear speckles in which unspliced RNAs usually accumulate. At larger distances from the gene, the *CCAT1* transcripts appeared spliced, implying that most *CCAT1* transcripts undergo post-transcriptional splicing in the zone of the active gene. Finally, we show that unspliced *CCAT1* transcripts can be detected in the cytoplasm during splicing inhibition, which suggests that there are several *CCAT1* variants, spliced and unspliced, that the cell can recognize as suitable for export.

**Keywords** *MYC* · *CCAT1* · *PVT1* · Transcription site · Post-transcriptional splicing · RNA FISH

## Introduction

The regulation of gene expression in eukaryotes is complex and includes different levels of regulation including epigenetic, transcriptional, translational, and post-translational control (Willemin et al. 2024; Wang et al. 2023). Enhancers are non-coding regulatory elements that enhance the transcription of associated genes when bound by specific transcription factors and are not necessarily in proximity to the gene body. Rather, they can be found upstream, downstream,

or within the coding region and are brought into the vicinity of the promoter through genome folding (Deng et al. 2012). Super-enhancers (SE) are a class of regulatory sequences with enrichment for the binding of activators of transcription factors within the sequence (Pott and Lieb 2015). SEs tend to spread over large DNA regions and are found to be enriched next to genes with known oncogenic functions (Loven et al. 2013; Hnisz et al. 2013; Jia et al. 2020).

SEs can drive the expression of oncogenic long non-coding RNAs (lncRNAs) with tumor-promoting functions. For instance, the lncRNA urothelial cancer-associated 1 (*UCA1*), which is driven by a SE, is overexpressed in ovarian cancer and leads to tumorigenesis through YAP activation (Lin et al. 2019). *LINC01503* lncRNA is located at a SE, and the binding of the transcription factor TP63 to the SE locus activates the expression of *LINC01503*. This activation leads to squamous cell carcinoma development, and the lncRNA levels are correlated with the shorter survival time of patients (Xie et al. 2018). The lncRNA *LIMD1-AS1*, which is also activated by a SE, is upregulated in glioma

✉ Yaron Shav-Tal  
Yaron.Shav-Tal@biu.ac.il

<sup>1</sup> The Mina and Everard Goodman Faculty of Life Sciences and Institute of Nanotechnology, Bar-Ilan University, Ramat Gan, Israel

<sup>2</sup> Biochemistry Laboratory, Samson Assuta Ashdod University Hospital, Ashdod, Israel

<sup>3</sup> Ziv Medical Center, Safed, Israel

<sup>4</sup> Surgical Innovation Laboratory, The Chaim Sheba Medical Center, Tel Hashomer, Ramat Gan, Israel

through its activation of CDK7, which contributes to cell proliferation (Chen et al. 2023).

A well-known SE that can express various lncRNAs in different types of cancers is the SE at the *MYC* locus (Pott and Lieb 2015; Xiang et al. 2014; Iaccarino 2017). The *c-MYC* (*MYC*) oncogene is known to be upregulated in 50–60% of all tumors, while its overexpression can be achieved by a variety of mechanisms. The *MYC* gene is located at the 8q24 locus, an area containing different enhancers that are known to be involved in diverse types of cancers. The enhancers are organized in topologically associating domains (TADs), while the binding of CTCF and cohesion proteins to the TADs contributes to their stability (Pombo and Dillon 2015; Dixon et al. 2016; Dekker et al. 2023). Several studies have identified that the long-range chromatin looping at the *MYC* locus, which leads to *MYC* hyperactivation, plays a critical role in cancer progression (Huppi et al. 2012; Lancho and Herranz 2018). A major factor in *MYC* upregulation that leads to cancer development is the binding of *MYC* to active binding sites on nearby enhancers, which results in high transcriptional activity (Lin et al. 2012; See et al. 2022). *MYC* binding to enhancers activates protein-coding genes and lncRNAs, which can indirectly regulate gene expression (Wang et al. 2020b). Those ncRNAs, which are known as super-enhancer-derived ncRNAs, play a critical role in tumorigenesis, metastasis, drug resistance, and more (Peng et al. 2019; Ge et al. 2019; Lee et al. 2020).

Two of the many lncRNAs expressed from the *MYC* locus are *PVT1* and *CCATI*. Plasmacytoma variant translocation 1 (*PVT1*) lncRNA is located on chromosome 8q24.21 and 53 kb downstream of the *MYC* locus (Parolia et al. 2018). *PVT1* is overexpressed in various types of cancer (Liu et al. 2015; Kong et al. 2015; Zhang et al. 2018; Li et al. 2024). Through its interaction with *MYC*, *PVT1* promotes progression, invasion, metastasis, and chemoradiotherapy resistance in different tumors (Shigeyasu et al. 2020; Ansari et al. 2019). Colon cancer-associated transcript 1 (*CCATI*), also known as *CCATI-S* or as cancer-associated region long non-coding RNA-5 (*CARLo-5*), was first discovered to be overexpressed in patients with colorectal cancer (CRC), while the highest expression was observed in the small intestine and esophagus (Nissan et al. 2012). *CCATI* maps to chromosome 8q24.2 and contains 2795 nucleotides. The transcript has two exons and an intron that is spliced. *CCATI* has two isoforms: *CCATI-L* with an extended second exon, and *5L-CCATI* with an extended 5'UTR. The dysregulation of *CCATI* expression affects tumorigenesis and clinical manifestations such as tumor size, metastasis, invasion, and patient survival (Wang et al. 2019; Zhan and Xian 2023).

One of the reasons for *CCATI* upregulation is its location at the SE and its interaction with *MYC*. A DNA loop between the *MYC* promoter and the enhancer locus 335 kb upstream of *MYC* (*MYC-335*) has been demonstrated (Ahmadiyeh et al. 2010). More specifically, the *CCATI-L* gene is located 515 kb upstream of *MYC*, and *CCATI-L* transcripts have a functional role in chromatin looping at the *MYC* locus. DNA fluorescence in situ hybridization (DNA FISH) experiments confirmed that the *MYC* gene and the *CCATI* gene are co-localized in the tissues of CRC patients (Xiang et al. 2014). Further studies have identified high expression of *CCATI* in breast cancer, lung cancer, osteosarcoma, and other types of cancers (Chen et al. 2016; Lai et al. 2018; Liu et al. 2019, 2023; White et al. 2014; Alaiyan et al. 2013). At present, the exact role of *CCATI* in cancer development is unknown.

The expression of the RNAs from the *MYC* locus has not been studied on the single-molecule RNA level. We found that the *CCATI* lncRNA is highly detectable in HeLa cells, particularly in comparison to other cell lines that are known to overexpress *CCATI* but have less detectable transcript levels. We, therefore, examined *CCATI* expression patterns in HeLa cells and found that the large observed foci of *CCATI* transcription were formed due to transcript accumulation at the gene locus. In addition, we found that *CCATI* transcripts are post-transcriptionally spliced, and this occurs shortly after their release from the gene. Interestingly, under splicing inhibition conditions, unspliced transcripts were also found in the cytoplasm, suggesting that these transcripts are a variant that the cell can recognize as suitable for export.

## Materials and methods

### Cell culture

HeLa and RKO cells were maintained in high-glucose Dulbecco's modified Eagle's medium (DMEM) (Biological Industries, Beit-Haemek, Israel) containing 10% fetal bovine serum (FBS; HyClone Laboratories, Logan, UT, USA), 100 IU/mL penicillin, and 100 µg/mL streptomycin (Biological Industries). Cells were grown at 37 °C and 5% CO<sub>2</sub>. HT-29 and HCT116 were maintained in McCoy's 5A Medium (Biological Industries) containing 10% FBS, 100 IU/mL penicillin, and 100 µg/mL streptomycin. For transcription inhibition, cells were grown on coverslips and incubated at 37 °C for 2, 3, and 4 h with either actinomycin D (ActD) (5 µg/mL, Sigma-Aldrich, Rehovot, Israel) or 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) (25 µg/mL, Sigma-Aldrich) before fixation for 20 min in 4% paraformaldehyde (PFA). For splicing inhibition, cells were grown on coverslips and incubated at 37 °C for 1 or 6 h with pladienolide B (PLB) (0.5 µM, Santa Cruz Biotechnology, Dallas, TX, USA) before fixation for 20 min in 4% PFA.

## Total RNA purification

Total RNA was produced using TRI Reagent (Sigma-Aldrich), and DNA was removed using the TURBO DNA-free kit (Invitrogen) according to the manufacturer's instructions. Synthesis of complementary DNA (cDNA) was performed using the RevertAid™ First Strand cDNA Synthesis Kit (Fermentas/Thermo Fisher Scientific), by taking 1 µg RNA for each sample. Semi-quantitative reverse transcription polymerase chain reaction (RT-PCR) was performed using an Eppendorf thermocycler amplification for 20–35 cycles (depending on the saturation level of the genes amplified) using 1 min denaturation at 94 °C, 1 min annealing at 50 °C, 1 min extension at 72 °C; and 72 °C for 10 min for final extension. The following primers were used:

CCAT1 forward: TCCATCTGGAGCATTCCTG  
 CCAT1 reverse: AGCCATACAGAGCCAACCTG  
 c-MYC forward: AATGAAAAGGCCCAAGGTA  
 GTTATCC  
 c-MYC reverse: GTCGTTTCCGCAACAAGTCCT  
 CTTC  
 PVT1 forward: GCTGTCAAAGAGGCCTGAAG  
 PVT1 reverse: ACATTCCTGCTGCCGTTTT  
 18S forward: TGTGCCGCTAGAGGTGAAATT  
 18S reverse: TGGCAAATGCTTTCGCTTT

## RNA fluorescence in situ hybridization (RNA FISH)

Cells were seeded on 18 mm coverslips and fixed for 20 min in 4% PFA, then washed in 70% ethanol overnight. Coverslips were then washed twice with 10% formamide for probes purchased from Stellaris or fluorescent light-up aptamer (FLAP) probes (Tsanov et al. 2016) diluted in 4× saline-sodium citrate (SSC). For single-molecule FISH (smFISH) on endogenous transcripts, fluorescence-labeled DNA probes targeting the *c-MYC* exon sequence (570 nm, ~10 ng probe, Stellaris), *CCAT1* intron sequence (670 nm, ~10 ng probe, Stellaris), *CCAT1-L* sequence (570 nm, ~10 ng probe, FLAP), *CCAT1-5L* (570 nm, ~10 ng probe, FLAP), *PVT1* exon sequence (570 nm, ~10 ng probe, FLAP) were hybridized overnight at 37 °C in a dark chamber in 10% formamide. The next day, cells were washed twice with 10% formamide diluted in 4× SSC for 30 min at 37 °C and then washed with 1× phosphate buffered saline (PBS). To reduce photobleaching, the slides were mounted in GLOX (catalase, glucose oxidase) buffer (pH 8.1 mM, 2× SSC, 0.4% glucose) supplemented with 3.7 ng of glucose oxidase (Sigma-Aldrich G2133-10KU) and 1 µL catalase (Sigma-Aldrich 3515) prior to imaging.

## RNA FISH and immunofluorescence

For smFISH on endogenous transcripts, fluorescence-labeled DNA probes targeting the *CCAT1* exon sequence (570 nm, 10 ng, Stellaris) and the *CCAT1* intron sequence (670 nm, 10 ng, Stellaris) were hybridized overnight at 37 °C in a dark chamber in 10% formamide together with the primary antibody, anti-SRRM2 (rabbit, Abcam, ab122719). The next day, cells were washed with 10% formamide, and a secondary antibody (Alexa Fluor 488 goat anti-rabbit IgG; Abcam) was added for 30 min at 37 °C, followed by washing with 1× PBS. To reduce photobleaching, the slides were mounted in GLOX buffer.

## Quantitative RNA FISH

Following RNA FISH experiments, three-dimensional (3D) stacks of cells were acquired using a wide-field fluorescence microscope at 60× magnification. Specifically, 51 z-planes were acquired for each cell with 300 nm steps. After the acquisition, the images underwent deconvolution using Huygens software and were transferred to Imaris software (Oxford Instruments, Abingdon, UK) for image processing. In Imaris, the signal of each RNA spot was evaluated using “spot object” and the transcription site was designated using “surface object.” To calculate the number of free *CCAT1* transcripts during mitosis vs. interphase, each spot was counted, following the average calculation for each group. To measure the distances of the RNAs from the transcription site, spots in the Cy5 channel were analyzed under “shortest distance from surface.”

