bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# A high-content RNAi screen reveals multiple roles for long noncoding RNAs in cell division

Lovorka Stojic[1], Aaron T L Lun[1]*, Patrice Mascalchi[1]*, Christina Ernst[1], Aisling M Redmond[1], Jasmin Mangei[1], Alexis R Barr[2], Vicky Bousgouni[2], Chris Bakal[2], John C Marioni[1,3,4], Duncan T Odom[1]** and Fanni Gergely[1]**

[1] Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK.

[2] Institute of Cancer Research, 237 Fulham Road London SW3 6JB, UK.

[3] European Bioinformatics Institute, European Molecular Biology Laboratory (EMBL-EBI), Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK.

[4] Wellcome Trust Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK

* Equal contributions

** To whom correspondence should be addressed

Fanni.gergely@cruk.cam.ac.uk

Duncan.odom@cruk.cam.ac.uk

Present Address: *Christina Ernst*: School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland. *Aisling M Redmond*: MRC Cancer Unit, Hutchison/MRC Research Centre, University of Cambridge, UK. *Jasmin Mangei:* German Cancer Research Center (DKFZ), Molecular Genetics, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany. *Alexis R Barr:* MRC London

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Institute of Medical Sciences (LMS), Hammersmith Hospital Campus, Du Cane Road, London W12 0NN, UK. *Duncan Odom:* German Cancer Research Center (DKFZ), Functional and Structural Genomics, Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## ABSTRACT

Genome stability relies on proper coordination of mitosis and cytokinesis, where dynamic microtubules capture and faithfully segregate chromosomes into daughter cells. The role of long noncoding RNAs (lncRNAs) in controlling these processes however remains largely unexplored. To identify lncRNAs with mitotic functions, we performed a high-content RNAi imaging screen targeting more than 2,000 human lncRNAs. By investigating major hallmarks of cell division such as chromosome segregation, mitotic duration and cytokinesis, we discovered numerous lncRNAs with functions in each of these processes. The chromatin-associated lncRNA, *linc00899,* was selected for in-depth studies due to the robust mitotic delay observed upon its depletion. Transcriptome analysis of *linc00899*-depleted cells together with gain-of-function and rescue experiments across multiple cell types identified the neuronal microtubule-binding protein, *TPPP/p25,* as a target of *linc00899*. *Linc00899* binds the genomic locus of *TPPP/p25* and suppresses its transcription through a *cis*-acting mechanism. In cells depleted of *linc00899,* the consequent upregulation of *TPPP/p25* alters microtubule dynamics and is necessary and sufficient to delay mitosis. Overall, our comprehensive screen identified several lncRNAs with roles in genome stability and revealed a new lncRNA that controls microtubule behaviour with functional implications beyond cell division.

**Keywords:** long noncoding RNA, transcription, cell division, microtubules, genome stability

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## INTRODUCTION

Long noncoding RNAs (lncRNAs) are defined as RNAs longer than 200 nucleotides that lack functional open reading frames, and represent a major transcriptional output of the mammalian genome [1, 2]. LncRNAs have been implicated in the control of numerous cellular processes including the cell cycle, differentiation, proliferation and apoptosis [3-5] and their deregulation has been associated with human disease including cancer [6, 7]. Several lncRNAs regulate the levels of key cell cycle regulators such as cyclins, cyclin-dependent kinases (CDK), CDK inhibitors and p53 [8, 9]. LncRNAs have been also linked to cell division by regulating the levels of mitotic proteins [10, 11] or by modulating the activity of enzymes involved in DNA replication and cohesion [12]. In addition, lncRNAs can regulate chromosome segregation by controlling kinetochore formation via centromeric transcription [13] or by acting as decoys for RNA binding proteins involved in maintaining genome stability [14, 15]. All of these lncRNA-dependent functions can occur through transcriptional and post-transcriptional gene regulation, chromatin organisation and/or post-translational regulation of protein activity [5]. These mechanisms usually involve lncRNAs establishing interactions with proteins and/or nucleic acids which allows lncRNA-containing complexes to be recruited to specific RNA or DNA targets [16]. Although lncRNAs represent more than 25% of all human genes (GENCODE v24), the biological significance of the majority of lncRNAs remains unknown.

Systematic screens in human cells have identified new protein-coding genes involved in cell survival, cell cycle progression and chromosome segregation [17-20]. Similar loss-of function screens have been performed to identify lncRNAs with new functions. For example, CRISPR interference (CRISPRi) has been used in high-throughput screens to identify lncRNA loci important for cell survival [21] and revealed

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

lncRNAs whose functions were highly cell type-specific. Similar results were obtained in a CRISPR/Cas9 screen targeting lncRNA splice sites [22]. In a recent RNAi screen targeting human cancer-relevant lncRNAs, Nötzold and colleagues used time-lapse microscopy imaging of Hela Kyoto cells and identified 26 lncRNAs linked to cell cycle regulation and cell morphology [23]. However, this screen only studied ~600 lncRNAs in the genome with respect to a limited number of phenotypes.

With an aim to identify lncRNAs with new functions in cell division, we performed a more comprehensive high-content imaging RNAi screen involving the depletion of 2231 lncRNAs in HeLa cells. We developed new image analysis pipelines to quantify a diverse set of mitotic features in fixed cells, and discovered multiple lncRNAs with roles in mitotic progression, chromosome segregation and cytokinesis. We focused on *linc00899*, a hitherto uncharacterised lncRNA that regulates mitotic progression by repressing the transcription of the microtubule-stabilising protein *TPPP/p25*. Our study demonstrates the regulatory function of *linc00899* in mitotic microtubule behaviour and provides a comprehensive imaging data resource for further investigation of the roles of lncRNAs in cell division.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## RESULTS

### High-content phenotypic screen identified new lncRNAs involved in mitotic progression, chromosome segregation and cytokinesis

To identify novel lncRNAs involved in regulating cell division, we performed two consecutive RNAi screens (screen A and B). Briefly, we transfected HeLa cells with the human Lincode small interfering RNA (siRNA) library targeting 2231 lncRNAs (Fig. 1a; Supplementary Table 1) and examined their effects using high-content screening of mitotic phenotypes. Each lncRNA was targeted with a SMARTpool of four different siRNAs. Following 48 hr incubation, cells were fixed and processed for immunostaining and subsequent automated image acquisition and analysis. In screen A, antibodies targeting CEP215 (to label centrosomes), α-tubulin (to label the microtubule cytoskeleton), phalloidin (to label the actin cytoskeleton) and Hoechst (to label nuclei) were used. In screen B (Fig. 1b-d), phospho-histone H3 (PHH3; to specifically label mitotic cells), α-tubulin, γ-tubulin (to label centrosomes) and Hoechst was used. We used these two screens as independent approaches to robustly identify lncRNAs with new functions in mitotic progression, chromosome segregation and cytokinesis.

In each screen, we employed automated image analysis to segment the cells and developed in-house pipelines to quantify defects in each of abovementioned categories upon lncRNA depletion (Supplementary Fig. 1a). First, for defects in mitotic progression, we determined the percentage of mitotic cells (also called mitotic index), because an increase in the mitotic index implies a delay or block in mitotic progression. We performed nuclear segmentation and computed the mitotic index (Supplementary Fig. 1b) where mitotic cells were identified by the presence of mitotic spindle staining (detected by α-tubulin and CEP215) in screen A or by positive PHH3 staining of

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

chromosomes in screen B. Second, for quantification of chromosome segregation defects, a category that includes chromatin bridges and lagging chromatids, we identified anaphase cells based on α-tubulin staining between the separating nuclei, in addition to Hoechst (DNA) signal (Supplementary Fig. 1c). Third, to evaluate defects in the execution of cytokinesis, we segmented the cytoplasm of interphase cells and scored the number of cells with cytokinetic bridges based on α-tubulin staining (Supplementary Fig. 1d).

For each category, we ascertained lncRNAs for which depletion increased the frequency of defects relative to the mean across all lncRNAs in the siRNA library. Negative control siRNA and cells without any siRNA treatment were used as controls, for which we observed no systematic differences in the frequencies of each phenotype (Fig. 1b-d). We identified candidate lncRNAs involved in mitotic progression (Fig. 1a; *linc00899* and *C1QTNF1-AS1*), chromosome segregation (Fig. 1b; *PP7080* and *linc00883*) and cytokinesis (Fig. 1c; *linc00840* and *loc729970*). As a positive control for mitotic progression defects, we used a SMARTpool against the protein-coding gene *Ch-TOG/CKAP5*, whose depletion leads to mitotic delay and increased mitotic index [24] (Fig. 1b). For chromosome segregation, we successfully identified the lncRNA *NORAD* (Fig. 1c), depletion of which increases the rate of chromosome segregation errors [14, 15]. Supplementary Table 2 contains raw data and computed Z-scores for each lncRNA and phenotype.

To confirm our findings, we conducted a validation screen targeting the top 25 lncRNA candidates identified in the initial screens for mitotic progression and cytokinesis (Supplementary Table 3). Depletion of each lncRNA was performed in two biological (and in total eight technical) replicates. For mitotic progression, this screen corroborated the increase in mitotic index following depletion of *linc00899 and*

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*C1QTNF1-AS1* (Supplementary Fig. 2a). Although *loc100289019*-depleted cells also displayed an elevated mitotic index, levels of the lncRNA did not change upon RNAi [25]. For the cytokinesis category, we observed an increase in the number of cells with cytokinetic bridges after *linc00840* depletion and a decrease after *loc729970 depletion*, but neither led to multinucleation (Supplementary Fig. 2b, c). Furthermore, elevated mitotic index and cytokinesis defects were not associated with reduced cell viability for these lncRNAs (Supplementary Fig. 2d). As positive controls, we used *Ch-TOG* and *ECT2* (a key regulator of cytokinesis [26], the depletion of which led to expected phenotypes: an increased number of mitotic and multinucleated cells, respectively (Supplementary Fig. 2a-c).

Mitotic perturbations caused by depletion of the lncRNA candidates were further characterised by time-lapse microscopy imaging to investigate the dynamics of each phenotype. As expected, a marked mitotic delay was observed in HeLa cells depleted of *linc00899* and *C1QTNF1-AS1,* lncRNAs associated with increased mitotic index (Supplementary Fig. 3). Next, we depleted *linc00883* and *PP7080,* lncRNAs with potential functions in chromosome segregation (Supplementary Fig. 4). Using time-lapse microscopy imaging of HeLa Kyoto cells stably expressing histone H2B-mCherry (a chromatin marker) and eGFP-α-tubulin (a microtubule marker) [18], we found that depletion of *linc00883* and *PP7080* increased the rate of chromosome segregation errors to a similar extent as that of *NORAD*. We then depleted *linc00840* and *loc729970* (Supplementary Fig. 5), lncRNAs from the cytokinesis category, and found that knockdown of *linc00840* doubled the time required for cells to cleave the cytokinetic bridge, whereas knockdown of *loc729970* resulted in shorter cytokinesis. Overall, our screen identified novel functions of lncRNAs in the control of cell division, supporting the idea that lncRNAs play an important role in cell cycle progression.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## Evolutionary and molecular characterisation of *linc00899* and *C1QTNF1-AS1*

Two lncRNAs associated with delayed mitotic progression, *linc00899 and C1QTNF1-AS,* were selected for in-depth functional analysis. *Linc00899* and *C1QTNF1-AS1* are spliced and polyadenylated lncRNAs. *Linc00899* (also known as *loc100271722* or *ENSG00000231711*) is a multi-exonic intergenic lncRNA located on chromosome 22 and is ~1.6kb long. *C1QTNF1-AS1* (also known as *ENSG00000265096*) is a lncRNA on chromosome 17 that is ~1kb long and is antisense to a protein-coding gene "C1q And TNF Related 1" (C1QTNF1/CTRP1). Both lncRNAs are annotated in Gencode, show signs of active transcription (Fig. 2a) and have low protein-coding potential (Supplementary Fig. 6a). Although, we did not identify a syntenic ortholog for *linc00899* in the mouse genome, short stretches of conserved regions [27] are present within exon 1 (Supplementary Fig. 6b). This places *linc00899* in a group of lncRNAs with conserved exonic sequences embedded in a rapidly evolving transcript architecture [28]. Based on the syntenic position of protein-coding gene *C1QTNF1*, we found a mouse ortholog for *C1QTNF1-AS1* (*GM11747, ENSMUG000000086514*) that is also antisense to the mouse *C1qtnf1*. Thus, *C1QTNF1-AS1* is a lncRNA that is conserved across mouse and human, while *linc00899* contains short conserved stretches at its 5' end representing possible functional domains [29, 30].

We then validated the expression of both lncRNAs in HeLa cells using a variety of techniques. cDNA generated from polyadenylated RNA was used for rapid amplification of cDNA ends (RACE) to define the locations of the 5' cap and 3' end identifying several isoforms for *linc00899* and *C1QTNF1-AS1* (Supplementary Fig. 6c). The diversity of isoforms for both lncRNAs is consistent with a previous study on

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

lncRNA annotation [27]. Expression data from ENCODE cell lines indicated that *linc00899* and *C1QTNF1-AS1* were present in both the nucleus and cytoplasm of HeLa-S3 cells (Fig. 2b), which we further confirmed by RNA fluorescence *in situ* hybridisation (RNA FISH) (Fig. 2c). qPCR analyses on RNA extracted from different cellular fractions revealed that *linc00899* but not *C1QTNF1-AS1* is associated with chromatin (Fig. 2d). Although some *linc00899* foci remain detectable in mitotic cells (Supplementary Fig. 6d), they do not associate with mitotic chromatin, arguing against a mitotic bookmarking role for these lncRNAs. *Linc00899* and *C1QTNF1-AS1* are estimated to occur, on average, in five and two copies per cell, respectively, and do not show cell cycle-dependency in their expression (Supplementary Fig. 6e, f). In the Genotype-Tissue Expression (GTEx) RNA-seq dataset [31] from human tissue, *C1QTNF1-AS1* was highly expressed in adrenal gland and spleen, while *linc00899* was broadly expressed in most of the tissues, with uterus being the highest (Supplementary Fig. 6g).

## ***Linc0889* and *C1QTNF1-AS1* facilitate timely mitotic progression**

To characterise the mitotic phenotype further, we depleted *linc00899* and *C1QTNF1-AS1* in HeLa and HeLa Kyoto cells and examined the effect on mitotic progression with immunofluorescence and time-lapse microcopy imaging, respectively. Quantification of lncRNA-depleted cells using RNAi revealed an increase in the mitotic index compared to cells treated with control siRNA (Fig. 3a, b). This was confirmed by time-lapse microscopy imaging of HeLa Kyoto cells where depletion of *linc00899* and *C1QTNF1-AS1* resulted in increased mitotic duration (Supplementary Fig. 7a-c). We found that cells treated with control siRNA initiated anaphase onset at 40 min ± 10 min (median ± SD), whereas the cells depleted of *linc00899* and *C1QTNF1-AS1* initiate

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

anaphase onset at 100 ± 173 and 240 ± 146 min, respectively (Fig. 3c). *Linc00899*-depleted cells assembled bipolar spindles and congressed chromosomes with normal kinetics, but exhibited a delay in metaphase to anaphase transition (Fig. 3d upper panels and Supplementary Movies S1, 2). By contrast, *C1QTNF1-AS1*-depleted cells showed a delay in chromosome congression to the metaphase plate (Fig. 3d lower panels and Supplementary Movie S3).

We next asked whether these phenotypes could be recapitulated by a loss-of-function (LOF) method other than RNAi. We used locked nucleic acid (LNA) oligonucleotides as a complementary method to target lncRNA expression due to their higher efficiency in depleting nuclear compared to cytoplasmic lncRNAs [32, 33]. We successfully depleted both *linc00899* and *C1QTNF1-AS1* with two different LNAs and selected the LNA with more efficient depletion (denoted LNA1) for subsequent studies (Fig. 3e). Similar to RNAi, we observed an increase in the mitotic index following LNA1-mediated depletion of *linc00899* and *C1QTNF1-AS1* (Fig. 3f) and confirmed the mitotic delay with time-lapse microcopy imaging of HeLa Kyoto cells (Fig. 3g, h; Supplementary Fig. 7d-f and Supplementary Movies S4-6). Together, these data indicate that *linc00899* and *C1QTNF1-AS1* have biological functions in regulation of mitotic progression, albeit through different mechanisms.

To further confirm the phenotype of *C1QTNF1-AS1* depletion, we inserted a polyadenylation poly(A) signal (pAS) [34] downstream of the *C1QTNF1-AS1* transcriptional starts site (TSS) using CRISPR-Cas9 gene editing. This method allows lncRNAs to be transcribed but terminates them prematurely due to the inserted pAS, preventing the expression of full-length lncRNA transcripts. We obtained four homozygous clones for *C1QTNF1-AS1* with similar knockdown efficiency, all of which displayed mitotic delay similar to RNAi- and LNA-mediated depletion of *C1QTNF1-*

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

*AS1* (Supplementary Fig. 8). These results suggest that the *C1QTNF1-AS1* transcript, and not transcription at the *C1QTNF1-AS1* locus, is required for the regulation of mitotic progression. Similar experiments for *linc00899* were hindered by the presence of over 4 copies of *linc00899* in the HeLa genome (https://cansar.icr.ac.uk/cansar/cell-lines/HELA/copy_number_variation/) preventing the generation of homozygous clones despite the use of two different guide RNAs targeting different regions of *linc00899* (data not shown).