To calculate the number of single *CCAT1* transcripts at the transcription site, the sum of the fluorescence intensity of transcription sites was measured using the Imaris “surface tracker.” Next, the common value for the fluorescence signal of single transcripts was measured using the Imaris “spot tracker” and defined as a single RNA molecule. Then, the mean intensity of a single transcript was multiplied by the number of pixels covered by the transcription site. This value was extracted from the sum of intensity at the transcription site (TS).

$$\frac{\text{TS intensity}}{\text{Single transcript intensity}} = \text{RNA number per TS}$$

## Fluorescence microscopy

Wide-field fluorescence images were obtained using the cellSens system based on an Olympus IX83 fully motorized inverted microscope (60× UPlanXApo objective, 1.42 NA) fitted with a Prime BSI scientific complementary

metal–oxide–semiconductor (sCMOS; Teledyne) driven by CellSens software. Images were then created using the FIJI software package.

## Statistical analysis

The experiments presented were carried out at least three times. Statistical analysis was performed using GraphPad Prism 10 software. For quantification of *CCAT1* unspliced transcripts in interphase versus mitotic cells, data were analyzed with independent-samples *t*-tests. For quantification of *CCAT1* transcripts at the transcription site, proximity to the transcription site, location of introns in the cytoplasm, percentage of transcription sites per cell, and semi-quantitative RT-PCR data were analyzed using one-way ANOVA, followed by Tukey's post hoc analysis. Treatment groups for which all values were constant (0% cells) were analyzed separately from other treatments using one-sample *t*-tests against a constant mean value of zero. Finally, a false discovery rate (FDR) correction was applied to adjust for multiple testing.

## Results

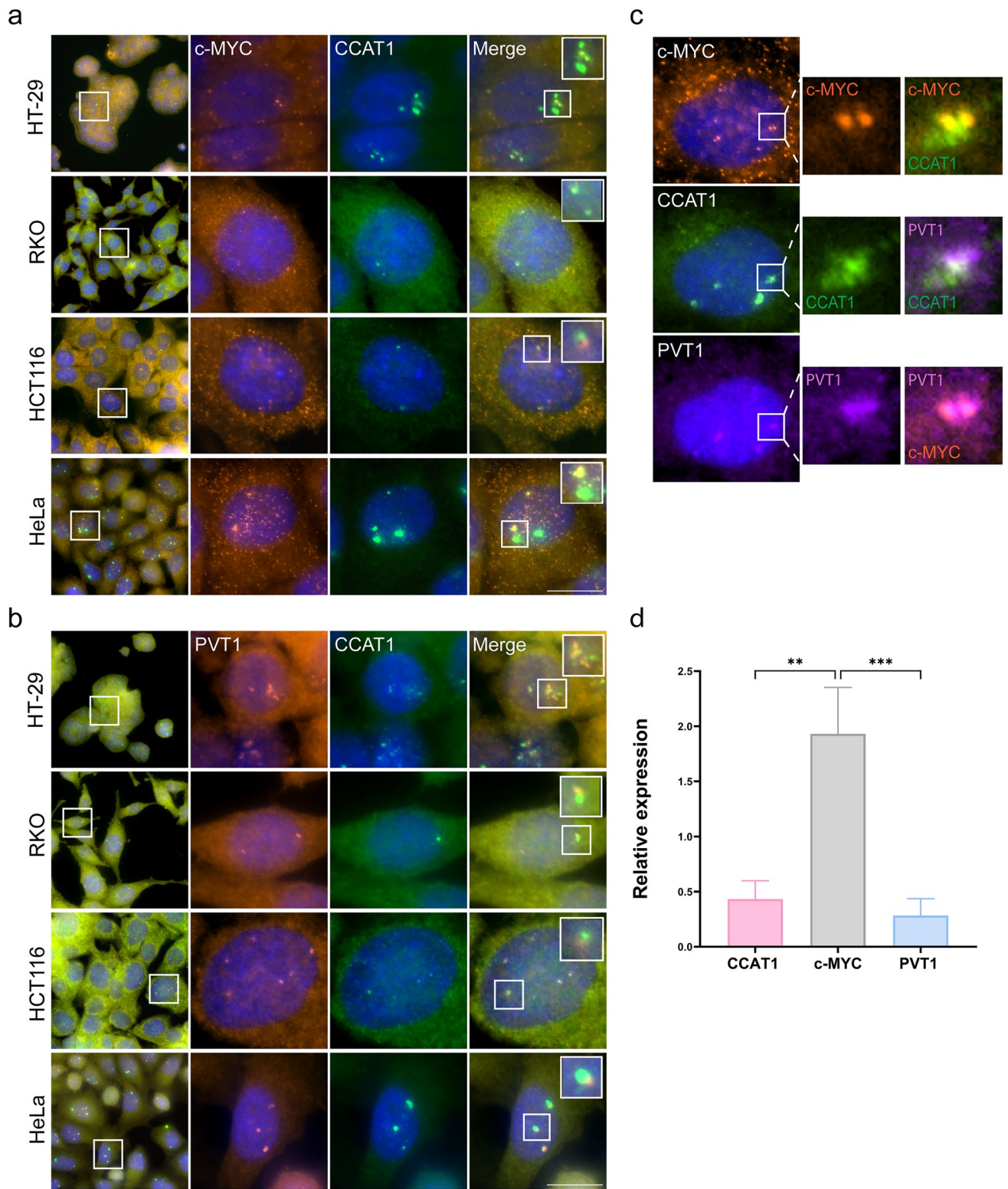
### *CCAT1* lncRNA is highly abundant in the nucleus of HeLa cells

The *MYC* 8q24 locus has been demonstrated to express cancer-specific lncRNAs, including *CCAT1* and *PVT1* (Xiang et al. 2014; Jin et al. 2019). *CCAT1* is located 515 kb upstream of the *MYC* locus, and *PVT1* is 53 kb downstream of *MYC* (Jin et al. 2019). We examined whether the three genes could be transcriptionally active at the same time by detecting their active transcription sites using RNA FISH probes designed to hybridize with their cognate endogenous RNAs. The expression of these genes was tested in four human cancer cell lines. We tested various colorectal cancer (CRC) cell lines, as it is known that *CCAT1* is expressed in CRC: HT-29 colorectal adenocarcinoma cells, RKO colon carcinoma cells, and HCT116 colorectal carcinoma cells (McClelland et al. 2016; Yang et al. 2019; Kam et al. 2014). We also decided to include a different type of cancer cell type, HeLa cervical cancer adenocarcinoma cells (Chen et al. 2021; Li et al. 2023). As expected, the active transcription sites of *MYC* and *CCAT1* were found in all colorectal cell lines and were in close proximity (Fig. 1a). Surprisingly, HeLa cells had very large *CCAT1* transcription foci relative to the other cell lines. *MYC* mRNAs were also abundant. Testing *PVT1* lncRNA expression and *CCAT1* yielded similar results, namely that *PVT1* was expressed in all cells, and the highest signal was observed in HeLa cells (Fig. 1b). Altogether, we found that the *MYC* locus was

transcriptionally active and that *CCAT1* expression seemed to be the highest of the three genes expressed from this locus. Notably for HeLa cells, the cells typically exhibited 3–4 active transcription sites expressing these genes. Indeed, chromosome 8 usually appears in three copies in HeLa cells, while chromosome translocation is also common (Landry et al. 2013). When we tested the expression of all three genes using three different probes, we found that *CCAT1*, *c-Myc*, and *PVT1* active genes were co-localized (Fig. 1c). Although *CCAT1* lncRNA showed the strongest signal at the microscopy level, we wanted to determine whether the expression levels of *CCAT1* were indeed relatively high. Semi-quantitative RT-PCR showed that *MYC* expression levels in HeLa cells were much higher than those of *CCAT1* (Fig. 1d). This finding implies that the strong *CCAT1* signals observed in the microscopy images at the gene locus do not mean that *CCAT1* is expressed at remarkably high expression levels. Rather, the lncRNA might be accumulating at the active gene locus, as it is known that many lncRNAs associate with chromatin and with their own gene loci (Calandrelli et al. 2023).

### *CCAT1* accumulates at the active gene locus

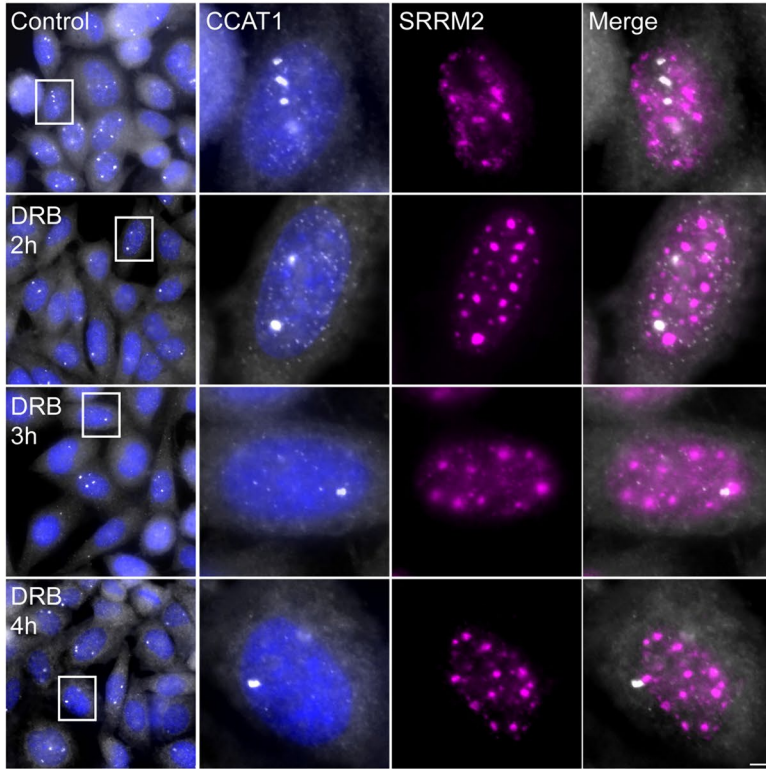
To test the hypothesis that the large size of the *CCAT1* site of transcription is due to lncRNA accumulation at the locus, we examined how transcription inhibition affects the size of these transcription sites. Namely, if the enlarged size of the transcription site is due to transcript accumulation, then after treatment with a transcription inhibitor, the signal should not disappear, as would be expected from a normal active gene under transcription inhibition conditions (Brody et al. 2011; Darzacq et al. 2007). First, we used the transcription inhibitor 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB), which inhibits the CDK9 kinase (kinase subunit of positive transcription elongation factor) (Bensaude 2011) to inhibit RNA polymerase II (Pol II) transcription. Even after 4 h of treatment with DRB, *CCAT1* transcription sites were still observed (Fig. 2a). However, the percentage of visible transcription sites per cell was dramatically decreased after DRB inhibition, and ~35% of the cells did not show any visible active genes (Fig. 2b). To confirm that the DRB treatment was effective, immunofluorescence against the SRRM2 protein was performed in parallel. As expected (Rino et al. 2007; Spector and Lamond 2011), nuclear speckles that were marked by SRRM2 and that usually have an irregular shape were transformed into rounded structures due to transcription inhibition. Many genes are known to associate with nuclear speckles when transcriptionally active (Faber et al. 2022; Kim et al. 2020; Belmont 2021). However, no association between the nuclear speckles and *CCAT1* transcription sites was detected under normal or DRB-treated conditions. Similarly, when the transcription inhibitor actinomycin



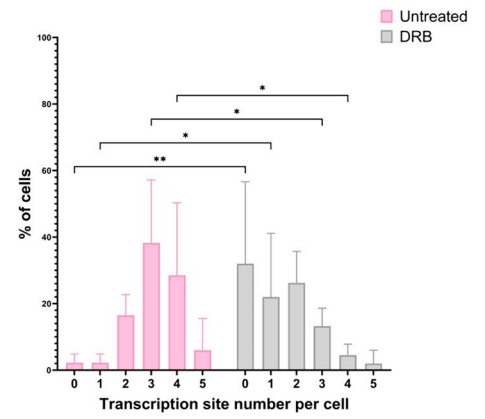
**Fig. 1** *CCAT1* lncRNA is highly abundant in nuclei of HeLa cells. Detection of **a** *CCAT1* lncRNA (exon; green) with *MYC* mRNA (orange) or with **b** *PVT1* lncRNA (orange) by RNA FISH in HT-29, RKO, HCT116, and HeLa cells. Large foci are the active genes and small dots are the single RNAs. Hoechst DNA stain is in blue. Boxed areas are enlarged. **c** RNA FISH in HeLa cells of *MYC* (orange), *CCAT1* (exon; green), and *PVT1* (purple) RNAs. Scale bars, 10  $\mu$ m.

**d** Expression levels of *CCAT1*, *MYC* and *PVT1* RNAs in HeLa cells measured by semi-quantitative RT-PCR. The *18S* gene was used as a housekeeping gene. Data were analyzed using one-way ANOVA, followed by Tukey's post hoc analysis. A significant difference was found in the relative expression levels between *MYC* to *CCAT1* and *PVT1*. \*\* $P < 0.01$ , \*\*\* $P < 0.001$

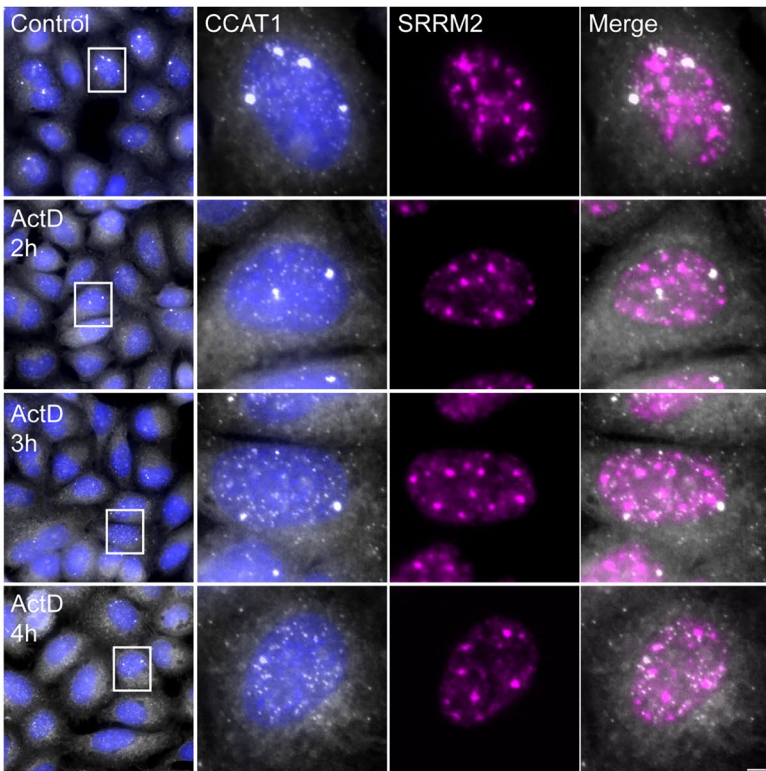
a



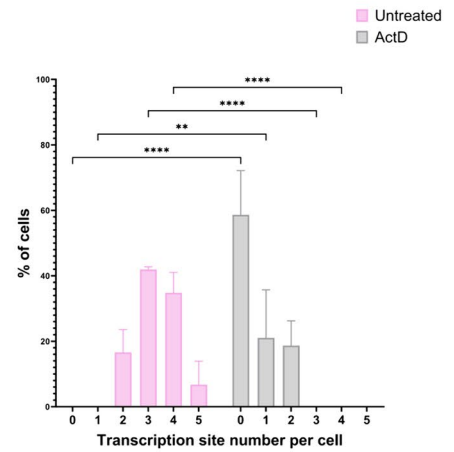
b



c



d



**Fig. 2** *CCATI* lncRNAs accumulate on the gene locus. The transcription inhibitors **a** DRB or **c** ActD decreased *CCATI* detection at the site of transcription, but small foci were still observed (4 h of treatment). *CCATI* (exon; gray) was detected together with anti-SRRM2 (magenta) that marks nuclear speckles. Hoechst DNA stain is in blue. Boxed areas are enlarged. Scale bars, 10  $\mu$ m. **b** The percentage of *CCATI* transcription foci per cell after treatment with DRB or **d** ActD. A minimum of  $n=50$  cells were selected for each analysis. Data were analyzed using one-way ANOVA, followed by Tukey's post hoc analysis. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\*\* $P < 0.0001$

D (ActD) was applied, a similar but stronger effect was observed, as no visible *CCATI* transcription sites were observed in ~60% of the cells (Fig. 2c,d). The persistence of the *CCATI* transcription sites after transcription inhibition, rather than completely disappearing as would be expected, indicated that lncRNA accumulation was occurring at the site of transcription. Since introns are prominently detected at the site of transcription (Cote et al. 2024; Darzacq et al. 2007; Mor et al. 2010), we examined whether the intron and exon sequences of *CCATI* overlapped at these foci. Indeed, using RNA FISH probe sets against the *CCATI* exons or intron, we found that the intron sequences were most prominent at these foci. The latter were reduced in size when cells were treated with the transcription inhibitors (Supplementary Figure S1a, b), implying that these foci are most likely the sites at which *CCATI* is transcribed. This means that under regular conditions, the active *CCATI* genes produce *CCATI* transcripts, and a subpopulation of these transcripts remains associated with the gene locus.

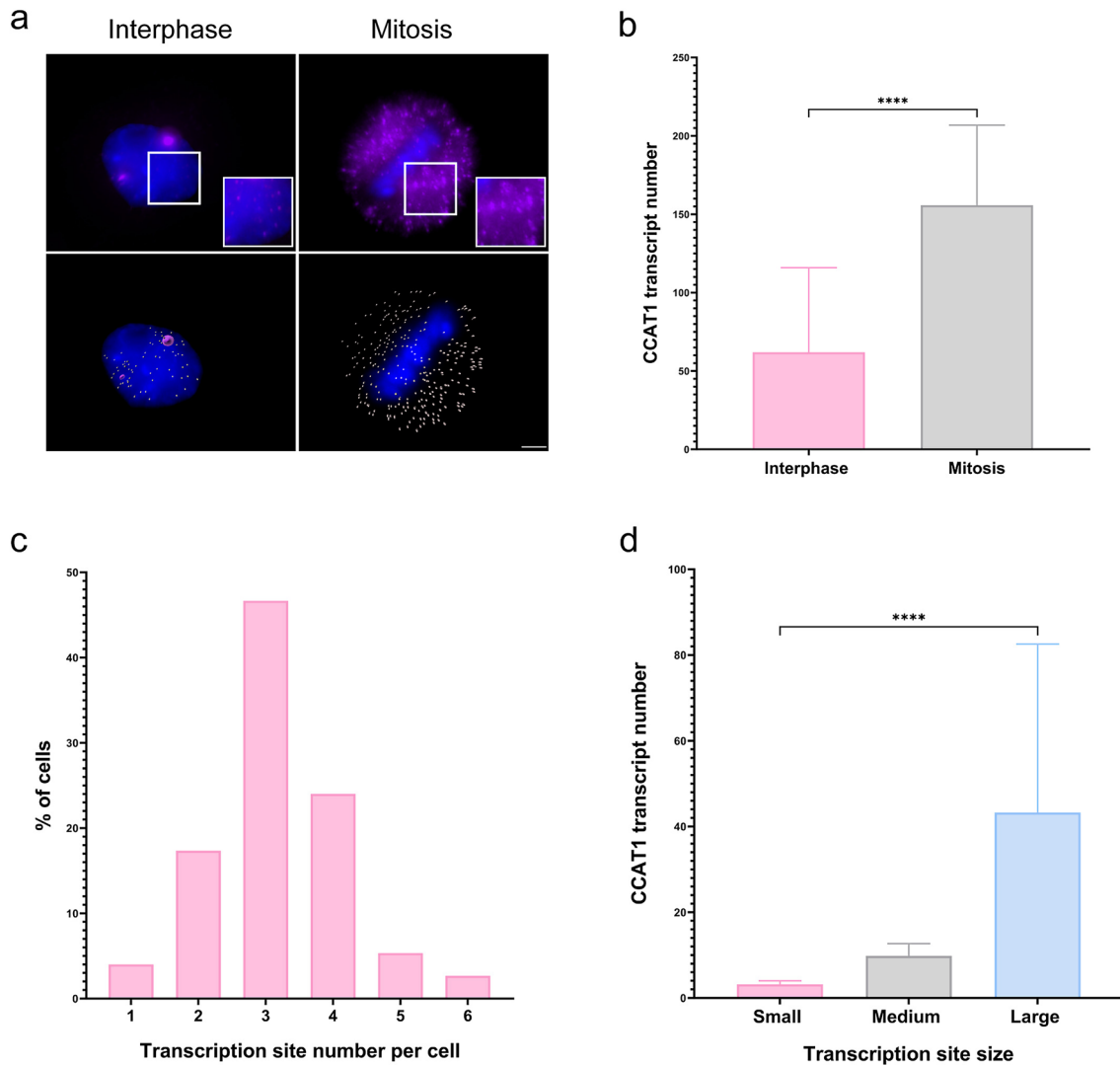
Since we speculated that *CCATI* transcripts were accumulating at the gene locus, we assumed that during cell division, these transcripts would dissociate from the chromosome. Also, active transcription sites were not expected to be seen since the rates of transcription decrease dramatically during cell division (Hartl et al. 1993; Palozola et al. 2017). Indeed, no *CCATI* transcription sites were observed during cell division (Supplementary Figure S2), and no accumulation of *CCATI* transcripts was observed on metaphase chromosomes (Fig. 3a). Rather, many *CCATI* transcripts were now seen throughout the mitotic cell. The number of free *CCATI* transcripts in cells during metaphase was counted and compared with the number of free transcripts in cells during interphase. A large increase in free *CCATI* transcripts in the cell was observed (Fig. 3a, b). This indicates that during interphase, *CCATI* transcripts accumulate at the locus, and during mitosis, they are released from the chromosomes. We could now quantify the number of *CCATI* transcripts that were associated with the locus in cells during interphase. First, the numbers of *CCATI* loci were counted ( $n=300$  cells), which showed that ~50% of the cells contained at least three active sites of transcription (Fig. 3c). Next, the estimated number of *CCATI* transcripts associated with a single transcription site was calculated using the

single *CCATI* transcript average intensity (Fig. 3d). Since the *CCATI* loci were not uniform in size, they were divided into several different subgroups. Large transcription sites contained ~40 transcripts, moderate-sized had ~10 transcripts, and the small ones contained less than 5 transcripts. Therefore, a cell that had three large loci would amount to a total of ~120 *CCATI* transcripts associated with the gene loci. This number is within the range of the average number of *CCATI* transcripts measured in the metaphase cells (Fig. 3b; ~150 transcripts). Altogether, these numbers correlate well, as there are also free *CCATI* transcripts in the nucleoplasm and cytoplasm of interphase cells.

### ***CCATI* transcripts undergo post-transcriptional splicing**

We next wanted to examine whether the transcripts associated with the *CCATI* gene loci were spliced transcripts in all cell lines expressing *CCATI*. Using the RNA FISH probe sets against the *CCATI* exons or intron, the intron signal overlapped with the site of transcription in HT-29, RKO, HCT116, and HeLa cells (Fig. 4a). Typically, intron signals are seen predominantly on sites of transcription, since much of the splicing occurs co-transcriptionally. Here, all transcription site signals contained intron signals, suggesting that the associated *CCATI* lncRNAs were unspliced at this stage. However, unspliced transcripts were also detected in the nucleoplasm of the HeLa cells (Fig. 4a, Supplementary Figure S1a, b). In general, the vast majority of the *CCATI* transcripts in the cells were spliced RNAs, and the unspliced transcripts were mostly found near the site of transcription. Unspliced transcripts were nuclear only. This was suggestive of post-transcriptional splicing occurring after release from the transcription site. Since unspliced RNAs can accumulate in nuclear speckles (Gordon et al. 2021; Mor et al. 2016), we examined whether there was an association of unspliced *CCATI* transcripts with nuclear speckles. RNA FISH applied to *CCATI* unspliced transcripts and staining for SRRM2, a core nuclear speckle protein, showed no association between the two (Fig. 4b).