**_Linc0889_-dependent regulation of _TPPP/p25_, a microtubule associated protein, is responsible for mitotic progression**

To reveal transcriptional regulatory functions of *linc00899* and *C1QTNF1-AS1*, we performed RNA sequencing (RNA-seq) of cells following RNAi- and LNA-mediated depletion of each lncRNA. We selected the subset of differentially expressed genes (DEG) that changed in a consistent direction with both LOF methods to minimise any potential method specific off-target effects as shown previously [25]. For *C1QTNF1-AS1*, a single DEG was detected, which was *C1QTNF1-AS1* itself (Fig. 4a) arguing against a transcriptional role for *C1QTNF1-AS1* in mitotic progression. For *linc00899*, we identified eight DEGs in common across the two LOF methods (Fig. 4a), of which four were changing in the same direction with both methods (Fig. 4b). One of these candidates was *TPPP/p25,* a tubulin polymerisation promoting protein with established roles in microtubule dynamics and mitosis [35, 36]. *TPPP* was upregulated upon RNAi and LNA-mediated depletion of *linc00899* (Fig. 4c, d), which we validated at the protein level in asynchronous (Fig. 4e, f) as well as in mitotic cells (Supplementary Fig. 9a).

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

To address whether the mitotic delay in *linc00899*-depleted cells is due to increased *TPPP* levels, we performed single and double knockdowns of *linc00899* and *TPPP* and analysed the mitotic progression (Fig. 4g). Whereas *TPPP* knockdown alone did not affect mitotic timing, depletion of *linc00899* alone in HeLa cells led to a mitotic delay, consistent with results from HeLa Kyoto cells (Fig. 3c, g). However, cells with simultaneous depletion of *linc00899* and *TPPP* progressed through mitosis with near-normal timing, a phenotype observed with both LOF methods. We confirmed that co-depletion of *linc00899* and *TPPP* rescued the previously observed upregulation of the latter with qPCR (Fig. 4h, i). Similar results were obtained with an additional LNA oligonucleotide targeting the first intron of *linc00899* (Supplementary Fig. 9b, c). These data suggest that *linc00899* needs to be depleted at least by 50% to attain *TPPP* upregulation and mitotic delay in HeLa cells (Supplementary Fig. 9d, e).

### *Linc00889* controls microtubule behaviour in mitotic cells

*TPPP* is known to localise to the mitotic spindle, and its overexpression influences microtubule dynamics and stability in mammalian cells [35-38]. Overexpression of *TPPP* increases tubulin acetylation, and also microtubule stability via microtubule-bundling [36]. Thus, we tested whether depletion of *linc00899,* which leads to upregulation of *TPPP*, had a similar effect. Immunofluorescence of *linc00899*-depleted cells using antibodies against α-tubulin (a microtubule marker) and acetylated α-tubulin (a marker of long-lived microtubules) [39] showed a marked increase in acetylated α-tubulin levels (Fig. 5a), consistent with *linc00899*-depleted cells containing more long-lived microtubules. This was confirmed by quantification of acetylated α-tubulin within the mitotic spindle in *linc00899*-depleted cells (Fig. 5b).

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

As *TPPP* influences the growth velocity of microtubules by affecting their stability [36], we examined the localisation of EB1 protein, which specifically associates with growing microtubule plus-ends where it regulates microtubule dynamics [40]. As expected, EB1 staining was apparent at the spindle pole and throughout the spindle in control cells but it was much reduced in intensity upon *linc00899* depletion (Fig. 5c). Importantly, reduced EB1 signal was not due to diminished microtubule levels, because α-tubulin staining of the mitotic spindle was comparable between *linc00899*-depleted and control cells. Quantification of EB1 signal in mitotic cells confirmed the decrease in EB1 levels, indicating that *linc00899* depletion lessens the number of growing microtubule ends, a phenotype consistent with a reduction in microtubule dynamics (Fig. 5d). Taken together, these data demonstrate that by controlling *TPPP* expression levels, *linc00899* limits the number of long-lived microtubules and maintains normal microtubule dynamics in mitotic HeLa cells.

We next asked whether *linc00899*-mediated regulation of *TPPP* occurs in cell lines other than HeLa. RNAi-mediated depletion of *linc00899* resulted in upregulation of *TPPP* in three normal diploid cell lines (hTERT-RPE1, retinal pigment epithelial cells; MCF10A, untransformed breast epithelial cells; HUVEC, primary umbilical vein endothelial cells) where mitosis is ~ 25 min (Fig. 5e, g, i). Elevated *TPPP* levels were accompanied with a mitotic delay from ~8 to 16 min upon *linc00899* knockdown (Fig. 5f, h, j). This delay was smaller compared to a mitotic delay in HeLa cells and most likely is due to the presence of at least 82 chromosomes in HeLa cells that need to be aligned at the metaphase plate, compared to 46 chromosomes present in normal diploid cells. Indeed, the mitotic timing in HeLa cells is at least ~15 min longer than in RPE1, MCF10A and HUVEC cells. Thus, *linc00899* regulates mitotic progression by controlling *TPPP* levels in multiple cell lines.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In addition to its mitotic functions, *TPPP* is crucial for microtubule organisation in the brain. Although TPPP is present in multiple tissues, expression of *TPPP* is highest in the brain in both mouse [41] and human (Supplementary Fig. 10a). In GTEx brain samples, high *TPPP* expression is accompanied by low levels of *linc00899*, while in multiple sclerosis patients *TPPP* and *linc00899* expression were negatively correlated (Supplementary Fig. 10b, c). This suggests that the regulatory relationship observed in our study may be physiologically relevant in human brain tissue and neuropathological diseases.

**Gain-of-function and rescue studies define *linc00899* as a repressor of *TPPP* expression through *cis*-acting mechanism**

*Linc00089* can regulate *TPPP* expression in *cis* (locally) or in *trans* (distally), as shown for multiple lncRNAs [3, 5, 42]. Despite *linc00089* and *TPPP* being located on different chromosomes, a *cis*-acting mechanism whereby lncRNA interacts with its target site by being tethered to its site of synthesis cannot be excluded. To elucidate the mode of action by which *linc00899* regulates *TPPP,* we compared effects of ectopic and endogenous overexpression of *linc00899* on *TPPP* levels. Despite an increase in *linc00899* expression after its ectopic overexpression using the expression plasmid encoding *linc00899* cDNA, no changes in *TPPP* levels were observed in HeLa and RPE1 cells (Fig. 6a). These results argue against a *trans* function of *linc00899* in *TPPP* regulation. Indeed, rescue experiments using expression plasmid after RNAi- or LNA-mediated depletion of *linc00899* failed to reduce *TPPP* levels (Fig. 6b). However, exogenous lncRNA expression is not expected to rescue phenotypes associated with depletion of *cis*-acting chromatin-associated lncRNAs [42-44]. To assess a possible *cis* effect, we employed the CRISPR activation (CRISPRa) system [45], which

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

uses catalytically inactive dCas9 fused to transcriptional activator VP64, and guide RNAs (gRNAs) targeting different regions of the *linc00899* promoter. We observed two-fold overexpression of *linc00899* from its endogenous locus, which led to downregulation of *TPPP* in the normal diploid RPE1 but not in HeLa cells (Fig. 6c). The lack of *TPPP* downregulation in HeLa cells may stem from the presence of multiple *linc00899* and *TPPP* alleles; for instance, CRISPRa may not induce overexpression of all (4 or more) *linc00899* loci. In summary, our data support the presence of a *cis*-acting mechanism for *linc00899* to repress *TPPP* expression.

### ***Linc0889*** **binds and regulates transcription of** ***TPPP***

As shown in Figure 2, *linc00899* is a nuclear and chromatin-enriched lncRNA, raising the possibility that it could directly regulate transcription of *TPPP*. To test this, we performed Cleavage Under Targets and Release Using Nuclease (CUT&RUN) [46] that allows genome-wide profiling of active and repressive histone modifications as well as transcription factors from low cell numbers. We observed an increase in trimethylation of lysine 4 on histone H3 (H3K4me3), a mark of active transcription, at the promoter of *TPPP* after RNAi- and LNA-mediated depletion of *linc00899* (Fig. 7a). Thus, elevated levels of *TPPP* mRNA in *linc00899*-depleted cells is likely to arise from increased transcription at the *TPPP* locus. No significant changes were seen in other active and repressive histone modifications at the *TPPP* locus (Supplementary Fig. 11).

Based on these findings, we further hypothesized that *linc00899* binds to regulatory regions of the *TPPP* locus. Therefore, we performed capture hybridization analysis of RNA targets with sequencing (CHART-seq), a method used to identify lncRNA binding sites on chromatin [47-50]. We first mapped antisense oligonucleotides

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

whose binding is accessible to *linc00899* transcript (Supplementary Fig. 12a) and using the antisense cocktail found ~10-fold enrichment of *linc00899* transcript compared to the control oligonucleotides (Fig. 7b). No enrichment upon pulldown was detected with the negative control transcript *5.8S*. After sequencing the enriched genomic DNA using the CHART-seq protocol [51] we identified prominent *linc00899* binding site in the intron of *TPPP* (Fig. 7c; Supplementary Fig. 12 b-d); however, we were not able to identify regions with sequence complementarity to the *linc00899* transcript (Supplementary Fig. 13). Thus, transcriptional upregulation of *TPPP* is likely to be mediated by *linc00899* binding to the *TPPP* locus through protein interactions. This could occur through 3D proximity-guided localisation mechanism which allows low-abundant lncRNAs, such as *linc00899*, to identify its target genes even on different chromosomes [52]. Such proximity-guided search has been observed for *Firre* and *CISTR-ACT* lncRNAs [53, 54] and could explain how *linc00899,* which is encoded on chromosome 22, may bind and regulate *TPPP* encoded on chromosome 5.

To corroborate that *linc00899* act as a transcriptional repressor of *TPPP*, we performed co-RNA FISH using intronic probes against the premature *TPPP* and the mature *linc00899* transcripts (Fig. 7d). *TPPP* has 8 alleles in the HeLa cell genome (https://cansar.icr.ac.uk/cansar/cell-lines/HELA/copy_number_variation), yet, premature *TPPP* transcripts were detected only in one third of the cells and presented mostly as a single focus. ~3% of cells showed colocalization between mature *linc00899* and premature *TPPP* transcript. Given that mature *linc00899* transcripts can be detected in most cells, this low level of colocalization is consistent with effective *linc00899*-mediated suppression of *TPPP* transcription.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## DISCUSSION

Previous studies have attributed the regulation of the cell cycle primarily to multi-protein networks. Here, we performed a high-content imaging screen to identify new lncRNAs with functions in cell division. Development of in-house image analysis pipelines coupled with targeted validation of lncRNA-induced phenotypes allowed us to quantify the impact of lncRNA depletion on cell division. Among other lncRNAs, this study identified *linc00899* and *C1QTNF1-AS1* as novel lncRNAs involved in the control of mitotic progression.

Our results revealed that *linc00899* controls mitotic progression by regulating *TPPP*, a protein that binds and stabilises the microtubule network at all stages of the cell cycle [35-38, 41, 55]. *TPPP* also binds to and inhibits histone deacetylase 6, an enzyme responsible for tubulin deacetylation. This binding results in increased tubulin acetylation [36], a phenotype also observed upon *linc00899* depletion. Fine-tuning TPPP protein levels seems particularly important for mitosis. Indeed, TPPP overexpression in human cells suppresses microtubule growth velocity and normal microtubule dynamics, thereby impeding timely spindle assembly and cell division [36]. Previous studies have shown that *TPPP* levels are subject to regulation by microRNAs [56] and protein kinases [41, 55, 57]; our study now reveals lncRNA-mediated transcriptional control as a new regulatory layer.

Our mechanistic studies indicate that *linc00899* regulates *TPPP* transcription in *cis.* In particular, this notion is supported by our findings that i) *linc00899* is a nuclear and chromatin-enriched lncRNA, ii) there is no change in *TPPP* levels after ectopic overexpression of *linc00899*, iii) H3K4me3 levels increase at the *TPPP* promoter in *linc00899*-depleted cells, and iv) *linc00899* binds at the *TPPP* genomic locus. It is possible that *linc00899* contributes to the repressive chromatin landscape at the *TPPP*

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

locus by altering local chromatin accessibility as observed with other lncRNAs [58]. The mechanism whereby *linc00899* binds to the *TPPP* genomic locus and represses its transcription remains to be fully defined. Given the limited colocalization between *linc00899* and the premature *TPPP* transcript, it is more likely that *linc00899* uses "proximity-guided search" where transcription site of *linc00899* resides in close spatial proximity of *TPPP* in the nucleus. This would allow *linc00899* to act immediately upon its transcription and suppress *TPPP* expression possibly via interactions with protein complexes such as chromatin regulators [52]. It is also possible that the *linc00899*-mediated regulation of *TPPP* transcription depends on *linc00899* release from the chromatin, with *linc00899* target specificity being guided by the pre-established chromosomal proximity, as shown for lncRNA *A-ROD* [59]. As we did not observe a strong sequence complementarity between the *linc00899* transcript and the *TPPP* DNA sequence, we excluded the possibility of direct binding of *linc00899* to *TPPP* locus. Instead, our data suggest that the *linc00899* function could be mediated through *linc00899*-protein interactions. Further studies will be required to determine the *in vivo* *linc00899*-protein interactome and the relevance of these interactions in *TPPP* transcriptional regulation and cell division.

*TPPP* is crucial for microtubule organization in the brain. In *Drosophila*, *TPPP* mutants (also known as *Ringmaker*) exhibit defects in axonal extension [60]. In mammals it is primarily expressed in oligodendrocytes where it stabilizes microtubule networks, and its depletion in progenitors inhibits oligodendrocyte differentiation [56]. Indeed, *TPPP*-deficient mice display convulsive seizures, consistent with a defect in myelinating oligodendrocytes [61]. In humans, an inverse correlation between *TPPP* and *linc00899* across all tissues is consistent with a regulatory relationship, with the highest expression of *TPPP* (and lowest expression of *linc00899*) being observed in

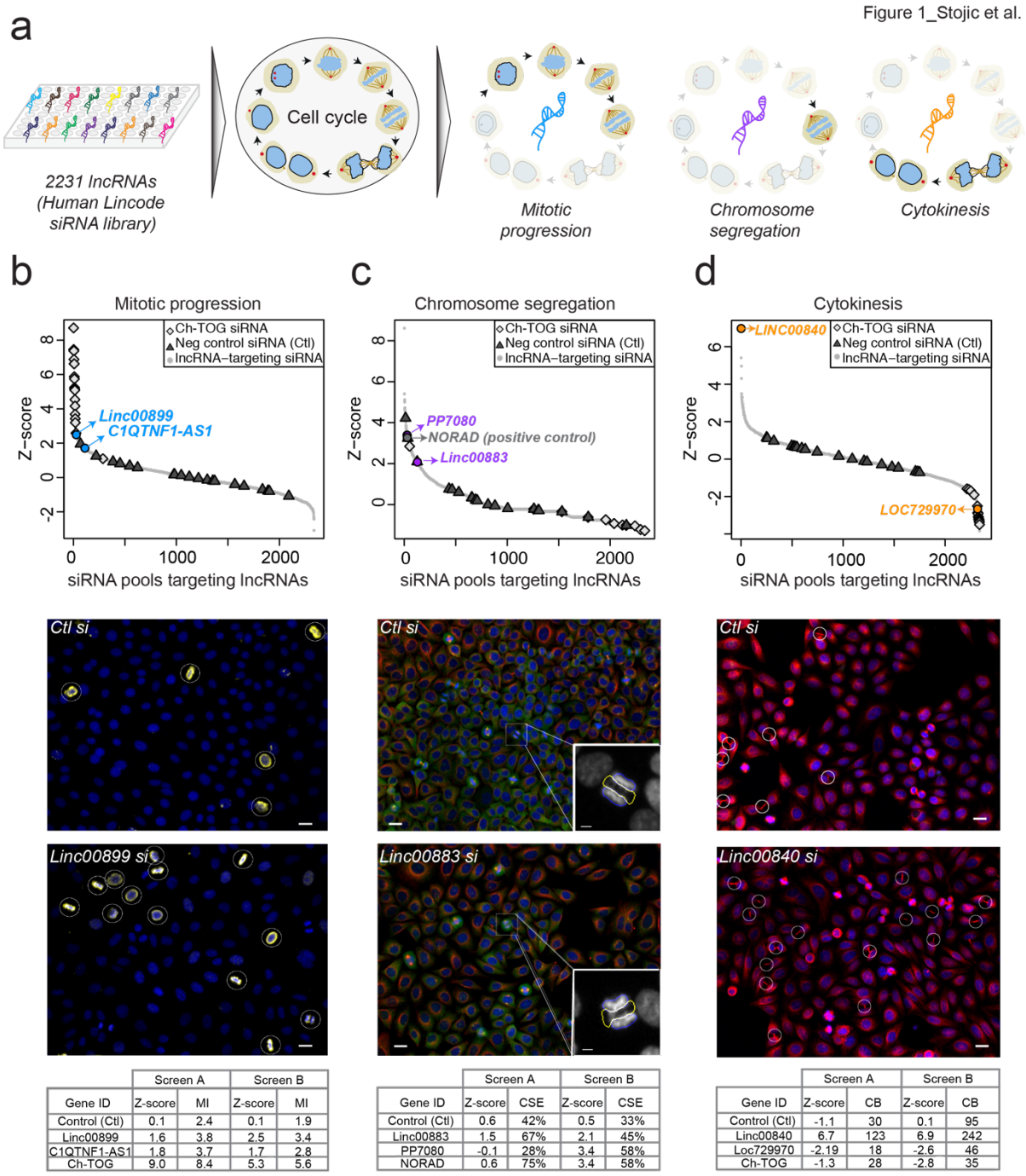
bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the brain. This suggests that *linc00899*-dependent suppression of *TPPP* could be used to fine-tune *TPPP* expression, and hence microtubule behaviour in a developmental stage- and tissue-specific manner. Intriguingly, altered TPPP protein levels have been observed in a number of neurodegenerative disorders including multiple sclerosis and Parkinson's disease [62, 63] .