Next, we focused on the location of the *CCATI* unspliced transcripts in the nucleus. Spliced *CCATI* transcripts were dispersed over the entire nucleus, while unspliced transcripts were localized mostly next to the site of transcription (Fig. 5a). Next, we measured the distance of *CCATI* unspliced transcripts from the gene foci (Fig. 5b). When calculating the distance of unspliced *CCATI* transcripts from the gene locus, we found that most transcripts were located less than 3  $\mu$ m from the transcription site, and less than 16% of the unspliced transcripts were located more than 6  $\mu$ m from the locus (Fig. 5c). In addition, to test whether the other *CCATI* isoforms *CCATI-L* and *5L-CCATI* were also post-transcriptionally spliced, RNA FISH was performed



**Fig. 3** *CCAT1* transcripts are released from chromosomes during cell division. **a** Detection of *CCAT1* RNAs by RNA FISH in cells at interphase (left) and metaphase (right). The top panels show original, with zoomed images presented at high signal intensity (boxes). The bottom panel shows analyzed images (Imaris). White spots show the *CCAT1* RNA. Large pink dots are sites of transcription. Scale bar, 3  $\mu$ m. **b** The total number of *CCAT1* RNAs counted in interphase versus met-

aphase cells. Data were analyzed with independent-samples *t*-tests. **c** The distribution of *CCAT1* transcription sites per cell under steady-state conditions. **d** Transcription foci were divided into subgroups according to their size. Each group contained different numbers of transcripts. A minimum of  $n=50$  cells were selected for each statistical analysis. \*\*\*\* $P < 0.0001$

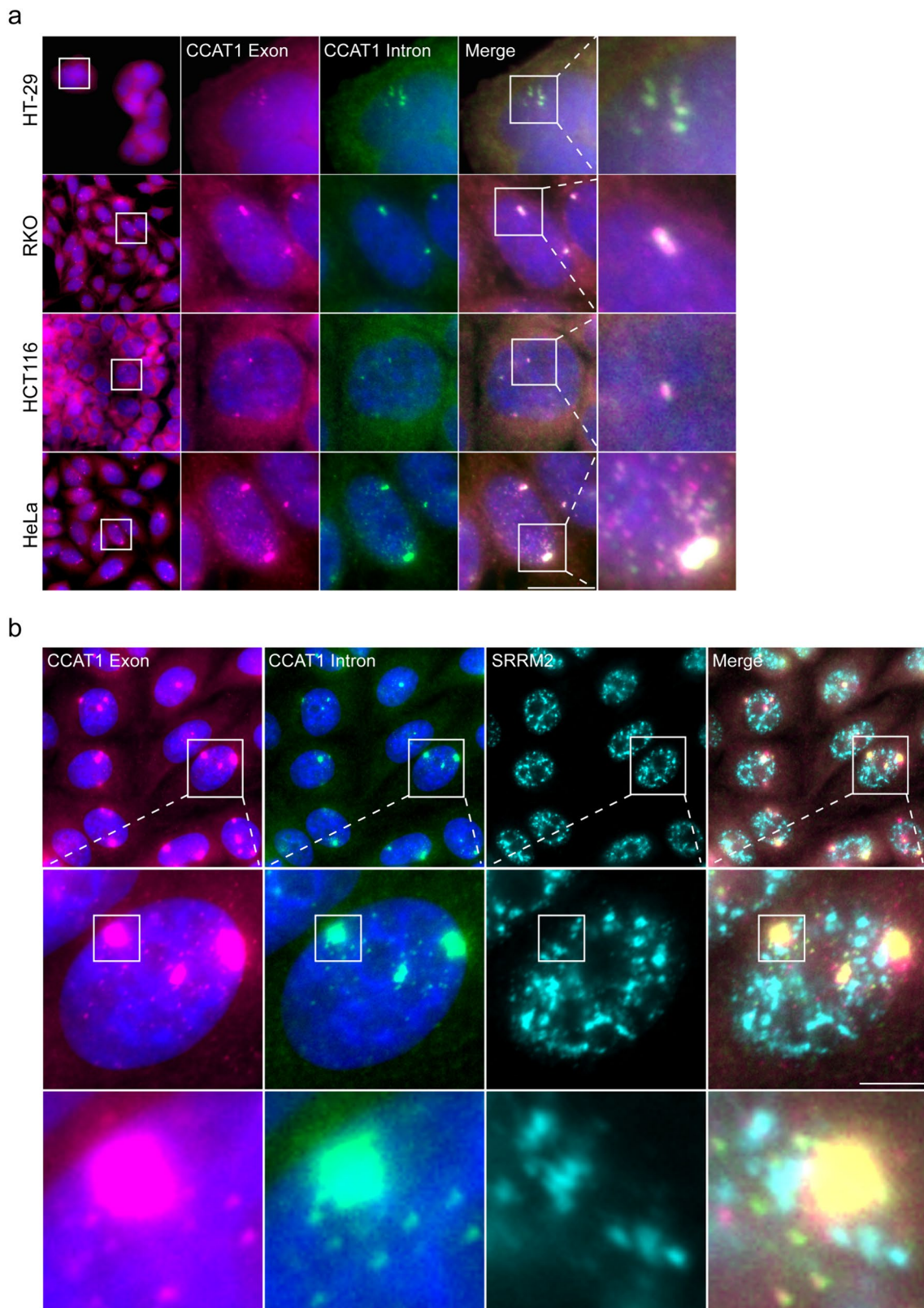
using probes that detect either *CCAT1-L* or *CCAT1-5L* only, together with probes directed to the intron. We observed that *CCAT1-L* and *CCAT-5L* expression overlapped with the intron signal in the nucleoplasm, and as observed with the *CCAT1* common isoform, the unspliced transcripts were localized close to the transcription site (Supplementary Figure S3). The presence of *CCAT1-L* and *CCAT1-5L* unspliced transcripts in proximity to the gene loci confirms that the post-transcriptional splicing is common to the three different isoforms. Taken together, these findings suggest that the splicing of *CCAT1* is post-transcriptional and occurs soon after release from the gene and before the transcripts

diffuse away from the gene. This finding is consistent with the recently published evidence demonstrating that highly transcribed genes are post-transcriptionally spliced (Cote et al. 2024).

### Unspliced *CCAT1* transcripts are exported to the cytoplasm during splicing inhibition

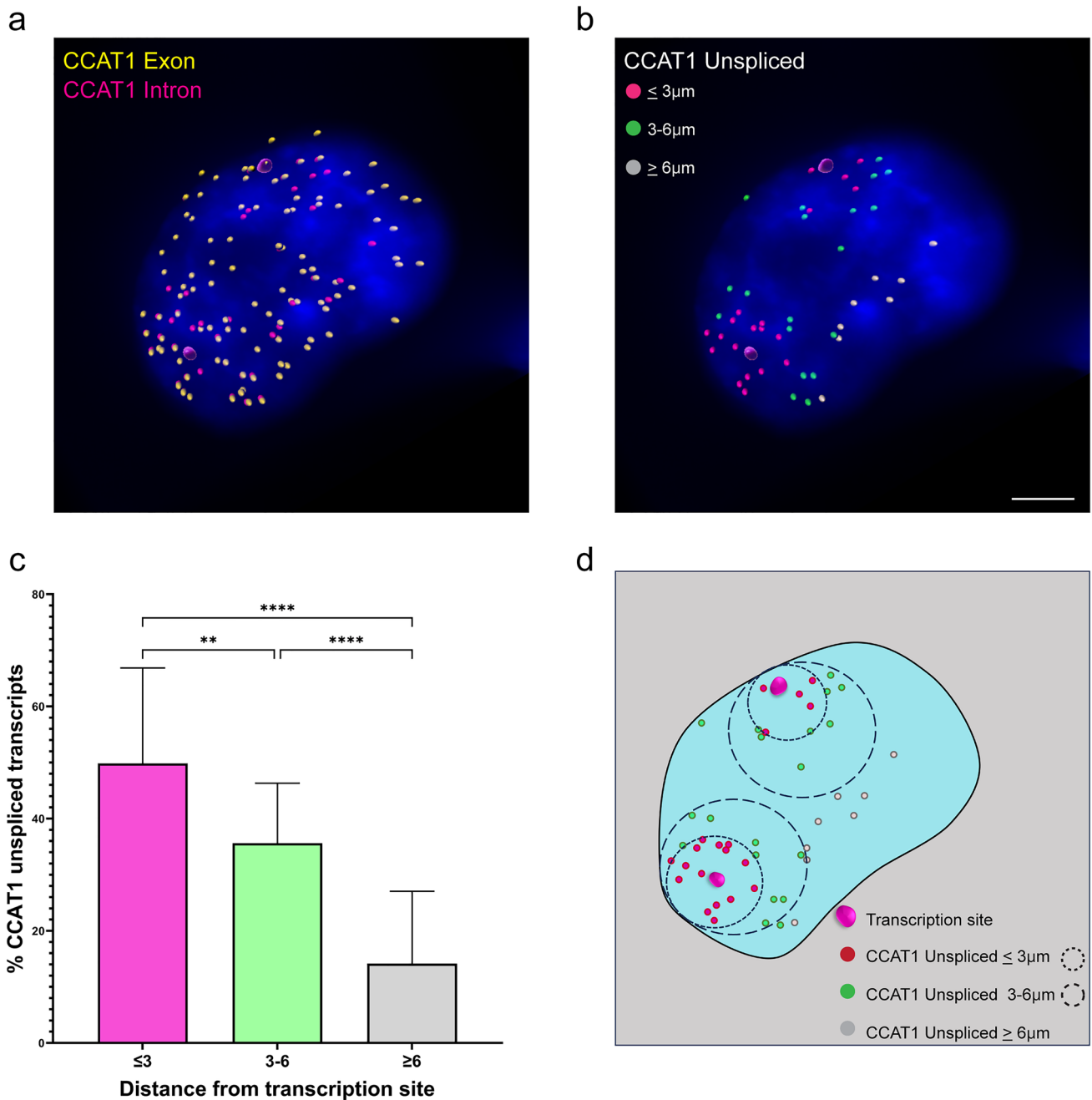
It was previously demonstrated that splicing inhibition can lead to the accumulation of unspliced transcripts at the site of transcription and in nuclear speckles (Cote et al. 2024). These transcripts are retained in the nucleus and are not





**Fig. 4** *CCAT1* unspliced transcripts are present in the nucleoplasm of HeLa cells and do not localize with nuclear speckles. **a** RNA FISH detection of *CCAT1* exon (pink) and *CCAT1* intron regions (green). HeLa cells displayed high levels of unspliced *CCAT1* transcripts at the sites of transcription and low levels of unspliced transcripts in the

nucleoplasm. HT-29, HCT116, and RKO cells showed low levels of *CCAT1* introns (unspliced transcripts) at the site of transcription. **b** *CCAT1* active genes and nucleoplasmic transcripts were not associated with nuclear speckles marked by anti-SRRM2 (cyan). Hoechst DNA stain is in blue. Boxed areas are enlarged. Scale bars, 10 μm