In this study, we have comprehensively explored the activity of thousands of lncRNAs in cell cycle regulation, and identified an assortment of lncRNAs that are involved in controlling mitotic progression, chromosome segregation and cytokinesis. While our analysis encompassed several cellular features, the imaging data from our screen allows us to extract phenotypes at any stage of cell division upon lncRNA depletion. As interest in the regulatory functions of lncRNAs increases, we anticipate that our data will serve as a powerful resource for prioritizing lncRNA candidates for further studies in the RNA and cell cycle fields.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# FIGURES AND FIGURE LEGENDS




| | Screen A | | Screen B | |
|---|---|---|---|---|
| Gene ID | Z-score | MI | Z-score | MI |
| Control (Ctl) | 0.1 | 2.4 | 0.1 | 1.9 |
| Linc00899 | 1.6 | 3.8 | 2.5 | 3.4 |
| C1QTNF1-AS1 | 1.8 | 3.7 | 1.7 | 2.8 |
| Ch-TOG | 9.0 | 8.4 | 5.3 | 5.6 |

| | Screen A | | Screen B | |
|---|---|---|---|---|
| Gene ID | Z-score | CSE | Z-score | CSE |
| Control (Ctl) | 0.6 | 42% | 0.5 | 33% |
| Linc00883 | 1.5 | 67% | 2.1 | 45% |
| PP7080 | -0.1 | 28% | 3.4 | 58% |
| NORAD | 0.6 | 75% | 3.4 | 58% |

| | Screen A | | Screen B | |
|---|---|---|---|---|
| Gene ID | Z-score | CB | Z-score | CB |
| Control (Ctl) | -1.1 | 30 | 0.1 | 95 |
| Linc00840 | 6.7 | 123 | 6.9 | 242 |
| Loc729970 | -2.19 | 18 | -2.6 | 46 |
| Ch-TOG | -1.3 | 28 | -2.8 | 35 |

**Fig. 1. Identification of lncRNAs involved in regulation of cell division.**

***a.*** Schematic representation of the high-throughput RNAi imaging screen for lncRNAs regulating three mitotic processes: mitotic progression, chromosome segregation and cytokinesis. The screen depleted each of 2231 lncRNAs in HeLa cells using the Human Lincode siRNA library (Dharmacon).

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

***b.*** Z-scores for mitotic progression defects upon depletion of each lncRNA in the siRNA library. Each point corresponds to a single lncRNA where the Z-score was computed based on the mean mitotic index (MI). siRNAs against the protein-coding gene *Ch-TOG* were used as positive controls, in addition to negative control siRNAs (Ctl, from Ambion). Representative images from the top candidate (*linc00899*) are also shown, with PHH3 in yellow indicating mitotic cells (white circles).

***c.*** Z-scores for chromosome segregation defects upon lncRNA depletion, similar to *(**b**).* The Z-score per lncRNA was computed from the mean number of chromosome segregation errors (CSE). Here, *NORAD* (grey) was used as a positive control. Top candidates are highlighted in purple. Representative images from one of the top candidates (*linc00883*) are also shown with staining for α-tubulin (red), PHH3 (green) and γ-tubulin (yellow). Inset depicts normal anaphase cell (blue area) or anaphase cell with CSE (yellow area).

***d.*** Z-scores for cytokinesis defects upon lncRNA depletion, similar to *(**a**).* The Z-score for each lncRNA was computed based on the mean number of cells with cytokinetic bridges (CB). Representative images from the top candidate (*linc00840*) are shown with staining for α-tubulin (red) and DNA (blue). CB are depicted in white circles.

All Z-scores shown here are from screen B. Some of the top candidates are shown in colour and labelled in each plot. The scale bar for all images is 20 μm. Tables below each panel represent the raw data and calculated Z-scores for top lncRNA candidates for each category from two independent screens (A and B).

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

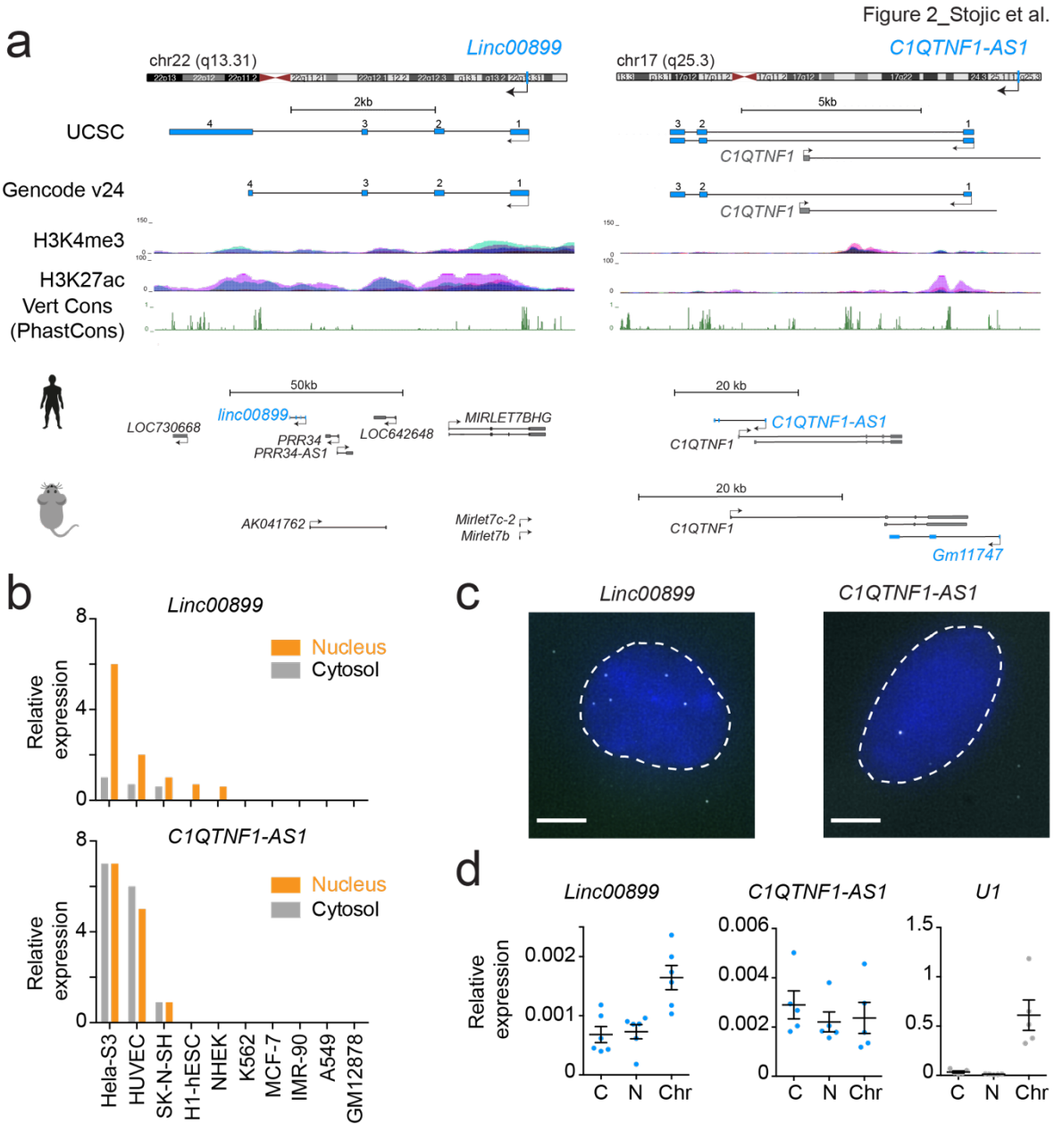



**Fig. 2. Molecular characterisation of the *linc00899* and *C1QTNF1-AS1* lncRNAs.**

***a.*** Schematic representation of the genomic landscape surrounding *linc00899* (annotated in RefSeq as NR_027036; Gencode gene *ENSG00000231711;* chr22:46039907-46044868, hg38) and *C1QTNF1-AS1* (annotated in RefSeq as NR_040018/NR_040019; Gencode gene *ENSG00000265096*; chr17:79019209-79027601, hg38). Marks of active transcription in HeLa cells (H3K4me3 and H3K27ac, obtained from ENCODE via the UCSC browser) and conservation

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

scores by PhastCons are also shown. Putative conserved mouse lncRNAs based on syntenic conservation are shown below.

***b.*** Expression of *linc00899* and *C1QTNF1-AS1* in the nucleus and cytosol of ENCODE cell lines, shown as reads per kilobase of exon per million reads mapped.

***c.*** Single-molecule RNA FISH using exonic probes (white) against *linc00899* and *C1QTNF1-AS1*. Nuclei were stained with DAPI (blue). Scale bar, 5μm.

***d.*** qPCR quantification of lncRNA levels in RNA fractions extracted from different cellular compartments (C, cytoplasm; N, nucleoplasm; Chr, chromatin). *U1* small nuclear RNA was used as a positive control for the chromatin fraction. Errors bars represent the standard error of the mean ±S.E.M from at least five independent experiments.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



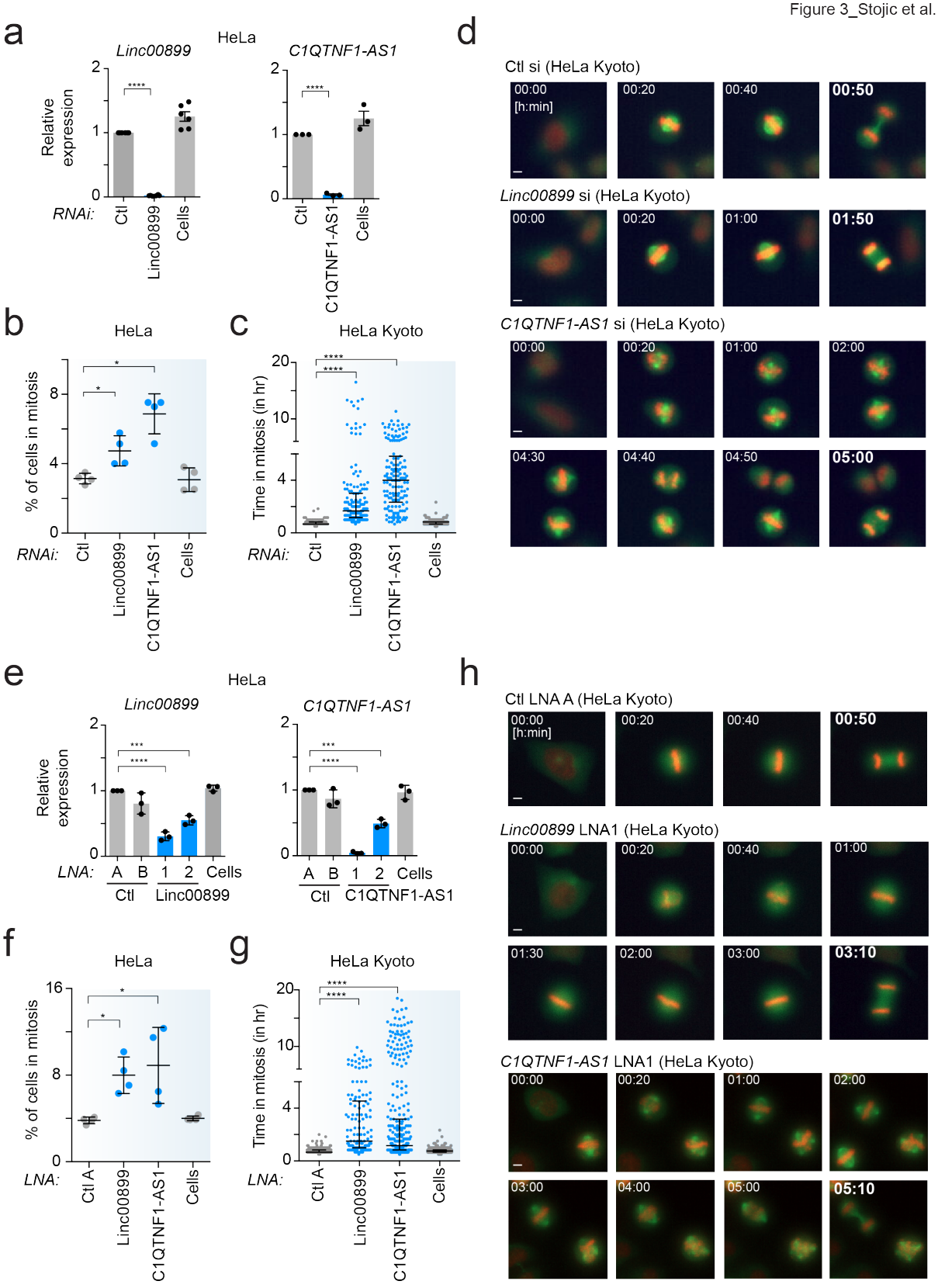



**Fig. 3. *Linc00899* and *C1QTNF1-AS1* regulate mitotic progression in HeLa cells through different mechanisms.**

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

***a.*** Expression of *linc00899* and *C1QTNF1-AS1* after RNAi depletion, as measured by qPCR. Results are also shown for negative control siRNAs (Ctl, from Ambion) and for cells treated with transfection reagent only (Cells). Data are shown as mean ± S.E.M. n = 3 (*C1QTNF1-AS1*) and 6 (*linc00899*) biological replicates, ****$P$<0.0001 by two-tailed Student's *t*-test. Expression values are presented relative to control siRNA.

***b.*** Changes in the mitotic index (MI, based on PHH3 and Hoechst staining) after RNAi-mediated depletion of *linc00899* and *C1QTNF1-AS1* determined 48 hours after siRNA transfection. Data are shown as mean ± S.E.M. n = 4 biological replicates, *$P$<0.05 by two-tailed Student's *t*-test.

***c.*** Quantification of mitotic duration from time-lapse microscopy imaging after depletion of *linc00899* and *C1QTNF1-AS1* in HeLa Kyoto cells. We analysed n=183 for cells treated with negative control siRNAs (Ctl), n=241 for cells with transfection reagent only (Cells), n=163 for *linc00899* siRNAs and n=147 for *C1QTNF1-AS1* siRNAs. Bars show the median and the interquartile range from 3 biological replicates. ****$P$<0.0001 by Mann-Whitney test.

***d.*** Representative still images from time-lapse microscopy from ***(c)***. Scale bar, 5μm.

***e.*** Expression of *linc00899* and *C1QTNF1-AS1* after depletion using two different LNA gapmers targeting the mature lncRNA transcripts, as measured by qPCR. Results are also shown for negative control LNAs (Ctl LNA A and B) and for cells treated with transfection reagent only (Cells). Data are shown as mean ± S.E.M. n = 3 biological replicates, ***$P$<0.001 and ****$P$<0.0001 by two-tailed Student's *t*-test. Expression levels are presented relative to Ctl LNA A.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

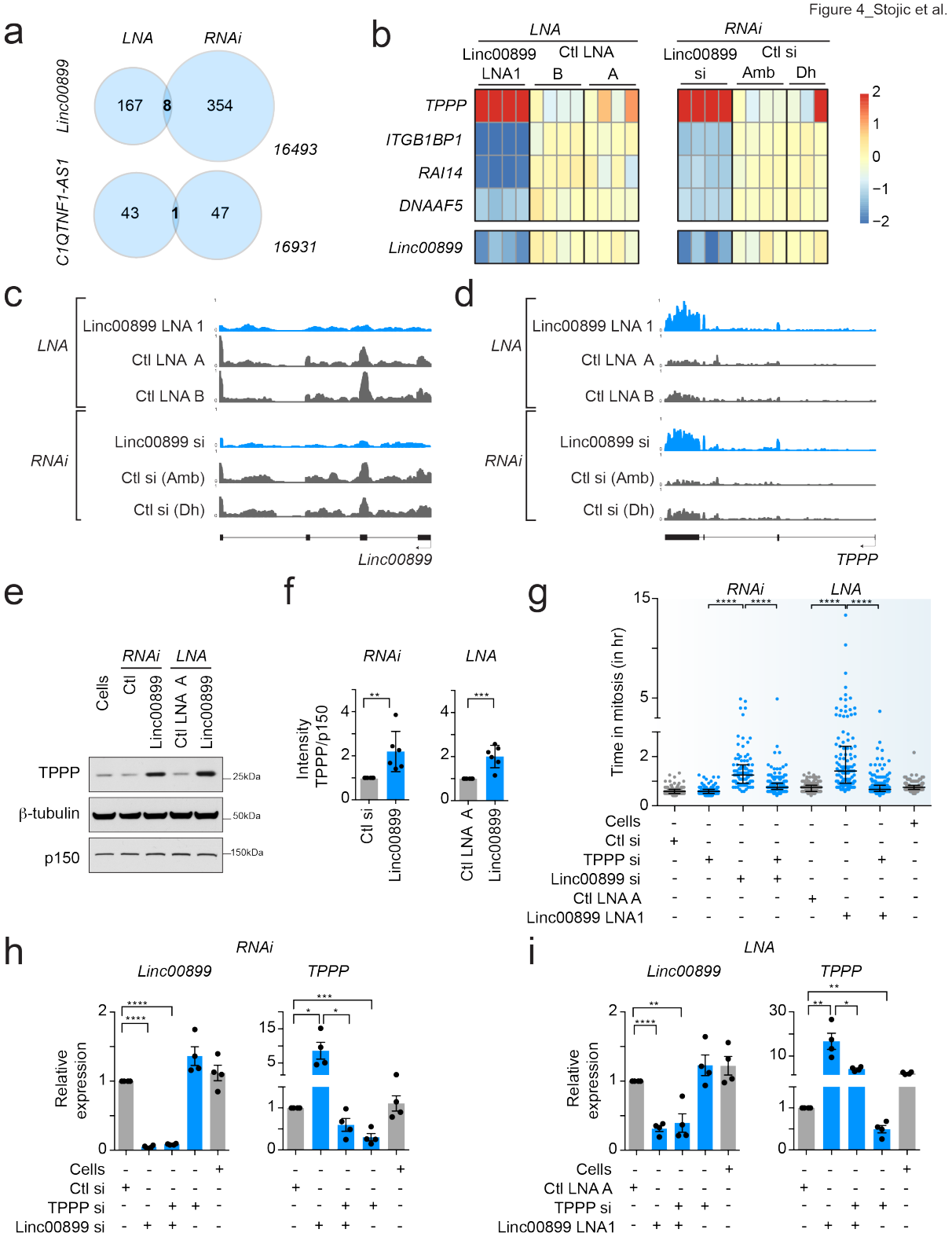
***f.*** Changes in the MI after depletion of *linc00899* and *C1QTNF1-AS1* using LNA1 gapmers, as measured in *(B).* Data are shown as mean ± S.E.M. n = 4 biological replicates, **P*<0.05 by two-tailed Student's *t*-test.

***g.*** Quantification of mitotic duration as in ***(c)*** after depleting *linc00899* and *C1QTNF1-AS1* with LNA1 gapmers. We analysed n=274 for negative control LNA (Ctl LNA A), n=312 for cells with transfection reagent alone (Cells), n=138 for *linc00899* LNA1 and n=349 for *C1QTNF1-AS1* LNA1. Bars show the median and interquartile range from 3 biological replicates. *****P*<0.0001 by Mann-Whitney test.

***h.*** Representative still images from time-lapse microscopy from ***(g)***. Scale bar, 5μm.

Mitotic duration in ***(c)*** and ***(g)*** was defined from NEBD (t=0 min) to anaphase onset.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.




5

**Fig. 4. Identification of *TPPP*, a microtubule-stabilising protein, as a target of *linc00899* in regulation of mitotic progression.**

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

***a.*** Venn diagram of DEGs detected by RNA-seq after depletion of *linc00899* and *C1QTNF1-AS1* using RNAi and LNA1 gapmers. The total number of genes in this analysis was 17022 and DEGs were detected at a FDR of 5% with a log2-fold change threshold of 0.5.

***b.*** Heat map of DEGs detected with RNA-seq after depletion of *linc00899* using RNAi and LNA gapmers. 3 - 4 biological replicates were generated for each condition and compared to appropriate negative controls (Ctl) – A and B for LNA gapmers, Ambion (Amb) and Dharmacon (Dh) for RNAi. Only DEGs changing in the same direction with both LOF methods are shown. *Linc00899* itself is shown as a reference.

***c.*** Genome tracks of RNA-seq coverage for *linc00899* before and after depletion of *linc00899* using RNAi or LNA1 gapmers. The tracks were constructed from averages of three to four biological replicates for each condition, see abbreviations in *(**b**)*.

***d.*** Genome tracks of RNA-seq coverage for *TPPP*, as in *(**c**)*.

***e.*** Representative Western blot of TPPP levels after depletion of *linc00899*. ß-tubulin and p150 were used as two loading controls. Unprocessed original scans of blots are shown in Supplementary Fig. 14.

***f.*** Densitometric analysis of TPPP levels in *(**e**)*. n = 5 biological replicates for each LOF method. Bar charts depict mean ± S.E.M.

***g.*** Quantification of mitotic duration from time-lapse microscopy imaging of HeLa cells after single and double knockdown of *linc00899* and *TPPP* using RNAi or LNAs. We analyzed n=154 for cells treated with transfection reagent alone (Cells), n=188 for negative control siRNA (Ctl, from Ambion), n=208 for *TPPP* siRNAs, n=98 for *linc00899* siRNAs, n=205 for *linc00899* and *TPPP* siRNAs,

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

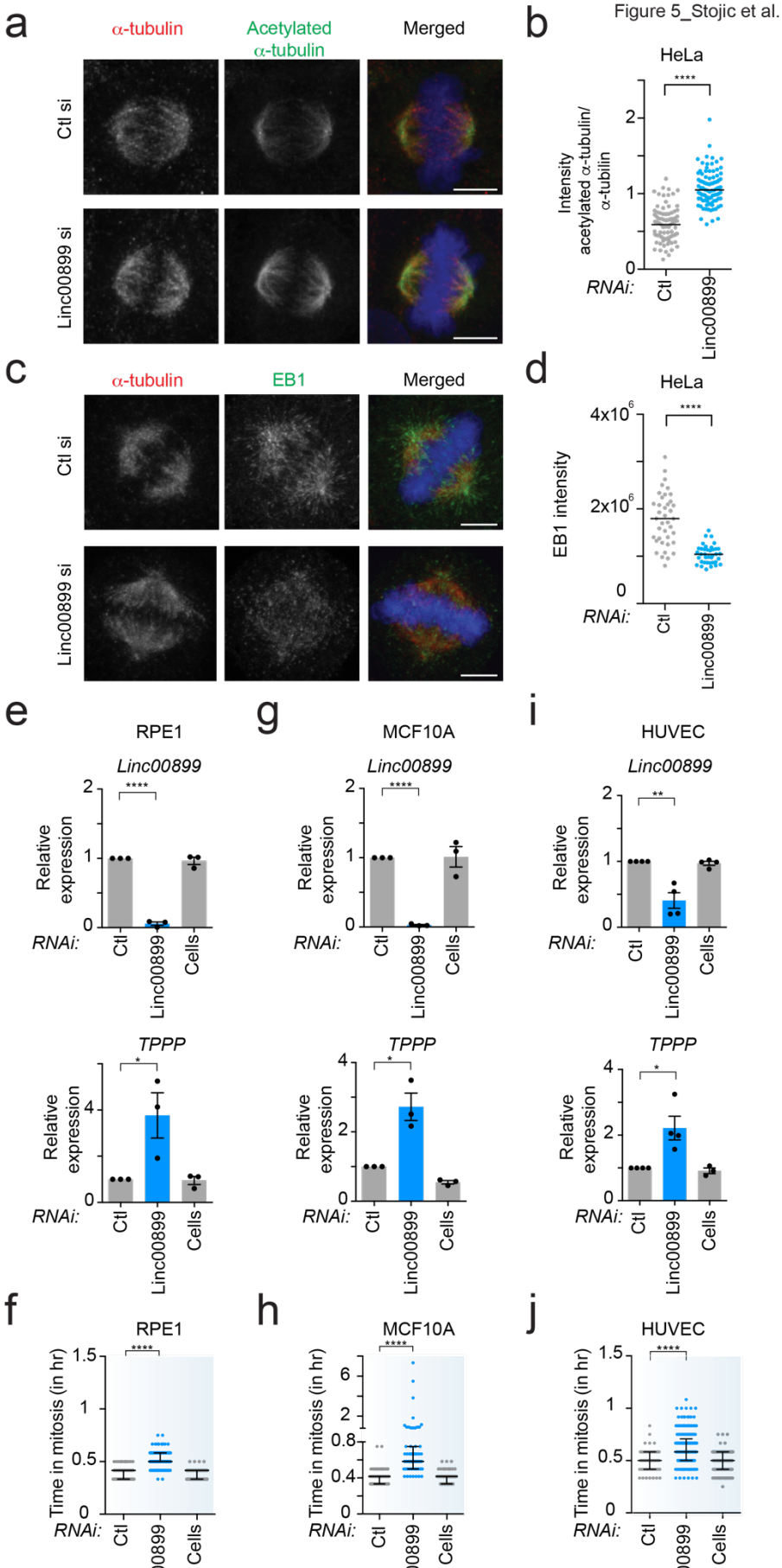
n=183 for negative control LNA (Ctl LNA A), n=116 for *linc00899* LNA1 and n=247 for *linc00899* LNA1 and *TPPP* siRNA. For each condition, we show the median with interquartile range from 2 biological replicates. ****$P$<0.0001 by Mann-Whitney test. Mitotic duration was defined from NEBD (t=0 min) to anaphase onset.

***h.*** Expression of *linc00899* and *TPPP* after single or double knockdown with RNAi, as quantified with qPCR. Results are also shown for cells treated with transfection reagent alone (Cells) or negative control siRNA (Ctl). Data are shown as mean ± S.E.M. n = 3 biological replicates, *$P$<0.05, ***$P$<0.001 and ****$P$<0.0001 by two-tailed Student's *t*-test. Expression values are presented relative to expression with the control siRNA.

***i.*** Expression of *linc00899* and *TPPP* after single or double knockdown as in ***(h)***, but using LNA1 gapmers to deplete *linc000899.* All expression values are presented relative to expression with the negative control LNA A (Ctl LNA A).

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5_Stojic et al.



bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**Fig. 5. Depletion of *linc00899* leads to long-lived mitotic microtubules, mitotic delay and *TPPP* upregulation in different cell lines.**

***a.*** Acetylated α-tubulin levels in HeLa cells after RNAi-mediated *linc00899* depletion or treatment with a negative control siRNA (Ctl, from Ambion), based on immunofluorescence after staining with antibodies against α-tubulin (microtubules, in red) and acetylated α-tubulin (long-lived microtubules, in green). Representative images are shown and correspond to maximum intensity projections of confocal micrographs. DNA was stained with Hoechst and is shown in blue in the merged images.

***b.*** Quantification of acetylated α-tubulin intensity over the total level of α-tubulin from maximum intensity projections obtained in ***(a)***. Numbers of cells analysed is n=85 for negative control siRNAs (Ctl) and n=92 for *linc00899* RNAi. Swarm plots represent values from single mitotic cells, with the median denoted by the horizontal line. ****$P$<0.0001 by two-tailed Student's *t*-test.

***c.*** EB1 signal in HeLa cells after RNAi-mediated *linc00899* depletion or treatment with a negative control siRNA (Ctl), based on immunofluorescence after staining with antibodies against α-tubulin (in red) and EB1 (in green). Images correspond to maximum intensity projections of confocal micrographs. DNA is shown in blue.

***d.*** Quantification of EB1 intensity from maximum intensity projections obtained in the experiment ***(c)***. Numbers of cells analysed is n=39 for negative control siRNAs (Ctl) and n=38 for *linc00899* RNAi. Swarm plots represent values from single mitotic cells, with the median represented by the horizontal line. ****$P$<0.0001 by two-tailed Student's *t*-test.

***e.*** Expression of *linc00899* and *TPPP* after RNAi-mediated depletion of *linc00899* in normal retina cell line (hTert-RPE1) based on qPCR. Results are also shown for

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

negative control siRNAs (Ctl, from Ambion) and cells treated with transfection reagent alone (Cells). Data are shown mean ± S.E.M. n = 3 biological replicates, *$P$<0.05 and ****$P$<0.0001 by two-tailed Student's $t$-test.

***f.*** Quantification of mitotic duration from time-lapse microscopy imaging after depletion of *linc00899* in RPE1 cells. Number of cells analysed is n=135 for negative control siRNAs (Ctl), n=98 for cells with transfection reagent alone (Cells) and n=103 for *linc00899* RNAi. Data are shown as median with interquartile range from 2 biological replicates. ****$P$<0.0001 by Mann-Whitney test. Depletion of *linc00899* in RPE1 leads to an ~12min delay (p<0.0001) compared to cells treated with control siRNA.

***g.*** Expression of *linc00899* and *TPPP* after RNAi-mediated depletion of *linc00899* in non-tumorigenic epithelial breast cells (MCF10A) based on qPCR. Controls are as described in ***(e)***. Data are shown mean ± S.E.M. n = 3 biological replicates, *$P$<0.05 and ****$P$<0.0001 by two-tailed Student's $t$-test.

***h.*** Quantification of mitosis duration from time-lapse microscopy imaging after depletion of *linc00899* in MCF10A cells, as described in ***(f)***. Number of cells analysed is n=143 for negative control siRNAs (Ctl), n=142 for cells with transfection reagent alone (Cells) and n=63 for *linc00899* siRNAs. Data are shown as median with interquartile range from 2 biological replicates. ****$P$<0.0001 by Mann-Whitney test. Depletion of *linc00899* in MCF10A leads to ~16min delay (p< 0.0001) compared to cells treated with control siRNA.

***i.*** Expression of *linc00899* and *TPPP* after RNAi-mediated depletion of *linc00899* in normal primary umbilical vein endothelial cells (HUVEC) based on qPCR. Controls are as described in ***(e)***. Data are shown mean ± S.E.M. n = 4 biological replicates, *$P$<0.05 and **$P$<0.01 by two-tailed Student's $t$-test.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

***j.*** Quantification of mitotic duration from time-lapse microscopy imaging after depletion of *linc00899* in HUVEC cells, as described in ***(f).*** Number of cells analysed is n=59 for negative control siRNAs (Ctl), n=129 for cells with transfection reagent alone (Cells) and n=181 for *linc00899* RNAi. Data are shown as median with interquartile range from 3 biological replicates. ****$P$<0.0001 by Mann-Whitney test. Depletion of *linc00899* in HUVEC leads to an ~8min delay (p=0.0002) compared to cells treated with control siRNA.

Mitotic duration in ***(f)***, ***(h)*** and ***(j)*** was defined from NEBD (t=0 min) to anaphase using bright-field microscopy. Scale bar, 5μm.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 6_Stojic et al.

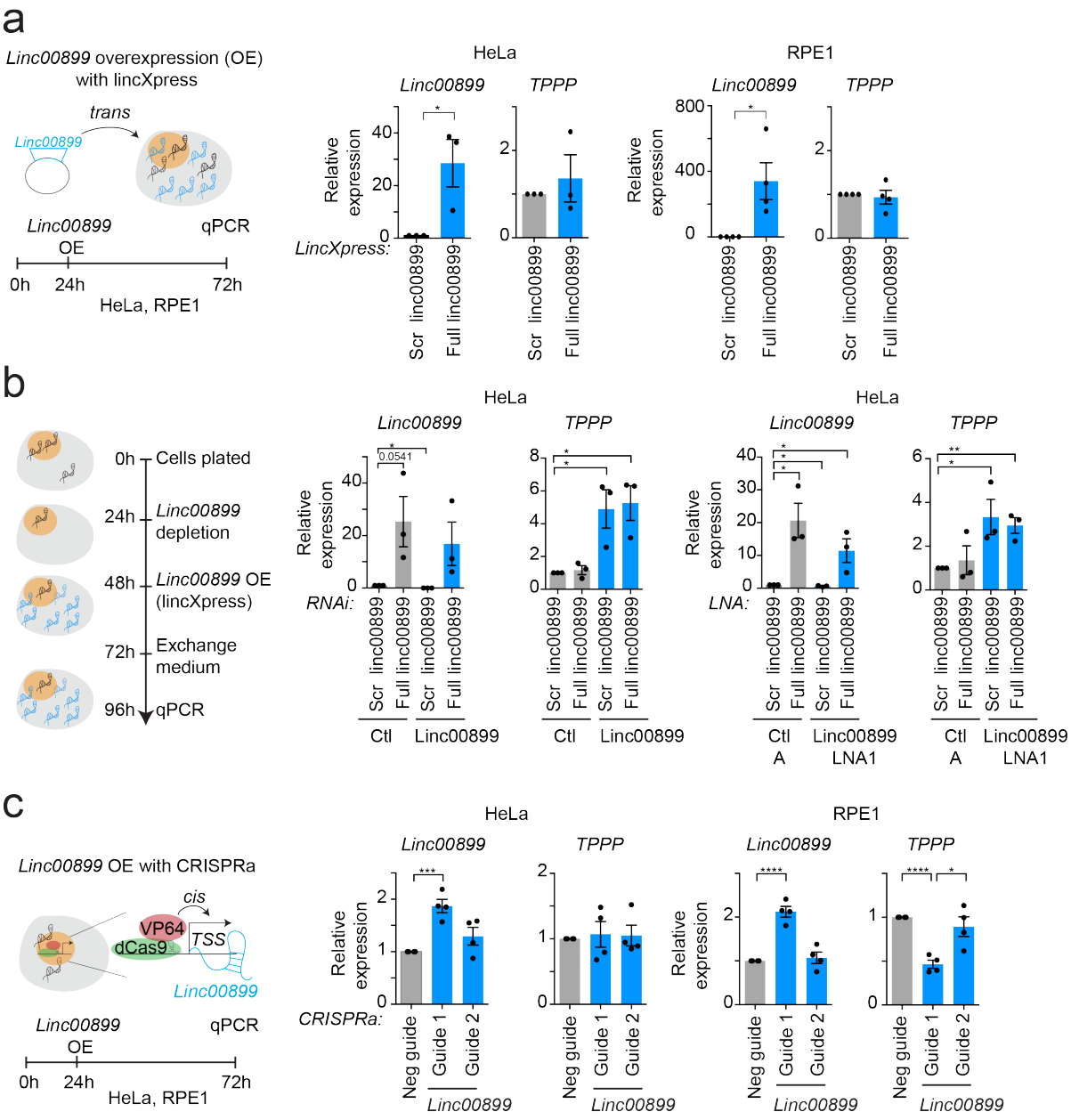



**Fig. 6. *Linc00899* represses *TPPP* expression in *cis*.**

***a.*** Schematic diagram of ectopic overexpression of *linc00899. Linc00899* and *TPPP* expression were analysed by qPCR after lentiviral overexpression using lincXpress vector encoding *linc00899* cDNA in HeLa (*left*) and RPE1 cells (*right*). The expression was normalised to the scrambled *linc00899* vector (negative

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

control). Data are shown mean ± S.E.M. n = 3 - 4 biological replicates, **P*<0.05 by two-tailed Student's *t*-test.

***b.*** Rescue of *linc00899* function after RNAi- or LNA-mediated depletion of *linc00899* in HeLa cells through ectopic overexpression of *linc00899*. Data are shown mean ± S.E.M. n = 3 biological replicates, **P*<0.05 and ***P*<0.01 by two-tailed Student's *t*-test.