**Fig. 5** Unspliced *CCAT1* transcripts are found in close proximity to the active gene. **a** *CCAT1* spliced and unspliced RNA was detected in untreated cells and analyzed by Imaris; *CCAT1* exon spots (cyan) and *CCAT1* intron spots (red). Large pink dots are sites of transcription. **b** Distances of *CCAT1* unspliced RNAs from the active genes in untreated cells are color-coded. Pink spots show the unspliced transcripts located less than 3  $\mu\text{m}$  from the active gene; green spots between 3 to 6  $\mu\text{m}$ ; and white spots more than 6  $\mu\text{m}$ . Large pink dots

are sites of transcription. **c** Measurements of the distances of *CCAT1* unspliced transcripts from the active genes. Data were analyzed with a one-way ANOVA, followed by Tukey's post hoc analysis.  $**P < 0.01$ ,  $****P < 0.0001$ . A minimum of  $n = 50$  cells were selected for statistical analysis. Scale bar, 3  $\mu\text{m}$ . **d** Schematic illustration of the measured distances of unspliced *CCAT1* transcripts from the active genes

exported to the cytoplasm. We did not find any accumulation of *CCAT1* unspliced transcripts in nuclear speckles under regular conditions (Fig. 4b), and so we examined their fate under conditions of splicing inhibition. We used

pladienolide B (PLB), a splicing inhibitor that binds to the SF3B1 subunit of the U2 small nuclear ribonucleoproteins (snRNP) and blocks spliceosome activity (Effenberger et al. 2017). When PLB was added to HeLa cells, the presence

of *CCATI* at the gene locus was reduced already after 1 h (Fig. 6a). After 6 h of treatment with PLB, *CCATI* transcription sites not only did not show any visible accumulation of unspliced transcripts, but in a considerable percentage of the cells the large loci disappeared (Fig. 6b). Surprisingly, in some cells, unspliced transcripts appeared in the cytoplasm (Fig. 6a, c). As a control, we examined another transcript, the *MKI* transcript that encodes Ki-67, under splicing inhibition conditions, and unspliced *MKI* mRNAs were not detected in the cytoplasm (Supplementary Figure S4). Additionally, unspliced *CCATI* transcripts did not appear to be associated with nuclear speckles (Fig. 6a). Quantification of the unspliced transcripts showed that ~40% of the introns were found in the cytoplasm during splicing inhibition, while the rest were in the nucleus. We postulated that following splicing inhibition, *CCATI* transcripts, which normally are highly transcribed, undergo nuclear export despite intron retention. Taken together, these findings might suggest that *CCATI* at steady-state conditions is post-transcriptionally spliced and that the unspliced version is not detected as an aberrant RNA with a splicing defect but might be a normal candidate for export under certain conditions.

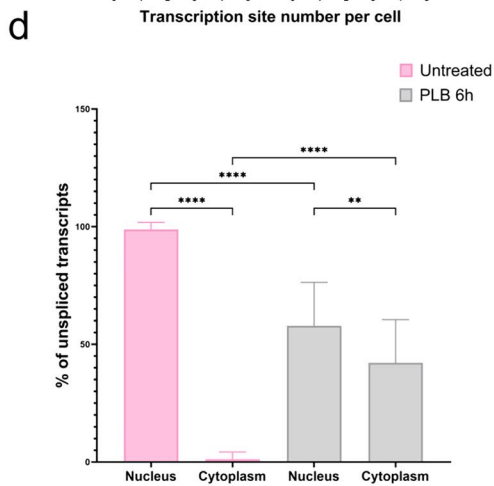
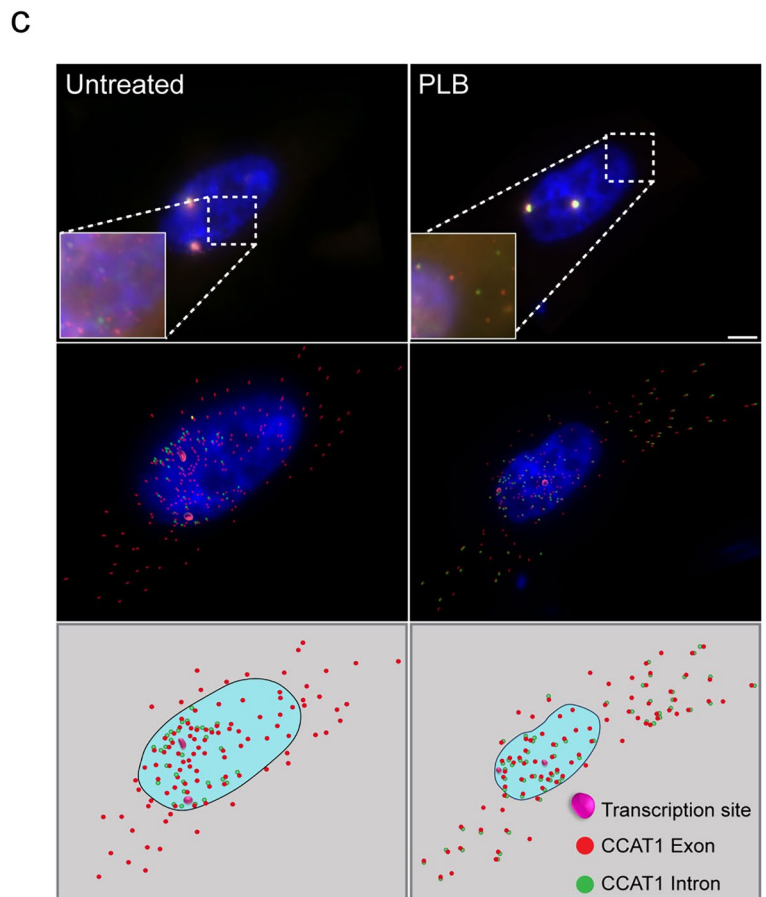
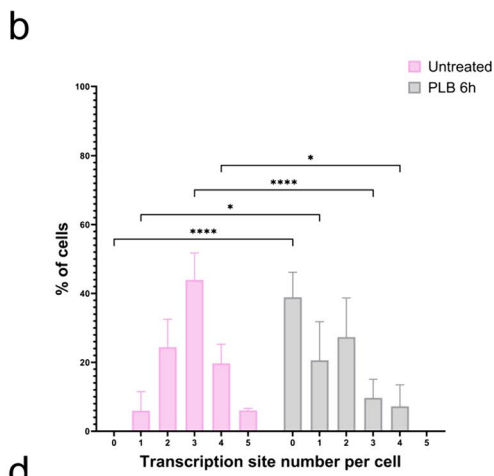
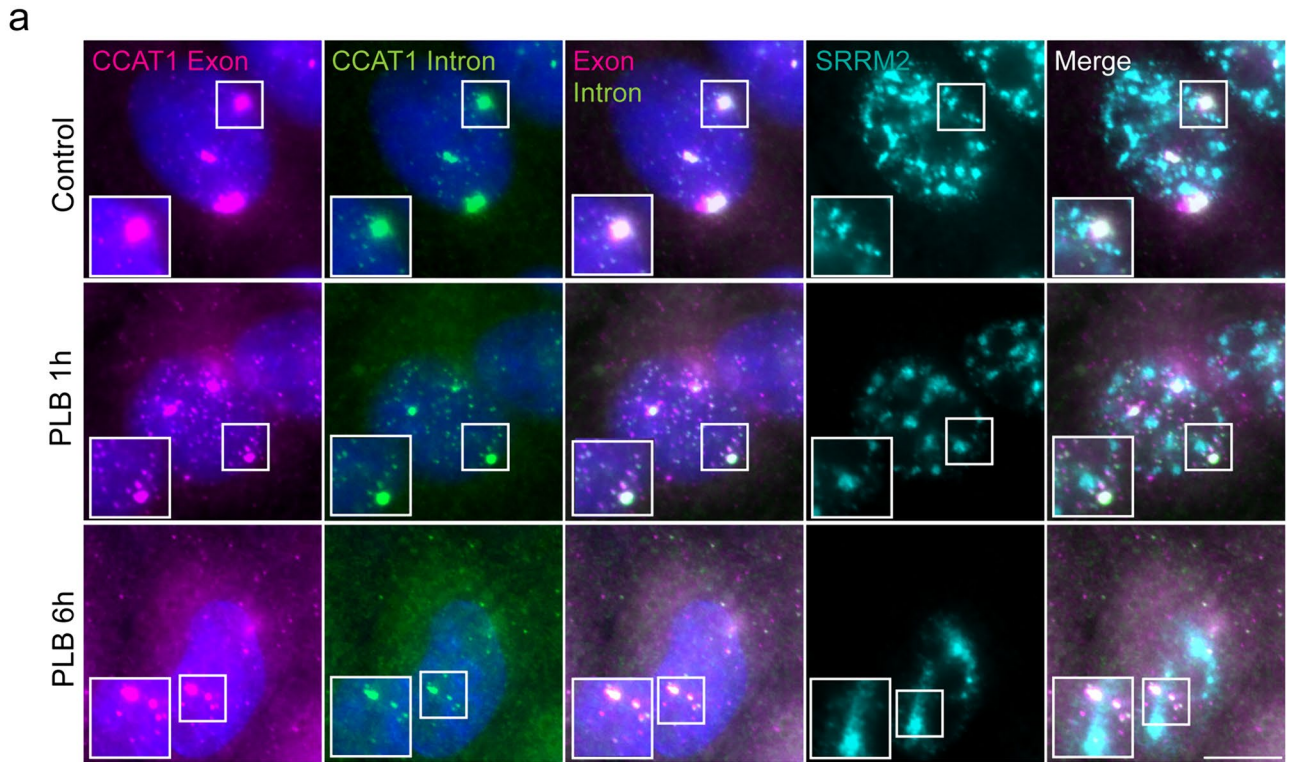
## Discussion

Super-enhancers are large clusters of enhancers spread over large DNA regions that regulate gene expression via a variety of mechanisms. Similar to typical enhancers, SEs can be located far away from their associated gene, and in such cases, genome folding brings the enhancer into close proximity to the gene promoter, where their interactions are mediated through transcription activators and the transcriptional initiation complex, to control gene expression (Deng et al. 2012; Ye et al. 2020). SEs display enrichment of binding sites for transcription factors and activators relative to a typical enhancer (Whyte et al. 2013; Pott and Lieb 2015). SEs in cancer cells have been shown to display different expression patterns and enhancer usage relative to healthy cells (Loven et al. 2013). Therefore, SEs are considered key regulators of oncogene expression in different tumors (Tang et al. 2020).

The *MYC* locus is a tumor type-specific super-enhancer that expresses different super-enhancer-derived ncRNAs (Hnisz et al. 2013; Amjadi-Moheb et al. 2021). Within the course of our research, we focused on the expression of the *CCATI* lncRNA at the *MYC* locus. *CCATI* is upregulated in CRC already in the early phase of tumorigenesis as well as at late stages of the disease, and hence, *CCATI* can be used as a potential biomarker for screening, diagnosis, and prognosis for patients with CRC (Mizrahi et al. 2015; Xiao et al. 2021). In the current study, we examined lncRNA localization in human cell lines on the single RNA level and characterized

*CCATI* expression, with an emphasis on its detection at the site of transcription in HeLa cells, which showed very prominent sites of *CCATI* transcription. We were surprised to observe this significant upregulation of *CCATI* in HeLa cells, since *CCATI* is mostly correlated with colon cancer (Mizrahi et al. 2015; Shang et al. 2020). However, *CCATI* is also expressed in ovarian cancer (Wang et al. 2020a). We found that *CCATI* was highly detectable relative to other cell lines that express *CCATI*; several large sites of transcription were observed and high numbers of transcripts were detected in the nucleus and cytoplasm (Fig. 1a, b). We assumed that the large foci were transcription sites since active transcription sites contain the highest number of transcripts when examined by RNA FISH, and the overlapping with the *MYC* and *PVT1* transcription sites strengthened this assumption. The relative RNA expression levels of *MYC* in HeLa cells were twofold higher than *CCATI* levels, while the *PVT1* levels were similar to *CCATI* (Fig. 1d).