***c.*** Schematic diagram of endogenous overexpression of *linc00899* using CRISPRa. *Linc00899* and *TPPP* expression were analysed by qPCR after transduction of dCas9-VP64 and gRNAs targeting different regions of the *linc00899* promoter in HeLa (*left*) and RPE1 cells (*right*). The expression was normalised to the negative guide (negative control). Data are shown mean ± S.E.M. n = 4 biological replicates, **P*<0.05, ****P*<0.001 and *****P*<0.0001 by two-tailed Student's *t*-test.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 7_Stojic et al.

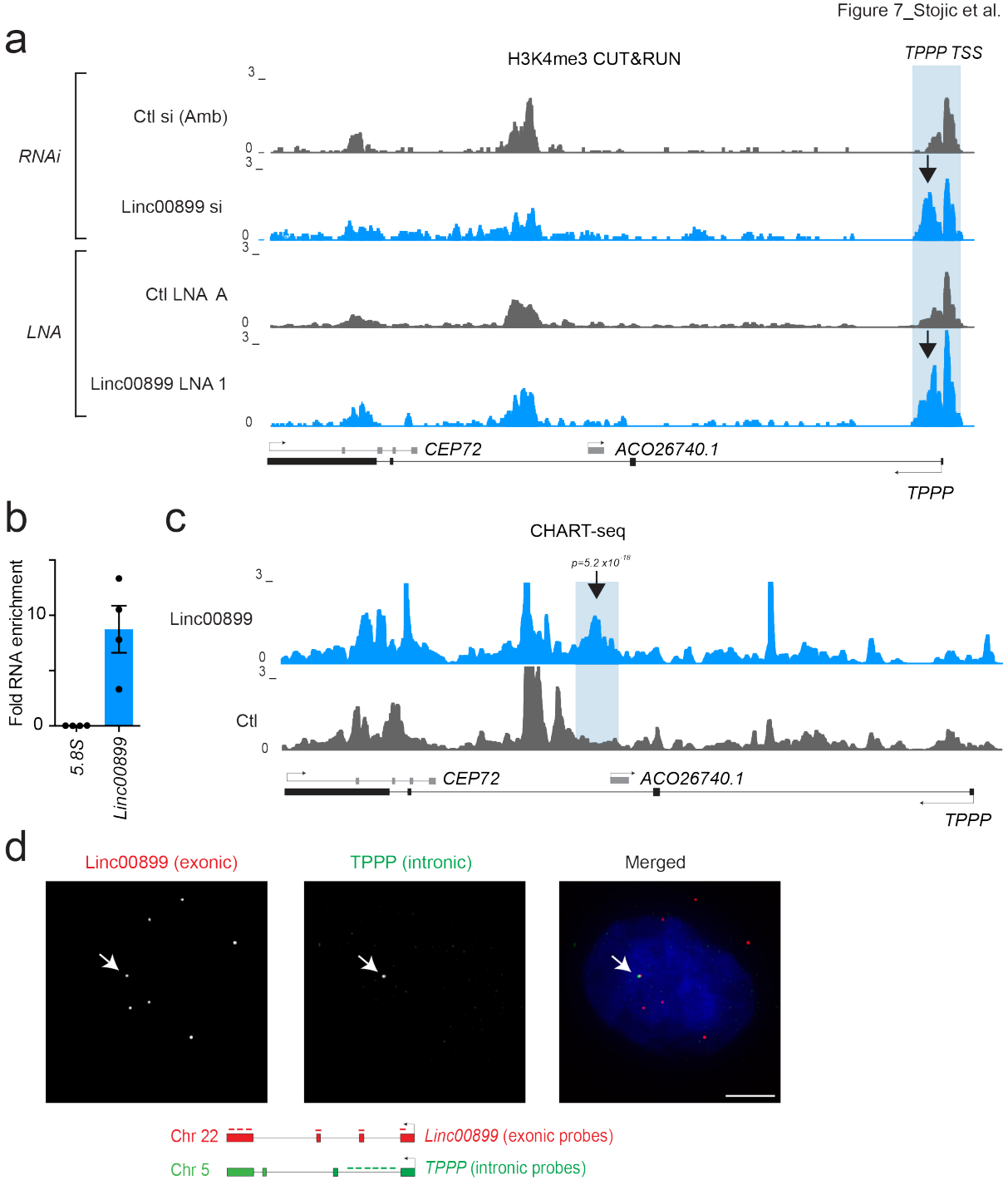



**Fig. 7. *Linc00899* binds to and represses transcription of *TPPP.***

***a.*** CUT&RUN profiling of H3K4me3 enrichment after depletion of *linc00899* with RNAi and LNAs at the *TPPP* locus in HeLa cells. An increase in H3K4me3 was observed at the *TPPP* promoter upon depletion of *linc00899* with both LOF methods

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(arrow). Tracks were constructed from averages of two biological replicates for each condition. TSS, transcriptional start site.

***b.*** Enrichment of *linc00899* using a cocktail of 4 oligonucleotides complementary to accessible regions of *linc00899* (as determined by RNase H mapping, Supplementary Fig. 12a), compared to a mixture of two control oligonucleotides containing parts of the *linc00899* sense sequence, as measured by qPCR. The *5.8S* was used as a negative control region. Data are shown mean ± S.E.M. n = 4 biological replicates.

***c.*** Profiling of *linc00899* binding by CHART-seq after RNase H elution at the *TPPP* locus in HeLa cells. The region with the most significant increase in *linc00899* binding compared to the sense control was present in the second intron of *TPPP* (arrow). All binding sites were detected using an empirical FDR threshold of 30%.

***d.*** Validation of *linc00899* colocalization with the *TPPP* locus by co-RNA FISH in HeLa cells. *Linc00899* (exonic probes against the mature transcript) colocalized to the genomic region of *TPPP* (intronic probes against the premature transcript) in ~3% of the cells (n=376). The nucleus was stained with DAPI (blue). Scale bar, 5 μm.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# METHODS

## Cell lines

HeLa and 293FT cells were maintained in Dulbecco's modified Eagle's medium (Gibco, 41966-029) supplemented with 10% Fetal bovine serum (FBS, Thermo Fisher Scientific, 10500064). MCF10A (human breast epithelial cell line) were cultured in Mammary Epithelial Cell Growth Medium (Lonza, CC-3151) with supplements and growth factors (Lonza, CC-4136). HUVEC (primary Human Umbilical Vein Endothelial Cells, CC2517, LOT 0000482213) were maintained in Endothelial Cell Growth Medium Lonza (Lonza, CC-3121) with supplements and growth factors (Lonza, CC-4133). RPE1 cells were maintained in Dulbecco's modified Eagle's medium F12 Nutrient Mixture (Gibco, 31331-028) supplemented with 10% Fetal bovine serum (FBS, Thermo Fisher Scientific, 10500064, LOT 2025814K). All cell lines were obtained from the American Type Culture Collection (ATCC) and were cultured at 37°C with 5% $CO_2$. HeLa Kyoto (EGFP-α-tubulin/ H2B-mCherry) cells were obtained from ATCC/Jan Ellenberg (EMBL Heidelberg) [18] and cultured in DMEM with 10% FBS. All cell lines were verified by short tandem repeat (STR) profiling and tested negative for mycoplasma contamination.

## High-content imaging screen: Lincode siRNA library

The Lincode siRNA Library (GE Dharmacon, G-301000) is a collection of siRNA reagents targeting 2231 human lncRNAs (1860 unique human lncRNA genes and 371 lncRNA transcripts associated with protein-coding genes). The design of this library is based on RefSeq version 54 and the siRNAs are arrayed as SMARTpools. The library was purchased at 0.1 nmol in a 96-well format. The library was diluted to a 5 µM stock with 1x siRNA buffer (GE Dharmacon, B-002000-UB-100) and arrayed onto black 384-well PerkinElmer Cell Carrier plates (Perkin Elmer, 6007550) using the

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Echo Liquid Handler (Labcyte). The black 384-well PerkinElmer Cell Carrier plates were prepared in advance and stored at -80°C. The final siRNA concentration per well was 20 nM. An siRNA targeting exon 1 of lncRNA *GNG12-AS1* (Silencer select, Life Technologies, S59962)[64] and SMARTpool siRNAs targeting protein-coding gene *CKAP5/Ch-TOG* (GE Dharmacon, L-006847-00) were also included on each plate. *CKAP5/Ch-TOG* was used as a positive control as its depletion leads to mitotic delay and increased mitotic index [24]. *ECT2* SMARTpool siRNAs (GE Dharmacon, L-006450-00-0005) were used as a positive control in the third validation screen as its depletion results in multinucleated cells [26].

**High-content imaging screen: Reverse transfection**

To redissolve the siRNA in the black 384-well PerkinElmer Cell Carrier plates (10 plates in total), 5 µl of OptiMEM medium (Thermo Fisher Scientific, 31985047) was added. The plates were centrifuged (1 min, 900g) and incubated at room temperature (RT) for 5 min. Lipofectamine RNAimax (Thermo Fisher Scientific, 13778150) was added in OptiMEM medium to a final concentration of 8 µl Lipofectamine to 1 ml OptiMEM medium and incubated at RT for 5 min. 5 µl of OptiMEM/Lipofectamine mix was then added to the plates. Plates were centrifuged and incubated at RT for 20 min. In the meantime, HeLa cells were trypinised and counted using an automated cell counter (Countess, Thermo Fisher Scientific). Cells were centrifuged at 1000g for 4 min, the medium was removed and the cells were resuspended in OptiMEM medium to a final concentration of 2000 cells/well. 10 µl of cell suspension was added to the plates, plates were centrifuged and incubated at 37°C for 4 hours, before adding 10 µl of DMEM + 30% FBS + 3% P/S (penicillin/streptomycin, P/S) was added (final concentration 10% FBS, 1% P/S). Plates were centrifuged and incubated at 37°C for 48 hours before fixation. A

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Multidrop Combi Reagent dispenser (Thermo Fisher Scientific) was used throughout the transfection protocol to ensure even liquid addition.

**High-content imaging screen: Fixation and immunostaining**

For screen A, the plates were fixed by adding an equal volume of pre-warmed (37°C) 8% formaldehyde (Thermo Fischer Scientific, 28908)/PBS solution to the wells and incubated at 37°C for 10 min. The cells were permeabilised with pre-warmed PBS/0.2% Triton X-100 (Acros Organics, 327371000) for 15 min at RT. The cells were then blocked in 1% BSA/PBS for 1 hour at RT. To perform the immunostaining, the cells were incubated with primary antibodies against α-tubulin (Dm1α, Sigma, TUB9026), CEP215/CDK5RAP2 [65] and Alexa-Fluor® 568 Phalloidin (Thermo Fisher Scientific, A12380) for 2 hours at RT. The cells were washed three times in 1X PBS and incubated for 1 hour at RT with secondary antibodies Alexa Fluor® 488 (Thermo Fisher Scientific, A21206) and Alexa Fluor® 647 (Thermo Fisher Scientific, A31571). After three washes in 1X PBS, the cells were incubated with 1 µg/ml Hoechst (Sigma, H 33258, diluted in PBS) for 15 min at RT before a final wash in 1X PBS and imaging.

For screen B, the same fixation protocol as described for screen A was used. For permeabilization, PBS/0.05% SDS was used for 20 min at RT. Blocking was performed as described above. For the immunostaining, cells were incubated with primary antibodies against γ-tubulin (Sigma, GTU88) and phospho-histone H3 serine 10 (PHH3, Millipore, 06-570) for 2 hours at RT, washed three times with 1X PBS before incubation with secondary antibodies (Thermo Fisher Scientific, A-31571 and A-21206). After three 1X PBS washes, the cells were stained with α-tubulin (Serotec, MCA78G) and incubated with secondary antibody (Thermo Fisher Scientific, A-21434). All primary and secondary antibodies were diluted in 1% BSA/PBS.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The fixation and staining for both replicates was carried out at the Institute for Cancer Research (ICR, London) using the PerkinElmer Cell:Explorer system coupled to automated liquid handling equipment. Solutions were dispensed using a Multidrop Combi Reagent dispenser (Thermo Fisher Scientific) and aspirated/washed using a Biotek washer with 96 pins. All plates were imaged using the PerkinElmer Opera high-content confocal screening platform with spinning disc. Thirty fields of view per well were captured using a 20x air objective, numerical aperture (NA) 0.45.

**High-content imaging screen*:* Image analysis**

All image analysis was performed using custom workflows created with the Columbus software (PerkinElmer). Several output parameters were evaluated from high-content images: mitotic index (number of cells in mitosis), multinucleation index (number of multinucleated cells), number of viable cells, number of chromosome segregation errors (chromatin bridges and lagging chromatids) and number of cells with cytokinetic bridges (Supplementary Fig. 1). These are defined in more detail below:

i) **Mitotic and multinucleation index.** Nuclei were first segmented using Hoechst staining (which defines the total cell number). The false positives (e.g., dead cells) were discarded based on the nucleus area, α-tubulin and γ-tubulin/CEP215 staining intensity. Multinucleated non-dividing cells were retained as a separate subpopulation using a two-step detection process: binucleated cells were isolated using size, aspect ratio and roundness parameters of close nucleus pairs. Other multinucleated cells were then identified among remaining cells for which α- or γ-tubulin intensity was low in the perinuclear region. Further identification of mitotic cells/stages was

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

accomplished using filters based on Hoechst for the nucleus shape and size, in combination with high PHH3 (screen B) or high α-tubulin and low CEP215 staining intensity (screen A). Notably, nuclei of cells in anaphase/telophase stage of mitosis were small, had elongated shape and exhibited low Hoechst integrated intensity (low amount of DNA among mitotic cells). The distance between both nuclei of cells in anaphase/telophase stage was the main criteria to discard two daughter non-mitotic cells (maximum of 0.65 µm and 2.6 µm, respectively). From all these sub-populations, we calculated mitotic and multinucleation index relative to the total number of live cells.

ii) **Chromosome segregation errors.** We started from the previous identified subpopulation of cells in anaphase/telophase stage. We filtered cells according to α-tubulin staining intensity between nuclei, as cells in anaphase have lower α-tubulin intensity compared to the cells in telophase. This allowed us to identify only the cells in anaphase. To calculate the number of anaphase cells with chromosome segregation errors, the inter-nuclei space was used as the measuring area to calculate the remaining Hoechst signal. This captures both chromatin bridges and lagging chromatids.

iii) **Number of viable cells.** The total number of viable cells was determined after removal of dead cells and cell debris with anaphase and telophase cells counted as one (despite exhibiting two nuclear segments). The same rule was applied for multinucleated cells.

iv) **Number of cytokinetic bridges.** Cytokinetic bridges were defined as elongated high-intensity objects split into two parts that are positive for α-tubulin staining. We used prior segmentation of the cytoplasm and the Spot Finder feature to identify bridge half parts and sorted them as doublets by calculating distance

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

between them. We discarded the false positive candidates based on the shape criteria and γ-tubulin staining. The final number of cytokinetic bridges was divided by total number of viable cells.

To minimize the variation in the cell density between different wells among all ten 384-well plates, we divided the output numbers by the total number of cells per well (dead cells were not included). Multinucleated cells were considered to be single cells during counting. The ratios were then normalized between screen plates by calculating the average value per plate and finally the grand average of all ten plates, giving a reference mean ratio. Per-well ratios were scaled so that the per-plate average was equal to the reference. Z-scores ($z$) were calculated as follows for each parameter:

$$z = (x - \mu)/\sigma$$

where $x$ represents the ratio for the feature of interest (e.g., mitotic/multinucleation index), $\mu$ represents the reference ratio and $\sigma$ represents the standard deviation of ratios across wells. All the scripts used for the image analysis are available at https://github.com/MarioniLab/LncScreen2018.

**High-content imaging screen: Third pass validation screen**

The third-pass validation screen was performed in two replicates with four technical replicates using the top 25 candidates from each of the categories (mitotic progression, cytokinesis) using the same antibodies as in screen B. Correlation coefficients between replicate plates in third screen were calculated by Spearman's rank correlation (mitotic index = 0.967888, viability = 0.93249, multinucleation index = 0.995897, cytokinetic bridges = 0.898324).

**Single-molecule RNA fluorescent *in situ* hybridization (FISH)**

Cells were grown on coverslips in 12-well plates, briefly washed with 1X PBS and fixed with PBS/3.7% formaldehyde at RT for 10 min. Following fixation, cells were washed

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

twice with PBS. The cells were then permeabilized in 70% ethanol for at least 1 hour at 4°C. Stored cells were briefly rehydrated with Wash Buffer (2X SSC, 10% formamide, Biosearch) (Formamide, Thermo Fischer Scientific, BP227-100) before FISH. The Stellaris FISH Probes (*linc00899* and *C1QTNF1-AS1* exonic probes Q570) were added to the hybridization buffer (2X SSC, 10% formamide, 10% dextran sulfate, Biosearch) at a final concentration of 250 nM. Hybridization was carried out in a humidified chamber at 37°C overnight. The following day, the cells were washed twice with Wash Buffer (Biosearch) at 37°C for 30 min each. The second wash contained DAPI for nuclear staining (5 ng/ml, Sigma, D9542). The cells were then briefly washed with 2X SSC and then mounted in Vectashield (Vector Laboratories, H-1000). Images were captured using a Nikon TE-2000 inverted microscope with NIS-elements software, a Plan Apochromat 100x objective and an Andor Neo 5.5 sCMOS camera. We acquired 25 optical slices at 0.3 µm intervals. Images were projected in two dimensions using ImageJ and deconvolved with Huygens Professional.

For validation of CHART-sequencing, Stellaris FISH Probes for intronic region of *TPPP* (Q670) were combined with *linc00899* exonic probe (Q570) at a final concentration of 250 nM per probe set. To score whether *TPPP* (intronic signal) and *linc00899* (exonic signal) colocalize, we only considered cells in which both signals were present. The sequences of RNA FISH probes are presented in the Supplementary Table 4.