To verify that the foci were indeed sites of transcription, we used transcription inhibitors. Following treatment with DRB or ActD, the *CCATI* foci were still observed to a certain extent (Fig. 2a, c), although the number of cells exhibiting these foci decreased significantly. Also, the intron and exon signals predominantly co-localized at these foci, as would be expected from transcription sites that are the areas that contain most intronic sequences. The fact that the foci did not completely disappear, as would be expected of transcription sites during transcription inhibition, suggested that in addition to transcription, there is an accumulation of the lncRNAs on their gene locus. The accumulation of lncRNAs on chromatin is known (Guo et al. 2020; Schlackow et al. 2017; Calandrelli et al. 2023). The treatment with ActD had a stronger effect on transcription site disappearance (Fig. 2b, d). This effect may be due to the ActD inhibition mechanism, which functions through intercalation into GC-rich sequences and prevents RNA polymerase progression. The intercalation of ActD into the double helix may interrupt the binding of *CCATI* to chromatin, which subsequently reduces foci size. Therefore, we speculate that the large transcription sites observed are formed due to *CCATI* transcript accumulation on the gene locus and not due to an unusually high transcription rate modulated by RNA Pol II activity. These results correlate with previous studies showing that lncRNAs can modulate gene expression locally by accumulation at or near their site of transcription (Kopp and Mendell 2018; Gil and Ulitsky 2020). lncRNA accumulation on chromatin can be associated with its regulation or function. One of the earliest examples of RNA accumulation is the lncRNA *XIST* (Brockdorff et al. 1992; Brown et al. 1991), a key regulator of X inactivation, which mediates the silencing of the inactive X chromosome through the subsequent recruitment of epigenetic regulators (Clemson et al. 1996; Lee 2012). Another example is polyadenylated nuclear



**Fig. 6** Splicing inhibition reduces *CCATI* expression and leads to the appearance of unspliced transcripts in the cytoplasm. **a** Pladienolide B (PLB) treatment for 1 and 6 h decreased *CCATI* detection at the active genes. Unspliced transcripts were detected at the cytoplasm after 6 h. *CCATI* exon (pink), *CCATI* intron (green), anti-SRRM2 for marking nuclear speckles (cyan). **b** The percentage of *CCATI* gene foci per cell after treatment with PLB for 6 h. **c** RNA FISH of *CCATI* exon and intron regions in untreated and splicing-inhibited conditions. *CCATI* exon (red), *CCATI* intron (green). The top panel represents original and zoomed images at higher intensity (boxes). The middle panel shows analyzed images (Imaris). Red spots show *CCATI* exon and green spots show *CCATI* intron. The bottom row presents a schematic illustration. Scale bar, 4  $\mu\text{m}$ . **d** The percentage of unspliced transcripts in the cytoplasm compared with nuclei in control and under splicing inhibition conditions. Data were analyzed with **b** independent-sample *t*-tests and **d** one-way ANOVA, followed by Tukey's post hoc analysis. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\*\* $P < 0.0001$ . A minimum of  $n = 50$  cells were selected for statistical analysis

RNA (*PAN*) lncRNA, which is a viral RNA that accumulates and binds to its own promoter and robustly activates *PAN* lncRNA expression (Campbell and Izumiya 2020). Multiple studies have indicated that RNAs with processing defects, particularly in their splicing, accumulate in nuclear foci located near their site of transcription. For example, the mutant  $\beta$ -globin mRNA that has a defect in either splicing or 3' formation has been reported to accumulate close to its site of transcription, even in the presence of transcription inhibitors, which under normal conditions cause the rapid release of wild-type human  $\beta$ -globin RNA from the vicinity of the gene (Custodio et al. 1999).

Co-transcriptional splicing of nascent RNA is a central mechanism for gene regulation in normal cells, while in some cancer cells, defective splicing machinery can shift splicing to occur post-transcriptionally (Bentley 2014). For example, a single missense mutation (Ser34Phe) in the zinc finger domain of the conserved splicing factor U2AF1 is common in different types of cancers, leading to the post-transcriptional splicing of  $\beta$ -globin and *FXR1* mRNAs (Coulon et al. 2014). A major question in the splicing field is how tightly associated are the processes of transcription and splicing. The balance between co-transcriptional and post-transcriptional splicing is regulated. Deep sequencing studies have shown that many RNAs undergo co-transcriptional splicing (Tilgner et al. 2012; Sanchez-Escabias et al. 2022). lncRNAs, however, have been reported to be less efficiently spliced than protein-coding genes, regardless of transcript activity (Mukherjee et al. 2017). One possible explanation for this phenomenon is the lack of splicing enhancer sequences at lncRNA splice sites (Krchnakova et al. 2019). Typically, splice sites are abundant along the transcribed region, while the recognition of those sites is mediated by different serine- and arginine-rich (SR) proteins and hnRNPs (De Conti et al. 2013). In HeLa cells, splicing was found to be less efficient and could occur post-transcriptionally, and in some cases the lncRNAs remain unspliced (Tilgner

et al. 2012; Schlackow et al. 2017). Also, intron excision in lncRNAs is slower than protein-coding genes, and in contrast, their exon skipping is higher (Mukherjee et al. 2017). This phenomenon can be partially explained by the existence of longer lncRNA introns relative to protein-coding genes (Krchnakova et al. 2019), or also by the finding that there is less RNA Pol II pausing at transcription start sites (TES) of lncRNAs versus protein-coding genes (Schlackow et al. 2017).

A recently published study has demonstrated that highly expressed genes can undergo post-transcriptional splicing close to the transcription site zone in addition to co-transcriptional splicing which occurs during transcription (Cote et al. 2024). This study revealed that the mobility of the RNA is slower at the transcription site proximal zone than within the nucleoplasm, suggesting that the slow-moving zone is where transcripts are finally spliced. Using RNA FISH applied to the *CCATI* exon and intron sequences revealed *CCATI* unspliced transcripts that were located in the nucleoplasm of the cells. This was observed only in HeLa cells. The existence of unspliced transcripts was not specific to the *CCATI* short isoform but was found for all three isoforms (Supplementary Figure S3a, b). We measured the distance that *CCATI* unspliced transcripts traveled from the transcription site and found that most of the unspliced transcripts were localized near the site of transcription, suggesting that this transcript undergoes post-transcriptional splicing not long after release from the gene. As *CCATI* is highly transcribed, a high number of transcripts should undergo splicing, and perhaps the delay in the splicing events led to the abundance of transcripts at the gene locus and the release of unspliced transcripts from the gene, and therefore to splicing in the nucleoplasm. This assumption is consistent with the data showing that lncRNA exhibits lower splicing rates than coding genes (Mukherjee et al. 2017). It is also possible that the accumulation of many unspliced *CCATI* transcripts on the gene locus form a structure that sequesters the RNAs from the splicing machinery. Only after release into the nucleoplasm can the spliceosome function on these transcripts. This scenario agrees with the observation of unspliced *CCATI* transcripts only in the region close to the gene. As mentioned, lncRNAs are less efficiently spliced since their exons contain fewer putative binding sites for SR proteins, and hence cannot generate the cooperative network of positive signals that is needed for recruiting the spliceosome to splice sites. Therefore, their splicing would be more dependent on how optimal their splice sites were (Krchnakova et al. 2019). However, the removal of introns from enhancer lncRNAs did not change their enhancing activity, meaning that introns are not essential for the activating function of lncRNAs. We find that the majority of *CCATI* transcripts are co-transcriptionally spliced,

while the minority are post-transcriptionally spliced. This can result from inefficient splicing at the site of transcription, caused by weak splicing factor interactions with the *CCATI* transcripts.

When the cells were treated with splicing inhibitors, there was a significant decrease in *CCATI* accumulation at the transcription site, and surprisingly, unspliced transcripts were observed in the cytoplasm of some cells. This is a rare occurrence, since introns are non-coding sequences that do not reach the cytoplasm. Moreover, when splicing is inhibited, the unspliced transcripts are usually retained at the site of transcription or accumulate in nuclear speckles (Hasenson et al. 2022; Barutcu et al. 2022; Hall et al. 2006; Johnson et al. 2000; Mor et al. 2016). These transcripts do not exit the nucleus since they are detected as aberrant transcripts. The fact that unspliced versions of *CCATI* could be found in the nucleoplasm under normal conditions, and even in the cytoplasm under conditions of splicing inhibition, suggests that the unspliced versions are not defective but are recognized as legitimate variants and might serve roles that await discovery.

Although we do not know to what extent *CCATI*-regulated looping engages in crosstalk with other aspects of *MYC* regulation, this study represents yet another component of the complicated *MYC* region. Our results indicate that the unusual *CCATI* transcription site size at the gene locus is due to the accumulation of *CCATI* transcripts, and not to exceptionally high transcription levels. This may suggest that a function of *CCATI* transcripts on the gene locus is in the maintenance of the long-range chromatin interactions at the *MYC* locus. This shows that the accumulation of *CCATI* at its site of transcription is poised to act as a modulator of gene expression in a locus-specific manner. Finally, as this region also expresses distinct lncRNAs in other types of human cancers, it will be of interest to learn whether other 8q24 lncRNAs behave similarly to *CCATI*.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00418-024-02294-w>.

**Acknowledgements** We acknowledge the help of Rakefet Ben-Yishay and Vera Pavlov.

**Author contributions** C.B., E.V., E.P., G.R. conducted the experiments. N.K. prepared the figures, D.H., A.N., Y.S.T. conceived the project. C.B. and Y.S.T. wrote the paper.

**Funding** Open access funding provided by Bar-Ilan University. This work was supported by the Israel Science Foundation (1278/18; Y. Shav-Tal) and a Bar-Ilan University - Rabin Medical Center Biomedical Research Award. The authors declare no competing financial interests.

**Data availability** The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ahmadiyeh N, Pomerantz MM, Grisanzio C, Herman P, Jia L, Almenro V, He HH, Brown M, Liu XS, Davis M, Caswell JL, Beckwith CA, Hills A, Macconail L, Coetzee GA, Regan MM, Freedman ML (2010) 8q24 prostate, breast, and colon cancer risk loci show tissue-specific long-range interaction with MYC. *Proc Natl Acad Sci U S A* 107(21):9742–9746. <https://doi.org/10.1073/pnas.0910668107>
- Alaiyan B, Ilyayev N, Stojadinovic A, Izadjoo M, Roistacher M, Pavlov V, Tzivin V, Halle D, Pan H, Trink B, Gure AO, Nissan A (2013) Differential expression of colon cancer associated transcript1 (*CCAT1*) along the colonic adenoma-carcinoma sequence. *BMC Cancer* 13:196. <https://doi.org/10.1186/1471-2407-13-196>
- Amjadi-Moheb F, Paniri A, Akhavan-Niaki H (2021) Insights into the links between MYC and 3D chromatin structure and epigenetics regulation: implications for cancer therapy. *Cancer Res* 81(8):1925–1936. <https://doi.org/10.1158/0008-5472.CAN-20-3613>
- Ansari H, Shahriza A, Birgani YT, Birgani MT, Hajjari M, Asl JM (2019) Long noncoding RNAs in colorectal adenocarcinoma; an in silico analysis. *Pathol Oncol Res* 25(4):1387–1394. <https://doi.org/10.1007/s12253-018-0428-2>
- Barutcu AR, Wu M, Braunschweig U, Dyakov BJA, Luo Z, Turner KM, Durbic T, Lin ZY, Weatheritt RJ, Maass PG, Gingras AC, Blencowe BJ (2022) Systematic mapping of nuclear domain-associated transcripts reveals speckles and lamina as hubs of functionally distinct retained introns. *Mol Cell* 82(5):1035–1052 e1039. <https://doi.org/10.1016/j.molcel.2021.12.010>
- Belmont AS (2021) Nuclear compartments: an incomplete primer to nuclear compartments, bodies, and genome organization relative to nuclear architecture. *Cold Spring Harb Perspect Biol*. <https://doi.org/10.1101/cshperspect.a041268>
- Bensaude O (2011) Inhibiting eukaryotic transcription: which compound to choose? how to evaluate its activity? *Transcription* 2(3):103–108. <https://doi.org/10.4161/trns.2.3.16172>
- Bentley DL (2014) Coupling mRNA processing with transcription in time and space. *Nat Rev Genet* 15(3):163–175. <https://doi.org/10.1038/nrg3662>
- Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S (1992) The product of the mouse *Xist* gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* 71(3):515–526. [https://doi.org/10.1016/0092-8674\(92\)90519-i](https://doi.org/10.1016/0092-8674(92)90519-i)
- Brody Y, Neufeld N, Bieberstein N, Cause SZ, Bohnlein EM, Neugebauer KM, Darzacq X, Shav-Tal Y (2011) The in vivo kinetics of RNA polymerase II elongation during co-transcriptional splicing.