**Plasmids**

To insert poly(A) signal into lncRNAs, pSpCas9(BB)-2A-GFP vector was used (PX458, Addgene, #48138). For CRISPR activation (CRISPRa), pHAGE EF1alpha dCAS-VP64-HA (Addgene, #50918), pU6-sgRNA EF1Alpha-puro-T2A-BFP (Addgene, #60955), second-generation packaging plasmid psPAX2 (Addgene,

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

#12260) and the envelope plasmid pMD2.G (Addgene, #12259) were used. *Linc00899* and *C1QTNF1-AS1* constructs were synthesized by Labomics. The full sequence of *linc00899* was synthesised as CS-LNC233J-T7 (4283bp, insert was 1610bp based on NR_027036), scrambled *linc00899* was CS-LNC236J-T7, full *C1QTNF1-AS1* was CS-LNC237JT7 (3638bp, insert was 970bp based on NR_040018) and scrambled *C1QTNF1-AS1* was CS-LNC238JT7). Sequences of Labomics lncRNA vectors (pUCLOMT) are presented in the Supplementary Information.

**Lentiviral overexpression of lncRNAs**

For overexpression of *linc00899* in HeLa and RPE1 cells, *linc00899* sequence (Labomics) and negative control vector (scrambled *linc00899* sequence) were cloned into pLenti6.3/TO/V5-DEST vector (also known as lincXpress; kindly provided by John Rinn, University of Colorado) using Gateway cloning strategy as described previously [64]. To overexpress *linc00899* from its endogenous locus using CRISPRa, we used pHAGE EF1alpha dCAS9-VP64-HA (Addgene, #50918). Two gRNA sequences targeting different regions of *linc00899* (guide 1 and 2) were cloned into pU6-sgRNA EF1Alpha-puro-T2A-BFP (Addgene, #60955). All clones were verified by Sanger sequencing using mU6 forward primers (Supplementary Methods). Viral overexpression was performed as described previously [25]. Briefly, the cells were transduced with lentivirus containing dCAS9-VP64 and two gRNA targeting *linc00899,* or with lentivirus contacting negative guide RNA (NC2) in the presence of polybrene (5µg/ml, Sigma). 24hr after viral transduction the medium was exchanged and RNA was collected 48 hr after for qPCR analysis.

**siRNA and LNA depletion experiments**

HeLa cells were transfected with Lipofectamine RNAiMax reagent (Thermo Fischer Scientific) following the manufacturer's instructions. All experiments were done 48

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

hours after transfection. The pool of 4 siRNA sequences (SMARTpool, Thermo Fischer Scientifc) and LNA Gapmers (Exiqon) were used at a final concentration of 50 nM and 25 nm, respectively. For double knockdown experiments, HeLa cells (10,000 cells/well) were plated on 8-well chamber slides (Ibidi, 80826) for time-lapse microscopy imaging or in 6-well for RNA extraction (Corning, 120,000 cells/well). The cells were transfected the next day with either negative control siRNA (Ctl, from Ambion), a SMARTpool of siRNAs targeting *linc00899* or *TPPP* or siRNAs targeting *linc00899* in combination with *TPPP*. The same final concentration of 50 nM was achieved for both single and double knockdown by adding equal amount of control siRNA sequence (Ctl) to the single SMARTpool targeting *linc00899* or *TPPP* separately. siRNA and LNA sequences are listed in the Supplementary Information.

**Western Blot analysis**

Cells were grown in a 6-well plate, trypsinized, pelleted and washed twice with 1X PBS. The pellet was lysed in lysis buffer (50 mM Tris-HCl, pH 8, 125 mM NaCl, 1% NP-40, 2 mM EDTA, 1 mM PMSF [Sigma, 93482-50ML-F], and protease inhibitor cocktail [Roche, 000000011836170001) and incubated on ice for 25 min. The samples were centrifuged for 3 min at 12,000g and 4°C. Supernatant was collected and protein concentration was determined using the Direct Detect® Spectrometer (Merck Millipore). The proteins (30 µg) were denatured, reduced, and separated with Bolt® 4-12% Bis-Tris Plus Gel (Thermo Fisher Scientific, NW04120BOX) in MOPS buffer (Thermo Fisher Scientific, B0001-02). Precision Plus Protein Standards (161-0373, Bio-Rad) was used as a protein standard. The proteins were then transferred to nitrocellulose membrane and blocked with 5% nonfat milk in TBS-T (50mM Tris, 150mM NaCl, 0.1% Tween-20) for 1 hour at RT. The membranes were incubated with primary antibodies in 5% milk in TBS-T. After overnight incubation at 4°C, the

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

membranes were washed with TBS-T and incubated with HRP secondary antibodies (GE Healthcare Life Sciences, 1:2000), and immunobands were detected with a Pierce ECL Western Blotting Substrate (Thermo Fischer Scientific, 32106). Quantification of immunoblots normalized against appropriate loading controls was done using ImageJ. Uncropped scans of the immunoblots are shown in Figure S14. The list of primary and secondary antibodies is provided in the Supplementary Information.

**Time-lapse microscopy imaging**

HeLa (10,000 cells/well), HeLa Kyoto (10,000 cells/well), RPE1 (10,000 cells/well), MCF10A (15,000 cells/well) and HUVEC cells (15,000 cells/well) were cultured in 8-well chamber slides (Ibidi) with 200 µl/well of corresponding medium. Imaging was performed in their corresponding medium. Time-lapse microscopy imaging was performed for all cell lines 48 hours after transfection with RNAi or LNA gapmers. Mitotic duration was measured as the time from nuclear envelope breakdown (NEBD) until anaphase onset, based on visual inspection of the images. Cytokinesis was measured from anaphase onset to abscission completion. Live-cell imaging was performed using a Zeiss Axio Observer Z1 microscope equipped with a PL APO 0.95NA 40X dry objective (Carl Zeiss Microscopy) fitted with a LED light source (Lumencor) and an Orca Flash 4.0 camera (Hamamatsu). Four positions were placed per well and a z-stack was acquired at each position every 10 minutes (HeLa, HeLa Kyoto) or 5 minutes (MCF10A, HUVEC, RPE1) for a total duration of 12 hours. To detect chromosome segregation errors (chromatin bridges and/or lagging chromatids) Hela Kyoto cells were imaged every 4 min with only 2 positions/well. Voxel size was 0.325 µm x 0.325 µm x 2.5 µm. Zen software (Zeiss) was used for data collection and

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

analysis. Throughout the experiment, cells were maintained in a microscope stage incubator at 37°C in a humidified atmosphere of 5% $CO_2$.

**Cell cycle synchronization**

HeLa cells were grown to 50% confluency and then synchronized with thymidine for 16 hours (2 mM, Sigma, T1895). The cells were washed three times with 1X PBS and then released into thymidine-free medium for 5 hours. The cells were released for indicated timepoints for RNA collection. Treatment of HeLa cells with monastrol (100 µM for 16hr; Tocris, 1305) coupled with the mitotic shake-off was used to isolate mitotic cells.

**Copy number evaluation**

To calculate the *linc00899 and C1QTNF1-AS1* copy number, a standard curve of Ct values was generated by performing qPCR on a dilution series of known concentration of *linc00899* or *C1QTNF1-AS1* DNA templates (Labomics). cDNA was prepared from RNA (1 µg) extracted from known number of HeLa cells (500,000 cells). The observed Ct values were fitted on the standard curve and the number of lncRNA molecules per cell was calculated. The final value was multiplied by 2 to account for the fact that cDNA is single stranded and DNA templates used to make the standard curve were double stranded.

**RNA isolation, cDNA synthesis and qPCR**

RNA (1 µg) was extracted with the RNeasy Kit (QIAGEN, 74106) and treated with DNase I following the manufacturer's instructions (QIAGEN, 79254). The QuantiTect Reverse Transcription Kit (QIAGEN, 205313) was used for cDNA synthesis including an additional step to eliminate genomic DNA contamination. Quantitative real-time PCR (qPCR) was performed on a QuantStudio 6 Flex (Thermo Fischer Scientific) with Fast SYBR Green Master Mix (Life Technologies). Thermocycling parameters were

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

defined as 95°C for 20 sec followed by 40 cycles of 95°C for 1 sec and 60°C for 20 sec. Two reference genes (*GAPDH* and *RPS18*) were used to normalise expression levels using the $2^{-\Delta\Delta CT}$ method. Sequences of qPCR primers are provided in the Supplementary Information. For *linc00899* and *C1QTNF1-AS1* primers against exons 2-4 and 1-2, respectively, were used throughout the paper if not indicated otherwise.

**RNA library preparation, sequencing and analysis**

RNA-seq libraries were prepared from HeLa cells using TruSeq Stranded Total RNA Kit with Ribo-Zero Gold (IIllumina, RS-122-2303). We generated four biological replicates for RNAi and LNA-mediated depletion of *linc00899* and *C1QTNF1-AS1*. Indexed libraries were PCR-amplified and sequenced using 125 bp paired-end reads on an Illumina Hiseq 2500 instrument (CRUK CI Genomics Core Facility). Each library was sequenced to a depth of 20-30 million read pairs. Paired-end reads were aligned to the human genome hg38 with subread and the number of read pairs mapped to the exonic regions of each gene was counted for each library with using the featureCounts [66] function in Rsubread v1.30.0 with Ensembl GRCh38 version 91. Only alignments with mapping quality scores above 10 and with the first read pair on the reverse strand were considered during counting. Approximately 80% of read pairs contained one read that was successfully mapped to the reference, and 74% of all read pairs in each library were assigned into exonic regions. Any outlier samples with very low depth (resulting from failed library preparation or sequencing) were removed prior to further analysis.

The DE analysis was performed using the limma package v3.36.0 [67]. First, lowly expressed genes with average counts below 3 were filtered out. Normalization was performed using the TMM method to remove composition biases. Log-transformed expression values with combined precision/array weights were computed with the

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

voomWithQualityWeights function [68]. The experimental design was parametrized using an additive model with a group factor, where each group was comprised of all samples from one batch/treatment combination; and an experiment factor, representing samples generated on the same day. Robust empirical Bayes shrinkage [69] was performed using the eBayes function to stabilize the variances. Testing for DE genes was performed between pairs of groups using the treat function [70] with a log-fold change threshold of 0.5. Here, the null hypothesis was that the absolute $\log_2$-fold change between groups was less than or equal to 0.5. All pairwise contrasts involved groups from the same batch to avoid spurious differences due to batch effects. For each contrast, genes with significant differences in expression between groups were detected at an FDR of 5%. To identify DE genes that were consistent across LOF methods, an intersection-union test was performed [71] on one-sided p-values in each direction, followed by an FDR correction. Coverage tracks for each library were generated using Gviz as previously described.

**Capture Hybridization Analysis of RNA targets (CHART) sequencing and analysis**

CHART enrichment and RNase H (NEB, M0297S) mapping was performed as previously described [47, 51]. Briefly, five 150-mm dishes of HeLa cells were used to prepare the CHART extract for each pulldown. In the first sonication step, the samples were sonicated using Covaris S220 (Covaris, 500217) in microTUBEs (Covaris, 520045) in a final volume of 130 µl (Program conditions: 20% duty cycle, 200 bursts/cycle, Intensity 175, 8 min). The extracts were hybridised with a *linc00899* oligonucleotide cocktail (mix of 4 oligos used at final concentration of 25 µM) overnight at RT. MyOne streptavidin beads C1 (Thermo Fischer Scientifc, 65001) were used to capture complexes overnight at RT by rotation. Bound material was washed five times

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and RNase H (NEB) was added for 30 min at RT to elute RNA-chromatin complexes. To increase recovery yield, the remaining beads were saved and the bound material was eluted with Proteinase K (Thermo Fischer Scientifc, 25530-049) in 100 µl of XLR buffer for 60 min at 55°C. The supernatant was then collected and heated for additional 30 min at 65°C. The RNase H eluate samples were also treated with Proteinase K (Thermo Fischer Scientifc, 25530-049) and cross-links were reversed (55°C for 1hr, followed by 65°C for 30 min). One fifth of the total sample was used to purify RNA using miRNeasy kit (QIAGEN, 217004) and to calculate RNA enrichment, while the rest of the sample was used to purify DNA using Phenol-$ChCl_3$:isoamyl (Thermo Fischer Scientific, 15593-031) extraction and ethanol precipitation. The final CHART DNA was eluted in 12.5 µl of 1X TE buffer (low EDTA, pH 7.4). 10 µl from this reaction was used for sonication using Covaris LE220 (programme 250bp/10sec) to obtain average fragment size of 200-300 bp. For this, microTUBE-15 beads strips (520159, Covaris) were used in a total volume of 15 µl. After sonication, the DNA was measured by Nanodrop (input, diluted 1:20) and Qubit High Sensitivity DNA Assay (eluates were undiluted). CHART material (5 ng) was used for library preparation using the ThruPLEX DNA-Seq library preparation protocol (Rubicon Genomics, UK). For inputs, 6 PCR cycles were performed, while for eluates 8 PCR cycles were performed due to the lower quantity of input DNA. Library fragment size was determined using a 2100 Bioanalyzer (Agilent). Libraries were quantified by qPCR (Kapa Biosystems). Pooled libraries were sequenced on a HiSeq4000 (Illumina) according to manufacturer's instructions to generate paired-end 150bp reads (CRUK CI Genomics Core Facility). CHART-seq was performed from five biological replicates using a mix of two sense control oligos for *linc00899* as a negative control for the *linc00899* pulldown. Reads were aligned to the hg38 build of the human genome using subread in paired-end

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

genomic mode. A differential binding (DB) analysis was performed using csaw to identify lncRNA binding sites in the pulldown compared to the sense control. Filtered windows were obtained as described for the CUT&RUN data analysis. Normalization factors were computed by binning read pairs into 5 kbp intervals and applying the TMM method without weighting, to account for composition biases from greater enrichment in the antisense pulldown samples. The filtered windows and normalization factors were then used in a DB analysis with the quasi-likelihood framework in edgeR. This was performed using an additive design for the generalized linear model fit, containing terms for the batch and the pulldown (5 batches in total, sense and antisense pulldowns in each batch). A p-value was computed for each window by testing whether the pulldown term was equal to zero, i.e., no difference in coverage between the antisense pulldown and sense control. The above analysis was repeated using window sizes from 150 to 1000 bp to obtain DB results at varying resolutions. These results were consolidated into a single list of DB regions using consolidateWindows as previously described, yielding a combined p-value for each region. The combined p-values across all regions were then used to compute the empirical false discovery rate (eFDR), defined as the ratio of the number of clusters that exhibited increased coverage in the sense control (i.e., false positives) to the number of clusters with increased coverage in the antisense pulldown (i.e., potentially true discoveries). Putative lncRNA binding sites were defined at an eFDR of 30%. Coverage tracks for each library were generated using Gviz as previously described.

**CUT&RUN (Cleavage under targets and release using nuclease)**

Chromatin profiling was performed according to Skene *et al.* (2018) [46] with minor modifications. HeLa cells were plated in 30-mm dishes (220,000 cells/dish) and transfected the next day with either negative control siRNA, *linc00899* siRNA pool,

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

negative control LNA A or an LNA gapmer targeting *linc00899* (LNA 1). Two biological replicates were performed per condition. Cells were washed once with PBS, spun down at 600g for 3 minutes in swinging-bucket rotor and washed twice with 1.5 mL Wash buffer (20 mM HEPES-KOH (pH 7.5), 150 mM NaCl, 0.5 mM Spermidine and 1X cOmplete™ EDTA-free protease inhibitor cocktail (Roche)). During the cell washes, concanavalin A-coated magnetic beads (Bangs Laboratories, BP531) (10 µl per condition) were washed twice in 1.5 mL Binding Buffer (20 mM HEPES-KOH (pH 7.5), 10 mM KCl, 1 mM CaCl, 1 mM $MnCl_2$) and resuspended in 10 µl Binding Buffer per condition. Cells were then mixed with beads and rotated for 10 minutes at RT and samples were split into aliquots according to number of antibodies profiled per cell type. We used 80-100,000 cells per chromatin mark. Cells were then collected on magnetic beads and resuspended in 50 µl Antibody Buffer (Wash buffer with 0.05% Digitonin and 2mM EDTA) containing one of the following antibodies in 1:100 dilution: H3K4me3 (Millipore 05-1339 CMA304, Lot 2780484), H3K27me3 (Cell Signaling #9733S C36B11, Lot 8), H3K36me3 (Active Motif Cat#61101, Lot 32412003), H3K27ac (Abcam ab4729, Lot GR3211741-1) and Goat Anti-Rabbit IgG H&L (Abcam ab97047, Lot GR254157-8). Cells were incubated with antibodies overnight at 4°C rotating, and then washed once with 1 mL Digitonin buffer (Wash buffer with 0.05% Digitonin). For the mouse anti-H3K4me3 antibody, samples were incubated with 50 µl of a 1:100 dilution in Digitonin buffer of secondary rabbit anti-mouse antibody (Invitrogen, A27033, Lot RG240909) for 10 minutes at RT and then washed once with 1 mL Digitonin buffer. Samples were then incubated in 50 µl Digitonin buffer containing 700 ng/ml Protein A-MNase fusion protein (kindly provided by Steven Henikoff) for 10 minutes at RT followed by two washes with 1 mL Digitonin buffer. Cells were then resuspended in 100 µl Digitonin buffer and cooled down to 4°C before addition of