- PLoS Biol 9(1):e1000573. <https://doi.org/10.1371/journal.pbio.1000573>
- Brown CJ, Lafreniere RG, Powers VE, Sebastio G, Ballabio A, Pettigrew AL, Ledbetter DH, Levy E, Craig IW, Willard HF (1991) Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature* 349(6304):82–84. <https://doi.org/10.1038/349082a0>
- Calandrelli R, Wen X, Charles Richard JL, Luo Z, Nguyen TC, Chen CJ, Qi Z, Xue S, Chen W, Yan Z, Wu W, Zaleta-Rivera K, Hu R, Yu M, Wang Y, Li W, Ma J, Ren B, Zhong S (2023) Genome-wide analysis of the interplay between chromatin-associated RNA and 3D genome organization in human cells. *Nat Commun* 14(1):6519. <https://doi.org/10.1038/s41467-023-42274-7>
- Campbell M, Izumiya Y (2020) PAN RNA: transcriptional exhaust from a viral engine. *J Biomed Sci* 27(1):41. <https://doi.org/10.1186/s12929-020-00637-y>
- Chen J, Zhang K, Song H, Wang R, Chu X, Chen L (2016) Long noncoding RNA CCAT1 acts as an oncogene and promotes chemoresistance in docetaxel-resistant lung adenocarcinoma cells. *Oncotarget* 7(38):62474–62489. <https://doi.org/10.18632/oncotarget.11518>
- Chen J, Alduais Y, Zhang K, Zhu X, Chen B (2021) CCAT1/FABP5 promotes tumour progression through mediating fatty acid metabolism and stabilizing PI3K/AKT/mTOR signalling in lung adenocarcinoma. *J Cell Mol Med* 25(19):9199–9213. <https://doi.org/10.1111/jcmm.16815>
- Chen Z, Tian D, Chen X, Cheng M, Xie H, Zhao J, Liu J, Fang Z, Zhao B, Bian E (2023) Super-enhancer-driven lncRNA LIMD1-AS1 activated by CDK7 promotes glioma progression. *Cell Death Dis* 14(6):383. <https://doi.org/10.1038/s41419-023-05892-z>
- Clemson CM, McNeil JA, Willard HF, Lawrence JB (1996) XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *J Cell Biol* 132(3):259–275. <https://doi.org/10.1083/jcb.132.3.259>
- Cote A, O'Farrell A, Dardani I, Dunagin M, Cote C, Wan Y, Bayatpour S, Drexler HL, Alexander KA, Chen F, Wassie AT, Patel R, Pham K, Boyden ES, Berger S, Phillips-Cremins J, Churchman LS, Raj A (2024) Post-transcriptional splicing can occur in a slow-moving zone around the gene. *Elife*. <https://doi.org/10.7554/eLife.91357>
- Coulon A, Ferguson ML, de Turrís V, Palangat M, Chow CC, Larson DR (2014) Kinetic competition during the transcription cycle results in stochastic RNA processing. *Elife*. <https://doi.org/10.7554/eLife.03939>
- Custodio N, Carmo-Fonseca M, Geraghty F, Pereira HS, Grosveld F, Antoniou M (1999) Inefficient processing impairs release of RNA from the site of transcription. *Embo J* 18(10):2855–2866. <https://doi.org/10.1093/emboj/18.10.2855>
- Darzacq X, Shav-Tal Y, de Turrís V, Brody Y, Shenoy SM, Phair RD, Singer RH (2007) In vivo dynamics of RNA polymerase II transcription. *Nat Struct Mol Biol* 14(9):796–806
- De Conti L, Baralle M, Buratti E (2013) Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip Rev RNA* 4(1):49–60. <https://doi.org/10.1002/wrna.1140>
- Dekker J, Alber F, Aufmkolk S, Beliveau BJ, Bruneau BG, Belmont AS, Bintu L, Boettiger A, Calandrelli R, Disteche CM, Gilbert DM, Gregor T, Hansen AS, Huang B, Huangfu D, Kalhor R, Leslie CS, Li W, Li Y, Ma J, Noble WS, Park PJ, Phillips-Cremins JE, Pollard KS, Rafelski SM, Ren B, Ruan Y, Shav-Tal Y, Shen Y, Shendure J, Shu X, Strambio-De-Castillia C, Vertii A, Zhang H, Zhong S (2023) Spatial and temporal organization of the genome: Current state and future aims of the 4D nucleome project. *Mol Cell* 83(15):2624–2640. <https://doi.org/10.1016/j.molcel.2023.06.018>
- Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, Dean A, Blobel GA (2012) Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149(6):1233–1244. <https://doi.org/10.1016/j.cell.2012.03.051>
- Dixon JR, Gorkin DU, Ren B (2016) Chromatin domains: the unit of chromosome organization. *Mol Cell* 62(5):668–680. <https://doi.org/10.1016/j.molcel.2016.05.018>
- Effenberger KA, Urabe VK, Jurica MS (2017) Modulating splicing with small molecular inhibitors of the spliceosome. *Wiley Interdiscip Rev RNA*. <https://doi.org/10.1002/wrna.1381>
- Faber GP, Nadav-Eliyahu S, Shav-Tal Y (2022) Nuclear speckles - a driving force in gene expression. *J Cell Sci*. <https://doi.org/10.1242/jcs.259594>
- Ge X, Li GY, Jiang L, Jia L, Zhang Z, Li X, Wang R, Zhou M, Zhou Y, Zeng Z, Xiang J, Li Z (2019) Long noncoding RNA CAR10 promotes lung adenocarcinoma metastasis via miR-203/30/SNAI axis. *Oncogene* 38(16):3061–3076. <https://doi.org/10.1038/s41388-018-0645-x>
- Gil N, Ulitsky I (2020) Regulation of gene expression by cis-acting long non-coding RNAs. *Nat Rev Genet* 21(2):102–117. <https://doi.org/10.1038/s41576-019-0184-5>
- Gordon JM, Phizicky DV, Neugebauer KM (2021) Nuclear mechanisms of gene expression control: pre-mRNA splicing as a life or death decision. *Curr Opin Genet Dev* 67:67–76. <https://doi.org/10.1016/j.gde.2020.11.002>
- Guo CJ, Ma XK, Xing YH, Zheng CC, Xu YF, Shan L, Zhang J, Wang S, Wang Y, Carmichael GG, Yang L, Chen LL (2020) Distinct processing of lncRNAs contributes to non-conserved functions in stem cells. *Cell* 181(3):621–636 e622. <https://doi.org/10.1016/j.cell.2020.03.006>
- Hall LL, Smith KP, Byron M, Lawrence JB (2006) Molecular anatomy of a speckle. *Anat Rec A Discov Mol Cell Evol Biol* 288(7):664–675. <https://doi.org/10.1002/ar.a.20336>
- Hartl P, Gottesfeld J, Forbes DJ (1993) Mitotic repression of transcription in vitro. *J Cell Biol* 120(3):613–624. <https://doi.org/10.1083/jcb.120.3.613>
- Hasenson SE, Alkalay E, Atrash MK, Boucholez A, Gershbaum J, Hochberg-Laufer H, Shav-Tal Y (2022) The association of MEG3 lncRNA with nuclear speckles in living cells. *Cells* 11(12):1942. <https://doi.org/10.3390/cells11121942>
- Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, Sigova AA, Hoke HA, Young RA (2013) Super-enhancers in the control of cell identity and disease. *Cell* 155(4):934–947. <https://doi.org/10.1016/j.cell.2013.09.053>
- Huppi K, Pitt JJ, Wahlberg BM, Caplen NJ (2012) The 8q24 gene desert: an oasis of non-coding transcriptional activity. *Front Genet* 3:69. <https://doi.org/10.3389/fgene.2012.00069>
- Iaccarino I (2017) lncRNAs and MYC: An Intricate Relationship. *Int J Mol Sci* 18(7):1497. <https://doi.org/10.3390/ijms18071497>
- Jia Q, Chen S, Tan Y, Li Y, Tang F (2020) Oncogenic super-enhancer formation in tumorigenesis and its molecular mechanisms. *Exp Mol Med* 52(5):713–723. <https://doi.org/10.1038/s12276-020-0428-7>
- Jin K, Wang S, Zhang Y, Xia M, Mo Y, Li X, Li G, Zeng Z, Xiong W, He Y (2019) Long non-coding RNA PVT1 interacts with MYC and its downstream molecules to synergistically promote tumorigenesis. *Cell Mol Life Sci* 76(21):4275–4289. <https://doi.org/10.1007/s00018-019-03222-1>
- Johnson C, Primorac D, McKinstry M, McNeil J, Rowe D, Lawrence JB (2000) Tracking COL1A1 RNA in osteogenesis imperfecta. splice-defective transcripts initiate transport from the gene but are retained within the SC35 domain. *J Cell Biol* 150(3):417–432
- Kam Y, Rubinstein A, Naik S, Djavarsarav I, Halle D, Ariel I, Gure AO, Stojadinovic A, Pan H, Tsvivin V, Nissan A, Yavin E (2014) Detection of a long non-coding RNA (CCAT1) in living cells and human adenocarcinoma of colon tissues using FIT-PNA molecular beacons. *Cancer Lett* 352(1):90–96. <https://doi.org/10.1016/j.canlet.2013.02.014>