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$CaCl_2$ to a final concentration of 2 mM. Targeted digestion was performed for 30 minutes on ice until 100 µl of 2X STOP buffer (340 mM NaCl, 20 mM EDTA, 4 mM EGTA, 0.02% Digitonin, 250 mg RNase A, 250 µg Glycogen, 15 pg/ml yeast spike-in DNA (kindly provided by Steven Henikoff)) was added. Cells were then incubated at 37°C for 10 minutes to release cleaved chromatin fragments, spun down for 5 minutes at 16,000g at 4°C and collected on magnet. Supernatant containing the cleaved chromatin fragments were then transferred and cleaned up using the Zymo Clean & Concentrator Kit. Library preparation was performed as described [72] using the ThruPLEX® DNA-Seq Library Preparation Kit (Rubicon Genomics) with a modified Library Amplification programme: Extension and cleavage for 3 minutes at 72°C followed by 2 minutes at 85°C, denaturation for 2 minutes at 98°C followed by four cycles of 20 seconds at 98°C, 20 seconds at 67°C and 40 seconds at 72°C for the addition of indexes. Amplification was then performed for 12 cycles of 20 seconds at 98°C and 15 seconds at 72°C (14 cycles were used for the Goat Anti-Rabbit antibody due to the lower yield of input DNA). Double-size selection of libraries was performed using Agencourt AMPure XP beads (Beckman Coulter, A63880) according to manufacturer's instructions. Average library size was determined using an Agilent Tapestation DNA1000 High Sensitivity Screentape and quantification was performed using the KAPA Library Quantification Kit. CUT&RUN libraries were sequenced on a HiSeq2500 using a paired-end 125bp run at the CRUK CI Genomics Core Facility. Reads were aligned to a reference genome containing the hg38 build of the human genome and the R64 build of the yeast genome, using subread v1.6.1 [73] in paired-end genomic mode. A DB analysis was performed using the csaw package v1.14.0 [74] to identify changes in histone mark enrichment in *linc00899* depleted cells compared to the negative controls for each LOF method. Coverage was quantified by sliding a

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

window across the genome and counting the number of sequenced fragments overlapping the window in each sample. Windows were filtered to retain only those with average abundances that were 5-fold greater than the expected coverage due to background non-specific enrichment. Normalization factors were computed by applying the trimmed mean of M-values (TMM) method [75] without weighting to the filtered windows. The filtered windows and normalization factors were then used in a DB analysis with the quasi-likelihood framework in the edgeR package v3.22.0 [76]. This was performed using an additive design for the generalized linear model fit, containing terms for the batch and the depletion status (2 batches in total, depletion and negative control from LNA and RNAi in each batch). A p-value was computed for each window by first testing whether for a depletion effect in each LOF method, and then taking the larger of the two p-values [71] across both LOF methods. The above analysis was repeated using window sizes from 150 to 1000 bp to obtain DB results at varying resolutions. To consolidate these results into a single list of DB regions, overlapping windows of all sizes were clustered together based on their genomic locations, using a single-linkage approach in the consolidateWindows function. For each cluster of windows, a combined p-value was computed using Simes' method [77]. This represents the evidence against the global null hypothesis for each cluster, i.e., that none of the constituent windows are DB. The Benjamini-Hochberg method was used to define putative DB regions at an FDR of 5%. Coverage tracks were generated by computing the number of sequencing fragments per million overlapping each based, using library sizes adjusted according to the TMM normalization factors for each library. Tracks were visualized using the Gviz package [78].

**Generation of *C1QTNF1-AS1* CRISPR poly(A) site clones**

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Insertion of the transcriptional termination signal (poly(A) signal (pAS) [34] into the first exon of *C1QTNF1-AS1* was performed using a published CRISPR/CAS9 protocol [79]. Briefly, the CRISPR pAS guide oligonucleotides were phosphorylated (T4 poly nucleotide kinase; NEB), annealed and ligated (Quick Ligase kit, NEB) into *Bbs*I (NEB) digested pX458 vector (Addgene, #48138). The oligonucleotides (Sigma-Aldrich, diluted in low TE EDTA buffer at 100 µM) used for the gRNA target sequence are listed in the Supplementary Methods. All inserts were verified with Sanger sequencing. HeLa cells were plated in 6-well plates (200,000 cells/well) and on the following day, transfected with either 2.5 µg of empty PX458 vector as a control or 2.5 µg of PX458 vector where guide oligonucleotides targeting *C1QTNF1-AS1* were cloned together with a symmetric single stranded oligonucleotide donor [80] (ssODN; IDT Inc, 4 µl from 10 µM stock) containing the pAS flanked by homology arms (75bp) to the target site. Lipofectamine 3000 reagent (ThermoFisher Scientific, L3000015) was used as transfection reagent according to the manufacturer's instructions. The sequence of the symmetric ssODNs is as follows (5' to 3'):

75bp homology left arm-***aataaaagatctttattttcattagatctgtgtgttggttttttgtgtg***-75 bp homology right arm

*CTCACTGCGGGGCTCGGGAAGGAGGAAAGGAGTGAGCATGTCCTGCTCCTGC*

*ATGTCCCTGCTTAAGCTCAGGAC**aataaa**<u>**agatct**</u>**ttattttcatt**<u>**agatct**</u>**gtgtgttggttttttgtgt***

***g**CCCTTCCAGGCCAAGGACCCCAGCATAGACCCCAGGACAGGGCCCCAAGGA*

*TCCCTGGCTCATGAGAGCGGCTTGC*

The exon 1 of *C1QTNF1-AS1* is shown in upper case. The pAS site is in **bold** and contains *Bgl*II restriction sites (underlined) that were used for identification of the positive clones. The PAM site (TGG) was abrogated to avoid further cleavage by Cas9.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

After 48 hours, the GFP-positive cells were sorted using BD FACSAria IIu (CRUK Flow Cytometry Core Facility) and plated on 96-well plates with DMEM supplemented with 20% FBS. Single cell clones were expanded and genomic DNA was extracted using Direct PCR Lysis reagent (Viagen, 201-Y; 25 µl/well) with Proteinase K (Thermo Fisher Scientific, 25530-049; 0.4 mg/ml). The plates were incubated for 1 hour at 55°C, followed by 45 min incubation at 85°C. 2 µl of genomic DNA was used for screening by PCR amplification (Phusion® High-Fidelity PCR Master Mix with HF Buffer, NEB, M0531S) of the targeted genomic region. PCR conditions for *C1QTNF1-AS1*_guide 70 were: 98°C 30 sec, 98°C 30 sec, 70°C 30 sec, 72°C 30 sec (repeat 30X), 72°C 90 sec, 4°C forever. After PCR, 5 µl was used for *Bgl*II digestion (NEB, 2 hours at 37°C) and the products were loaded on a 2% agarose gel to confirm the pAS insertion. The rest of PCR reaction (15 µl) was also loaded on the gel as a control. The same genomic PCR product was also ligated into pJET1.2/Blunt and transformed into bacteria (CloneJET; ThermoFisher Scientific, K1231). To ensure representation by both alleles, plasmids were isolated and sequenced from a minimum of 5-10 bacterial colonies (Figure S8). This method revealed that *C1QTNF1-AS1* pAS clones were homozygously targeted clones. Four clones were analysed (clone 95, 136, 153 and 169) by qPCR and live-cell imaging for the mitotic phenotype, along with the wild type controls (clone 2 and clone 5) that were transfected with an empty vector PX458. We also attempted to perform pAS insertion into exon 1 of *linc00899* using two different sgRNAs. However, we failed to obtain homozygous targeted clones, most likely due to the presence of 4 copies of *linc00899* in HeLa cells (*https://cansar.icr.ac.uk/cansar/cell-lines/HELA/copy_number_variation/).*

**Mapping of 5′ and 3′ ends of *linc00899* and *C1QTNF1-AS1* by RACE**

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

HeLa total RNA (1 µg) was extracted and 5' and 3' rapid amplification of cDNA ends (RACE) was performed using the Smarter RACE kit (Clontech, 634858). cDNA was synthesized as described in the kit using 5′ and 3′ RACE CDS primers and SMARTer IIA oligo for template switching for 5′ RACE. cDNA ends were then amplified by touchdown PCR. The first PCR (touchdown) used the following conditions: 5 cycles of 94°C/30 sec, 72°C/3 min; 5 cycles of 94°C/30 sec, 70°C/30 sec, 72°C 3 min; 25 cycles of 94°C/30 sec, 68°C/30 sec, 72°C/3 min. For nested PCR, 2 µl of PCR reaction was diluted in 98 µl Tricine-EDTA buffer and used as a template for the second PCR reaction with following conditions: 25 cycles of 94°C/30 sec, 68°C/30 sec, 72°C/3 min. For *linc00899,* the first PCR used Universal Primer A mixed with *linc00899*_5R outer 1=GATTACGCCAAGCTTacatcccggttcccacgaaaagcaacc for the 5′ ends or *linc00899*_3Routerinner3=GATTACGCCAAGCTTccagggaggggaaaggagtcggcaat for the 3′ ends. The second (nested) PCR used Nested Universal Primer A and *linc00899*_5Routerinner1=GATTACGCCAAGCTTggagcaggcgaagagggagtgagggg for the 5′ ends or *linc00899*_3R outerinner1=GATTACGCCAAGCTTggtcacagcctagccaagcccagcca for the 3′ ends.

For *C1QTNF1-AS1*, the first PCR used Universal Primer A mixed with *C1QTNF1-AS1*_5_RACE_outer_1=GATTACGCCAAGCTTccaggcccctaatgatgtcctttga for the 5′ ends or *C1QTNF1-AS1*_3RACE_outer_1=GATTACGCCAAGCTTggaggaaaggagtgagcatgtcctg for the 3′ ends. The second nested PCR used Nested Universal Primer A and *C1QTNF1-AS1*_5RACE_inner_2 = GATTACGCCAAGCTTGTCCTGATCTCCACCTGTCCCAAGC for the 5′ ends or *C1QTNF1-AS1*_3RACE_inner_2= GATTACGCCAAGCTTGGAAACTTGGCAGACAGATCCAGCC for the 3′ ends.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

For *linc00899,* nine different 5' sites and four different 3' sites were identified. For *C1QTNF1-AS1*, six different 5' sites and four different 3' sites were identified. The fragments were purified after agarose gel electrophoresis with the QiaQuick Gel Extraction Kit (QIAGEN), cloned in pRACE vector as per kit instructions and transformed into Stellar component cells (Clontech, 636763). The inserts were verified by Sanger sequencing. Uncropped images of RACE results are presented in the Supplementary Fig. 14.

**RACE results for *linc00899*:** We have identified nine different 5' start sites (pink) and four 3' sites (yellow). The most common starting site for *linc00899* was CACGTCC, 87 bp upstream from the *linc00899* TSS. The most common termination site for *linc00899* was acagcaag. Capital italic letters indicate the region upstream of *linc00899* TSS (cggccgcccc) based on UCSC.

*TGGCCTCGGAGGGTCAACACCTGAGGGCCCGAGGGGCATCCATGCCCCCTCC*
*TCCTTCCCTCCAGCCAAAGGTGGGGGGCAAAGCGCAGGGAAGAAAACACCAG*
*GAGAAAACCAGAGAGCTCTCGGTTCTCTTTTAAGAGCCCTGATGGCCGCAGCG*
*CAGAGCCCGAGGGGAGGGAAAATGTCGGGAAAGATTCTCTTCCGAACTTTGCG*
*AGTCTTTGTTTGGGAGGCTGGGGGCTGACTTCGCCGGGGGCCGGGCCGCGG*
*GCTCGGCCGTGCGCTCCGGTGCAGCGGCCGAGGAGCCCCGGCGCCCGCCAC*
*CCCGGGACACGCCCTCGCAGTCGAGCCCGGACCCCGACCCGGACCCCAGCG*
*CCGCCGGGCGAGGGCGGGAGGGGGAGCGCTTACCAGATCGTCCCGAGCGCG*
*CCGCGGTCCAGGCGGGCACAGCGCAGGGTCAAGTTCACGTCCGGCCCGCGG*
*GCTGCCCGAGGTCCCCGGGCGCGGCTGGGGCAGCGGGAGGCGCGGGAGGC*
*CGAGGTCCGGGTGGCCGCCG*cggccgcccc cgaagcgctgctgtcaccccggccgcgccccccaa
ctttctgcacagtcgcggagctggaagtttccgggcttcgcggacacgctgggctgggtttcagtcgcggctccgaggt
tggcaacaaagagggaaagaaggaggaaaagcaggccggggaggggaggaagagaaccgcgcggaggcc

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

gcggcgccgagagccccagaacttccaattctacccagaagctttttttcgtcgtgtttttctcttagacatgatcctctctga
ggttggtcctgggcttccatacgtgattcatggaagaggtctcagccccaagagcccctgagggtactgtccactcccc
ctggaaacttccagaacctgacgtggggctgaagacatagaggctctgagagttacataattgattctgactttggctgt
tggtcaacagtgtcataaggtaaaataaggctgttgtagaatctgctcagccagggaggggaaaggagtcggcaat
caggtctcctcctgggcacctttgtgaggccagctggcgagagtgggggggtgacactgaggtcccagcagctccaa
atgcaggcagagccctgtcctcagagaaggtcacagcctagccaagcccagccaggtggatgggcccacggaac
gcacaggaacctggaacggaggttgaaagcaggaagcacagtctgtgactccccagcccactctgcattcgacca
cttggggcccagaagcttcaggaaaggtgcacaaggtcactgggtcccagtactcccaacaggaaggtctggtcca
gggacagggctcttcccgactcccttagccacacgcaccagaagttctgcagtgcccagtgggcatagcagtcccc
aagaatgacccagcactgaagctgagccaaagaacttggggagcgagccacacccccctcactccctcttcgcctg
ctccagacttgccaggcggttgcttttcgtgggaaccgggatgtcctcaccaccctgtccagggcccagccccatgtcc
ctggcctgctacagctggaaaaaaaaaagagagatgtttgtttttatttgtttataaaaagaaaagtgttatatatataac
atattatacctcatgaatacatacaattatttgtcaattaacaataaagaaaaatacagcaagcaaaaaagactctcttc
cacaaaaatagtgttcattacagaaaagtacaaaaaaaaaaaaaaaaaactaagaggatatttagaattaagaaa
aaactaagagggtatttagaattaaaaaataaaaagaaaacaattacccatgaggtaattactgaatgcatttggttg
aaagtccttctgctatattttccaaactgtatgtgtatatatgtgcgcatgcatatatgtgtttgtgtgtgtgtacacatatctat
atgaatggaccatatcgtaagttataaatgcatacatatattcatgtatatataatcattagatcatactataggttattttac
agtctttttgttgaatacgttgtgagcattttatgtcattattttctacaagatttgaaaaataaagtataaataccagttaa.

Based on our sequencing results, *linc00899* transcript varied from 1144bp to 1532bp in HeLa cells. We seqeunced at least 10 clones per 5‘ and 3‘ RACE.

**RACE results for *C1QTNF1-AS1*:**

We have identified six different 5‘ start sites (pink) and four 3‘ sites (yellow). The most common 5‘ site was AGAGAACTA, 54 bp from TSS. The most common 3‘site was

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

tctggaa. Capital italic letters indicate the region upstream of *C1QTNF1-AS1* TSS (gaaggagg) based on UCSC.

*ATCACCCCAGCACAAGTGTCACACAGCCGTGACCTTGACAAGGACCCAGAGAT*
*AAAGATCCTTCCCACATGGCTCCGAAGCCCCTCCCTTCCTGCTCACACCGCAT*
*GCCTCTCCCAGAGAACTAGAGGCTGCAGCGGCAGCAGTTGGAGCATCTCACTG*
*CGGGGCTCGG*gaaggaggaaaggagtgagcatgtcctgctcctgcatgtccctgcttaagctcaggactggc
ccttccaggccaaggaccccagcatagaccccaggacagggccccaaggatccctggctcatgagagcggcttg
ctgggctgccccaagagagcctgaaggaaacacattgttgagctgagctgacgtcgctgtttcttccagactgctctct
aaagtgggcagggtagcgaccggccggctccgatggtgacgtcccactgccaaggggtgggagtggggagagtc
tccacagagcttcggagaagctgctaagATGGAAAAGTGGAAACTTGGCAGACAGATCCAG
CCTCCCTGGCCACTGGCCCATGCTCGTGGCTCCTGGATGGCGCTGCCACGTTC
TGAGCAGCTTGGGACAGGTGGAGATCAGGACTGGCAGCTGCAAGGACACACC
AGAGCCACAGAAACTAAAGAGAATTTCCAAAAGGAGTCTATGGTGAAGTCTCTG
AGGATGCAAAGAAGACAAGGAGAATGAaaatccaatgaaagcctgattgtatttgttgaccttaagg
aaagtgattttatggtacagcctctctggaagggagggtgtgttcgctcacagaatgcaaataccctttgacccctaat
ctttcttctaggagtttctcctacagataaacttagaagggtgctcaaataagtaagttcaaggatatcctctgaagcattg
ccgtagtataaaaaaagcacagataccctcaaaggacatcattaggggcctggtaaaataaattccacacagtgga
acaccgtgtagctctttagagaataaacagctctctatatgtgatctggaaCAATCTCC.

Based on our sequencing results, *C1QTNF1-AS1* transcript varied from 864bp to 952bp in HeLa cells. We seqeunced at least 10 clones per 5‘ and 3‘ RACE.