- Kim J, Venkata NC, Hernandez Gonzalez GA, Khanna N, Belmont AS (2020) Gene expression amplification by nuclear speckle association. *J Cell Biol*. <https://doi.org/10.1083/jcb.201904046>
- Kong R, Zhang EB, Yin DD, You LH, Xu TP, Chen WM, Xia R, Wan L, Sun M, Wang ZX, De W, Zhang ZH (2015) Long non-coding RNA *PVT1* indicates a poor prognosis of gastric cancer and promotes cell proliferation through epigenetically regulating p15 and p16. *Mol Cancer* 14:82. <https://doi.org/10.1186/s12943-015-0355-8>
- Kopp F, Mendell JT (2018) Functional classification and experimental dissection of long noncoding RNAs. *Cell* 172(3):393–407. <https://doi.org/10.1016/j.cell.2018.01.011>
- Krchnakova Z, Thakur PK, Krausova M, Bieberstein N, Haberman N, Muller-McNicoll M, Stanek D (2019) Splicing of long non-coding RNAs primarily depends on polypyrimidine tract and 5' splice-site sequences due to weak interactions with SR proteins. *Nucleic Acids Res* 47(2):911–928. <https://doi.org/10.1093/nar/gky1147>
- Lai Y, Chen Y, Lin Y, Ye L (2018) Down-regulation of LncRNA CCAT1 enhances radiosensitivity via regulating miR-148b in breast cancer. *Cell Biol Int* 42(2):227–236. <https://doi.org/10.1002/cbin.10890>
- Lancho O, Herranz D (2018) The MYC enhancer-ome: long-range transcriptional regulation of MYC in cancer. *Trends Cancer* 4(12):810–822. <https://doi.org/10.1016/j.trecan.2018.10.003>
- Landry JJ, Pyl PT, Rausch T, Zichner T, Tekkedil MM, Stutz AM, Jauch A, Aiyar RS, Pau G, Delhomme N, Gagneur J, Korbel JO, Huber W, Steinmetz LM (2013) The genomic and transcriptomic landscape of a HeLa cell line. *Genes Genom Genetics* 3(8):1213–1224. <https://doi.org/10.1534/g3.113.005777>
- Lee JT (2012) Epigenetic regulation by long noncoding RNAs. *Science* 338(6113):1435–1439. <https://doi.org/10.1126/science.1231776>
- Lee JH, Xiong F, Li W (2020) Enhancer RNAs in cancer: regulation, mechanisms and therapeutic potential. *RNA Biol* 17(11):1550–1559. <https://doi.org/10.1080/15476286.2020.1712895>
- Li Z, Yuan J, Da Q, Yan Z, Qu J, Li D, Liu X, Zhan Q, Liu J (2023) Long non-coding RNA colon cancer-associated transcript 1-Vimentin axis promoting the migration and invasion of HeLa cells. *Chin Med J (engl)* 136(19):2351–2361. <https://doi.org/10.1097/CM9.0000000000002373>
- Li L, Chen J, Wang A, Yi K (2024) ALKBH5 regulates ovarian cancer growth via demethylating long noncoding RNA *PVT1* in ovarian cancer. *J Cell Mol Med* 28(2):e18066. <https://doi.org/10.1111/jcmm.18066>
- Lin CY, Loven J, Rahl PB, Paranal RM, Burge CB, Bradner JE, Lee TI, Young RA (2012) Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* 151(1):56–67. <https://doi.org/10.1016/j.cell.2012.08.026>
- Lin X, Spindler TJ, de Souza Fonseca MA, Corona RI, Seo JH, Dezem FS, Li L, Lee JM, Long HW, Sellers TA, Karlan BY, Noushmehr H, Freedman ML, Gayther SA, Lawrenson K (2019) Super-enhancer-associated LncRNA UCA1 interacts directly with AMOT to activate YAP target genes in epithelial ovarian cancer. *iScience* 17:242–255. <https://doi.org/10.1016/j.isci.2019.06.025>
- Liu E, Liu Z, Zhou Y, Mi R, Wang D (2015) Overexpression of long non-coding RNA *PVT1* in ovarian cancer cells promotes cisplatin resistance by regulating apoptotic pathways. *Int J Clin Exp Med* 8(11):20565–20572
- Liu Z, Chen Q, Hann SS (2019) The functions and oncogenic roles of CCAT1 in human cancer. *Biomed Pharmacother* 115:108943. <https://doi.org/10.1016/j.biopha.2019.108943>
- Liu F, Wang Y, Cao Y, Wu Z, Ma D, Cai J, Sha J, Chen Q (2023) Transcription factor B-MYB activates lncRNA CCAT1 and upregulates SOCS3 to promote chemoresistance in colorectal cancer. *Chem Biol Interact* 374:110412. <https://doi.org/10.1016/j.cbi.2023.110412>
- Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, Young RA (2013) Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* 153(2):320–334. <https://doi.org/10.1016/j.cell.2013.03.036>
- McClelland ML, Mesh K, Lorenzana E, Chopra VS, Segal E, Watanabe C, Haley B, Mayba O, Yaylaoglu M, Gnad F, Firestein R (2016) CCAT1 is an enhancer-templated RNA that predicts BET sensitivity in colorectal cancer. *J Clin Invest* 126(2):639–652. <https://doi.org/10.1172/JCI83265>
- Mizrahi I, Mazeh H, Grinbaum R, Beglaibter N, Wilschanski M, Pavlov V, Adileh M, Stojadinovic A, Avital I, Gure AO, Halle D, Nissan A (2015) Colon cancer associated transcript-1 (CCAT1) expression in adenocarcinoma of the stomach. *J Cancer* 6(2):105–110. <https://doi.org/10.7150/jca.10568>
- Mor A, Suliman S, Ben-Yishay R, Yunger S, Brody Y, Shav-Tal Y (2010) Dynamics of single mRNP nucleocytoplasmic transport and export through the nuclear pore in living cells. *Nat Cell Biol* 12(6):543–552. <https://doi.org/10.1038/ncb2056>
- Mor A, White A, Zhang K, Thompson M, Esparza M, Munoz-Moreno R, Koide K, Lynch KW, Garcia-Sastre A, Fontoura BM (2016) Influenza virus mRNA trafficking through host nuclear speckles. *Nat Microbiol* 1(7):16069. <https://doi.org/10.1038/nmicrobiol.2016.69>
- Mukherjee N, Calviello L, Hirsekorn A, de Pretis S, Pelizzola M, Ohler U (2017) Integrative classification of human coding and noncoding genes through RNA metabolism profiles. *Nat Struct Mol Biol* 24(1):86–96. <https://doi.org/10.1038/nsmb.3325>
- Nissan A, Stojadinovic A, Mitrani-Rosenbaum S, Halle D, Grinbaum R, Roistacher M, Bochem A, Dayanc BE, Ritter G, Gomceli I, Bostanci EB, Akoglu M, Chen YT, Old LJ, Gure AO (2012) Colon cancer associated transcript-1: a novel RNA expressed in malignant and pre-malignant human tissues. *Int J Cancer J Int Du Cancer* 130(7):1598–1606. <https://doi.org/10.1002/ijc.26170>
- Palozola KC, Donahue G, Liu H, Grant GR, Becker JS, Cote A, Yu H, Raj A, Zaret KS (2017) Mitotic transcription and waves of gene reactivation during mitotic exit. *Science* 358(6359):119–122. <https://doi.org/10.1126/science.aal4671>
- Parolia A, Cieslik M, Chinnaiyan AM (2018) Competing for enhancers: *PVT1* fine-tunes MYC expression. *Cell Res* 28(8):785–786. <https://doi.org/10.1038/s41422-018-0064-0>
- Peng L, Jiang B, Yuan X, Qiu Y, Peng J, Huang Y, Zhang C, Zhang Y, Lin Z, Li J, Yao W, Deng W, Zhang Y, Meng M, Pan X, Li C, Yin D, Bi X, Li G, Lin DC (2019) Super-enhancer-associated long noncoding RNA *HCCL5* is activated by ZEB1 and promotes the malignancy of hepatocellular carcinoma. *Cancer Res* 79(3):572–584. <https://doi.org/10.1158/0008-5472.CAN-18-0367>
- Pombo A, Dillon N (2015) Three-dimensional genome architecture: players and mechanisms. *Nat Rev Mol Cell Biol* 16(4):245–257. <https://doi.org/10.1038/nrm3965>
- Pott S, Lieb JD (2015) What are super-enhancers? *Nat Genet* 47(1):8–12. <https://doi.org/10.1038/ng.3167>
- Rino J, Carvalho T, Braga J, Desterro JM, Luhrmann R, Carmo-Fonseca M (2007) A stochastic view of spliceosome assembly and recycling in the nucleus. *PLoS Comput Biol* 3(10):2019–2031. <https://doi.org/10.1371/journal.pcbi.0030201>
- Sanchez-Escabias E, Guerrero-Martinez JA, Reyes JC (2022) Co-transcriptional splicing efficiency is a gene-specific feature that can be regulated by TGFbeta. *Commun Biol* 5(1):277. <https://doi.org/10.1038/s42003-022-03224-z>
- Schlackow M, Nojima T, Gomes T, Dhir A, Carmo-Fonseca M, Proudfoot NJ (2017) Distinctive patterns of transcription and RNA processing for human lincRNAs. *Mol Cell* 65(1):25–38. <https://doi.org/10.1016/j.molcel.2016.11.029>
- See YX, Chen K, Fullwood MJ (2022) MYC overexpression leads to increased chromatin interactions at super-enhancers and MYC



- binding sites. *Genome Res* 32(4):629–642. <https://doi.org/10.1101/gr.276313.121>
- Shang A, Wang W, Gu C, Chen W, Lu W, Sun Z, Li D (2020) Long non-coding RNA CCAT1 promotes colorectal cancer progression by regulating miR-181a-5p expression. *Aging (Albany NY)* 12(9):8301–8320. <https://doi.org/10.18632/aging.103139>
- Shigeyasu K, Toden S, Ozawa T, Matsuyama T, Nagasaka T, Ishikawa T, Sahoo D, Ghosh P, Uetake H, Fujiwara T, Goel A (2020) The PVT1 lncRNA is a novel epigenetic enhancer of MYC, and a promising risk-stratification biomarker in colorectal cancer. *Mol Cancer* 19(1):155. <https://doi.org/10.1186/s12943-020-01277-4>
- Spector DL, Lamond AI (2011) Nuclear speckles. *Cold Spring Harb Perspect Biol*. <https://doi.org/10.1101/cshperspect.a000646>
- Tang F, Yang Z, Tan Y, Li Y (2020) Super-enhancer function and its application in cancer targeted therapy. *NPJ Precis Oncol* 4:2. <https://doi.org/10.1038/s41698-020-0108-z>
- Tilgner H, Knowles DG, Johnson R, Davis CA, Chakraborty S, Djebali S, Curado J, Snyder M, Gingeras TR, Guigo R (2012) Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res* 22(9):1616–1625. <https://doi.org/10.1101/gr.134445.111>
- Tsanov N, Samacoits A, Chouaib R, Traboulsi AM, Gostan T, Weber C, Zimmer C, Zibara K, Walter T, Peter M, Bertrand E, Mueller F (2016) smiFISH and FISH-quant - a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res* 44(22):e165. <https://doi.org/10.1093/nar/gkw784>
- Wang N, Yu Y, Xu B, Zhang M, Li Q, Miao L (2019) Pivotal prognostic and diagnostic role of the long non-coding RNA colon cancer-associated transcript 1 expression in human cancer (Review). *Mol Med Rep* 19(2):771–782. <https://doi.org/10.3892/mmr.2018.9721>
- Wang DY, Li N, Cui YL (2020a) Long non-coding RNA CCAT1 sponges miR-454 to promote chemoresistance of ovarian cancer cells to cisplatin by regulation of surviving. *Cancer Res Treat* 52(3):798–814. <https://doi.org/10.4143/crt.2019.498>
- Wang Y, Nie H, He X, Liao Z, Zhou Y, Zhou J, Ou C (2020b) The emerging role of super enhancer-derived noncoding RNAs in human cancer. *Theranostics* 10(24):11049–11062. <https://doi.org/10.7150/thno.49168>
- Wang M, Li Q, Liu L (2023) Factors and methods for the detection of gene expression regulation. *Biomolecules* 13(2):304. <https://doi.org/10.3390/biom13020304>
- White NM, Cabanski CR, Silva-Fisher JM, Dang HX, Govindan R, Maher CA (2014) Transcriptome sequencing reveals altered long intergenic non-coding RNAs in lung cancer. *Genome Biol* 15(8):429. <https://doi.org/10.1186/s13059-014-0429-8>
- Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, Young RA (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* 153(2):307–319. <https://doi.org/10.1016/j.cell.2013.03.035>
- Willemin A, Szabo D, Pombo A (2024) Epigenetic regulatory layers in the 3D nucleus. *Mol Cell* 84(3):415–428. <https://doi.org/10.1016/j.molcel.2023.12.032>
- Xiang JF, Yin QF, Chen T, Zhang Y, Zhang XO, Wu Z, Zhang S, Wang HB, Ge J, Lu X, Yang L, Chen LL (2014) Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res* 24(5):513–531. <https://doi.org/10.1038/cr.2014.35>
- Xiao K, Dong Z, Wang D, Liu M, Ding J, Chen W, Shang Z, Yue C, Zhang Y (2021) Clinical value of lncRNA CCAT1 in serum extracellular vesicles as a potential biomarker for gastric cancer. *Oncol Lett* 21(6):447. <https://doi.org/10.3892/ol.2021.12708>
- Xie JJ, Jiang YY, Jiang Y, Li CQ, Lim MC, An O, Mayakonda A, Ding LW, Long L, Sun C, Lin LH, Chen L, Wu JY, Wu ZY, Cao Q, Fang WK, Yang W, Soukiasian H, Meltzer SJ, Yang H, Fullwood M, Xu LY, Li EM, Lin DC, Koeffler HP (2018) Super-enhancer-driven long non-coding RNA LINC01503, regulated by TP63, is over-expressed and oncogenic in squamous cell carcinoma. *Gastroenterology* 154(8):2137–2151 e2131. <https://doi.org/10.1053/j.gastro.2018.02.018>
- Yang C, Pan Y, Deng SP (2019) Downregulation of lncRNA CCAT1 enhances 5-fluorouracil sensitivity in human colon cancer cells. *BMC Mol Cell Biol* 20(1):9. <https://doi.org/10.1186/s12860-019-0188-1>
- Ye R, Cao C, Xue Y (2020) Enhancer RNA: biogenesis, function, and regulation. *Essays Biochem* 64(6):883–894. <https://doi.org/10.1042/EBC20200014>
- Zhan DT, Xian HC (2023) Exploring the regulatory role of lncRNA in cancer immunity. *Front Oncol* 13:1191913. <https://doi.org/10.3389/fonc.2023.1191913>
- Zhang R, Li J, Yan X, Jin K, Li W, Liu X, Zhao J, Shang W, Liu Y (2018) Long noncoding RNA plasmacytoma variant translocation 1 (PVT1) promotes colon cancer progression via endogenous sponging miR-26b. *Med Sci Monit* 24:8685–8692. <https://doi.org/10.12659/MSM.910955>

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