**Immunofluorescence**

HeLa cells were seeded on coverslips in 6-well plates (120,000/well), transfected the next day with siRNA or LNA gapmers and fixed 48 hours post-transfection in ice-cold methanol (Acros Organics, 167830025) for 10 min at -20°C. The cells were then washed once in 1X PBS and permeabilized with PBS/0.5%Triton-X100 (Acros

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Organics, 327371000)/0.5%Tween-20 (Promega, H5151) for 5 min at RT followed by blocking in 5% BSA/PBS/0.001% Sodium Azide for 30 min at RT. Cells were incubated with primary antibodies overnight at 4°C. Antibodies against acetylated α-tubulin (Sigma) and α-tubulin (Serotec) or against EB1 (CRUK,[81]) and α-tubulin (Serotec) were used. Cells were washed 3 × 10 min with PBS/0.1%Tween-20 and then incubated with secondary antibodies (Supplementary Methods) diluted in blocking-buffer for 1 hour at RT. After washing again 3 × 10 min with PBS/0.1%Tween-20, cells were stained with 1 µg/ml Hoechst (B2261, Sigma, diluted in PBS) for 10 min at RT. Coverslips were mounted onto glass slides using Prolong Diamond Antifade Mountant (ThermoFischer Scientifc, P36961). To determine the mitotic index, cells were stained with antibodies against PHH3 (Millipore) and α-tubulin (Sigma). The mitotic index was calculated by counting the cells in mitosis (positive for PHH3 and α-tubulin) and total number of cells (Hoechst positive). For each sample, at least 100 cells were randomly counted by fluorescence microscopy and mitotic cells were scored from prophase to anaphase/telophase.

**Image processing and quantification**

Imaging of fixed cells was performed on a Leica Sp8 confocal microscope using a 100 X 1.4 numerical aperture Leica oil objective where z-stacks with 0.5 µm step size were recorded. Images were taken at identical exposure times within each experiment and were imported into ImageJ and Photoshop (CS6, Adobe). Images shown here represent 3D maximum intensity projections. To analyse the ratio of α-tubulin to acetylated α-tubulin, raw integrated intensities were measured (ImageJ) over the total z-stack plane using a circle selection around the mitotic spindle (based on α-tubulin). The measured intensity values were divided by the area of the selection and the background signal was subtracted. Afterwards, the mean of the

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

signal intensity over all z-stacks was calculated and the ratio of the signal of acetylated-tubulin to α-tubulin was determined. The total EB1 levels were quantified as descried above.

**Subcellular fractionation**

RNA was fractionated as described previously [64]. Briefly, cells from a 150-mm dish were used to isolate RNA from cytoplasmic, nucleoplasmic and chromatin fractions by TRIzol extraction. Turbo DNA-free kit (Thermo Fischer Scientifc, AM1907) was used to remove any traces of DNA. Expression of target genes in each fraction was analysed by qPCR. Data were normalised to the geometric mean of *GAPDH* and *ACTB* levels in each cellular compartment. *U1* small nuclear RNA was used as a positive control for chromatin fraction.

**Coding potential assessment for lncRNAs**

The Coding-Potential Calculator[82] (CPC, http://cpc.cbi.pku.edu.cn) and Coding Potential Assessment Tool[83] (CPAT, http://lilab.research.bcm.edu/cpat/index.php) were used to determine noncoding potential. LncRNAs with CPC scores above 1 and CPAT scores above 0.364 were predicted to have protein-coding capacity. The PhyloCSF [84] score was taken from UCSC (https://github.com/mlin/PhyloCSF/wiki).

**Statistical analysis**

The statistical significance of data was determined by two-tailed Student's t-test in all experiments using GraphPad Prism unless indicated otherwise. P-values greater than 0.05 were considered statistically not significant.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## DATA AVAILABILITY

Sequencing data are available in the ArrayExpress database (http://www.ebi.ac.uk/arrayexpress) with the accession codes E-MTAB-7432 (RNA-seq), E-MTAB-7418 (CHART-seq) and E-MTAB-7419 (CUT&RUN). The imaging data have been submitted to the Image Data Resource (https://idr.openmicroscopy.org) under IDR accession number idr0056.

## CODE AVAILABILITY

All code used in this analysis is available at https://github.com/MarioniLab/LncScreen2018.

## FUNDING

This work was made possible by funding from Cancer Research UK (C14303/A17197 to F. G, A17197 to J.C.M and A20412 to D.T.O). We also acknowledge the support of the University of Cambridge, the Wellcome Trust (WT202878; D.T.O), European Research Council (615584; D.T.O) and Hutchison Whampoa Limited.

## ACKNOWLEDGMENTS

We thank all the members of Gergely and Odom groups for helpful discussions and Julia Tischer for critical reading of the manuscript. We also thank the Genomics, Microscopy, FACS and Research Instrumentation and Cell Service Core Facilities at the CRUK Cambridge Institute. We thank Keith Vance (University of Bath, UK) and Matthew Simon (Yale, USA) for the help with the CHART-seq protocol, John Rinn (University of Colorado) for the lincXpress vector, and Steve Henikoff (Fred Hutchinson Cancer Research Center, USA) for providing the reagents for CUT&RUN.

## AUTHOR CONTRIBUTIONS

Conception and design of study: LS, FG and DTO. Data acquisition: LS, ATLL, PM, CE, AMR, JM, ARB, VB. Data analyses and interpretation: LS, ATLL, PM, ARB, JCM,

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CB, DTO, FG. Writing the paper: LS, ATLL, DTO and FG with input from all of the authors wrote the manuscript.

## DECLARATION OF INTERESTS

None.

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

# REFERENCES


1. Iyer, M.K. *et al.* The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet* **47**, 199-208 (2015).
2. Hon, C.C. *et al.* An atlas of human long non-coding RNAs with accurate 5' ends. *Nature* **543**, 199-204 (2017).
3. Fatica, A. & Bozzoni, I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet* **15**, 7-21 (2014).
4. Kitagawa, M., Kitagawa, K., Kotake, Y., Niida, H. & Ohhata, T. Cell cycle regulation by long non-coding RNAs. *Cell Mol Life Sci* **70**, 4785-4794 (2013).
5. Quinn, J.J. & Chang, H.Y. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* **17**, 47-62 (2016).
6. Huarte, M. The emerging role of lncRNAs in cancer. *Nat Med* **21**, 1253-1261 (2015).
7. Wapinski, O. & Chang, H.Y. Long noncoding RNAs and human disease. *Trends Cell Biol* **21**, 354-361 (2011).
8. Huarte, M. *et al.* A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell* **142**, 409-419 (2010).
9. Liu, X., Li, D., Zhang, W., Guo, M. & Zhan, Q. Long non-coding RNA gadd7 interacts with TDP-43 and regulates Cdk6 mRNA decay. *EMBO J* **31**, 4415-4427 (2012).
10. Kim, J. *et al.* LncRNA OIP5-AS1/cyrano suppresses GAK expression to control mitosis. *Oncotarget* **8**, 49409-49420 (2017).
11. Wang, Z.D. *et al.* Long noncoding RNA lnc-RI is a new regulator of mitosis via targeting miRNA-210-3p to release PLK1 mRNA activity. *Sci Rep* **6**, 25385 (2016).
12. Marchese, F.P. *et al.* A Long Noncoding RNA Regulates Sister Chromatid Cohesion. *Mol Cell* **63**, 397-407 (2016).
13. Rosic, S., Kohler, F. & Erhardt, S. Repetitive centromeric satellite RNA is essential for kinetochore formation and cell division. *J Cell Biol* **207**, 335-349 (2014).
14. Lee, S. *et al.* Noncoding RNA NORAD Regulates Genomic Stability by Sequestering PUMILIO Proteins. *Cell* **164**, 69-80 (2016).
15. Tichon, A., Perry, R.B., Stojic, L. & Ulitsky, I. SAM68 is required for regulation of Pumilio by the NORAD long noncoding RNA. *Genes Dev* **32**, 70-78 (2018).
16. Marchese, F.P., Raimondi, I. & Huarte, M. The multidimensional mechanisms of long noncoding RNA function. *Genome Biol* **18**, 206 (2017).
17. Kittler, R. *et al.* An endoribonuclease-prepared siRNA screen in human cells identifies genes essential for cell division. *Nature* **432**, 1036-1040 (2004).
18. Neumann, B. *et al.* Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* **464**, 721-727 (2010).
19. Kittler, R. *et al.* Genome-scale RNAi profiling of cell division in human tissue culture cells. *Nat Cell Biol* **9**, 1401-1412 (2007).
20. Schmitz, M.H. *et al.* Live-cell imaging RNAi screen identifies PP2A-B55alpha and importin-beta1 as key mitotic exit regulators in human cells. *Nat Cell Biol* **12**, 886-893 (2010).
21. Liu, S.J. *et al.* CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* **355** (2017).
22. Liu, Y. *et al.* Genome-wide screening for functional long noncoding RNAs in human cells by Cas9 targeting of splice sites. *Nat Biotechnol* (2018).

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

23. Notzold, L. *et al.* The long non-coding RNA LINC00152 is essential for cell cycle progression through mitosis in HeLa cells. *Sci Rep* **7**, 2265 (2017).
24. Gergely, F., Draviam, V.M. & Raff, J.W. The ch-TOG/XMAP215 protein is essential for spindle pole organization in human somatic cells. *Genes Dev* **17**, 336-341 (2003).
25. Stojic, L. *et al.* Specificity of RNAi, LNA and CRISPRi as loss-of-function methods in transcriptional analysis. *Nucleic Acids Res* **46**, 5950-5966 (2018).
26. Su, K.C., Takaki, T. & Petronczki, M. Targeting of the RhoGEF Ect2 to the equatorial membrane controls cleavage furrow formation during cytokinesis. *Dev Cell* **21**, 1104-1115 (2011).
27. Hezroni, H. *et al.* Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Rep* **11**, 1110-1122 (2015).
28. Ulitsky, I. Evolution to the rescue: using comparative genomics to understand long non-coding RNAs. *Nat Rev Genet* **17**, 601-614 (2016).
29. Marin-Bejar, O. *et al.* The human lncRNA LINC-PINT inhibits tumor cell invasion through a highly conserved sequence element. *Genome Biol* **18**, 202 (2017).
30. Quinn, J.J. *et al.* Revealing long noncoding RNA architecture and functions using domain-specific chromatin isolation by RNA purification. *Nat Biotechnol* **32**, 933-940 (2014).
31. Consortium, G.T. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-585 (2013).
32. Wheeler, T.M. *et al.* Targeting nuclear RNA for in vivo correction of myotonic dystrophy. *Nature* **488**, 111-115 (2012).
33. Lennox, K.A. & Behlke, M.A. Cellular localization of long non-coding RNAs affects silencing by RNAi more than by antisense oligonucleotides. *Nucleic Acids Res* **44**, 863-877 (2016).
34. Levitt, N., Briggs, D., Gil, A. & Proudfoot, N.J. Definition of an efficient synthetic poly(A) site. *Genes Dev* **3**, 1019-1025 (1989).
35. Tirian, L. *et al.* TPPP/p25 promotes tubulin assemblies and blocks mitotic spindle formation. *Proc Natl Acad Sci U S A* **100**, 13976-13981 (2003).
36. Tokesi, N. *et al.* TPPP/p25 promotes tubulin acetylation by inhibiting histone deacetylase 6. *J Biol Chem* **285**, 17896-17906 (2010).
37. DeBonis, S., Neumann, E. & Skoufias, D.A. Self protein-protein interactions are involved in TPPP/p25 mediated microtubule bundling. *Sci Rep* **5**, 13242 (2015).
38. Otzen, D.E. *et al.* p25alpha is flexible but natively folded and binds tubulin with oligomeric stoichiometry. *Protein Sci* **14**, 1396-1409 (2005).
39. Webster, D.R. & Borisy, G.G. Microtubules are acetylated in domains that turn over slowly. *J Cell Sci* **92 ( Pt 1)**, 57-65 (1989).
40. Akhmanova, A. & Steinmetz, M.O. Control of microtubule organization and dynamics: two ends in the limelight. *Nat Rev Mol Cell Biol* **16**, 711-726 (2015).
41. Acevedo, K. *et al.* The phosphorylation of p25/TPPP by LIM kinase 1 inhibits its ability to assemble microtubules. *Exp Cell Res* **313**, 4091-4106 (2007).
42. Goff, L.A. & Rinn, J.L. Linking RNA biology to lncRNAs. *Genome Res* **25**, 1456-1465 (2015).
43. Ballarino, M. *et al.* Deficiency in the nuclear long noncoding RNA Charme causes myogenic defects and heart remodeling in mice. *EMBO J* **37** (2018).
44. Wang, L. *et al.* LncRNA Dum interacts with Dnmts to regulate Dppa2 expression during myogenic differentiation and muscle regeneration. *Cell Res* **25**, 335-350 (2015).

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

45. Gilbert, L.A. *et al.* Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell* **159**, 647-661 (2014).
46. Skene, P.J., Henikoff, J.G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc* **13**, 1006-1019 (2018).
47. Simon, M.D. *et al.* The genomic binding sites of a noncoding RNA. *Proc Natl Acad Sci U S A* **108**, 20497-20502 (2011).
48. Simon, M.D. *et al.* High-resolution Xist binding maps reveal two-step spreading during X-chromosome inactivation. *Nature* **504**, 465-469 (2013).
49. Vance, K.W. *et al.* The long non-coding RNA Paupar regulates the expression of both local and distal genes. *EMBO J* **33**, 296-311 (2014).
50. West, J.A. *et al.* The long noncoding RNAs NEAT1 and MALAT1 bind active chromatin sites. *Mol Cell* **55**, 791-802 (2014).
51. Vance, K.W. Mapping Long Noncoding RNA Chromatin Occupancy Using Capture Hybridization Analysis of RNA Targets (CHART). *Methods Mol Biol* **1468**, 39-50 (2017).
52. Engreitz, J.M., Ollikainen, N. & Guttman, M. Long non-coding RNAs: spatial amplifiers that control nuclear structure and gene expression. *Nat Rev Mol Cell Biol* **17**, 756-770 (2016).
53. Hacisuleyman, E. *et al.* Topological organization of multichromosomal regions by the long intergenic noncoding RNA Firre. *Nat Struct Mol Biol* **21**, 198-206 (2014).
54. Maass, P.G. *et al.* A misplaced lncRNA causes brachydactyly in humans. *J Clin Invest* **122**, 3990-4002 (2012).
55. Heng, Y.W. *et al.* TPPP acts downstream of RhoA-ROCK-LIMK2 to regulate astral microtubule organization and spindle orientation. *J Cell Sci* **125**, 1579-1590 (2012).
56. Lehotzky, A. *et al.* Tubulin polymerization-promoting protein (TPPP/p25) is critical for oligodendrocyte differentiation. *Glia* **58**, 157-168 (2010).
57. Hlavanda, E. *et al.* Phosphorylation blocks the activity of tubulin polymerization-promoting protein (TPPP): identification of sites targeted by different kinases. *J Biol Chem* **282**, 29531-29539 (2007).
58. Atianand, M.K. *et al.* A Long Noncoding RNA lincRNA-EPS Acts as a Transcriptional Brake to Restrain Inflammation. *Cell* **165**, 1672-1685 (2016).
59. Ntini, E. *et al.* Long ncRNA A-ROD activates its target gene DKK1 at its release from chromatin. *Nat Commun* **9**, 1636 (2018).
60. Mino, R.E. *et al.* Drosophila Ringmaker regulates microtubule stabilization and axonal extension during embryonic development. *J Cell Sci* **129**, 3282-3294 (2016).
61. Dickinson, M.E. *et al.* High-throughput discovery of novel developmental phenotypes. *Nature* **537**, 508-514 (2016).
62. Olah, J. *et al.* Interactions of pathological hallmark proteins: tubulin polymerization promoting protein/p25, beta-amyloid, and alpha-synuclein. *J Biol Chem* **286**, 34088-34100 (2011).
63. Olah, J. & Ovadi, J. Dual life of TPPP/p25 evolved in physiological and pathological conditions. *Biochem Soc Trans* **42**, 1762-1767 (2014).
64. Stojic, L. *et al.* Transcriptional silencing of long noncoding RNA GNG12-AS1 uncouples its transcriptional and product-related functions. *Nat Commun* **7**, 10406 (2016).

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

65. Barr, A.R., Kilmartin, J.V. & Gergely, F. CDK5RAP2 functions in centrosome to spindle pole attachment and DNA damage response. *J Cell Biol* **189**, 23-39 (2010).
66. Liao, Y., Smyth, G.K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923-930 (2014).
67. Ritchie, M.E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
68. Law, C.W., Chen, Y., Shi, W. & Smyth, G.K. voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).
69. Phipson, B., Lee, S., Majewski, I.J., Alexander, W.S. & Smyth, G.K. Robust hyperparameter estimation protects against hypervariable genes

and improves power to detect differential expression. *Annals of Applied Statistics* **10**, 946-963 (2016).
70. McCarthy, D.J. & Smyth, G.K. Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics* **25**, 765-771 (2009).
71. Berger, R.L. & Hsu, J.C. Bioequivalence trials, intersection-union tests and equivalence confidence sets. *Stat Sci* **11**, 283-302 (1996).
72. Ernst, C., Eling, N., Martinez-Jimenez, C.P., Marioni, J.C. & Odom, D.T. Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. *Nat Commun* **10**, 1251 (2019).
73. Liao, Y., Smyth, G.K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res* **41**, e108 (2013).
74. Lun, A.T. & Smyth, G.K. From reads to regions: a Bioconductor workflow to detect differential binding in ChIP-seq data. *F1000Res* **4**, 1080 (2015).
75. Robinson, M.D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11**, R25 (2010).
76. Lun, A.T., Chen, Y. & Smyth, G.K. It's DE-licious: A Recipe for Differential Expression Analyses of RNA-seq Experiments Using Quasi-Likelihood Methods in edgeR. *Methods Mol Biol* **1418**, 391-416 (2016).
77. Lun, A.T. & Smyth, G.K. De novo detection of differentially bound regions for ChIP-seq data using peaks and windows: controlling error rates correctly. *Nucleic Acids Res* **42**, e95 (2014).
78. Hahne, F. & Ivanek, R. Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol* **1418**, 335-351 (2016).
79. Ran, F.A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nature protocols* **8**, 2281-2308 (2013).
80. Richardson, C.D., Ray, G.J., DeWitt, M.A., Curie, G.L. & Corn, J.E. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol* **34**, 339-344 (2016).
81. Zyss, D., Ebrahimi, H. & Gergely, F. Casein kinase I delta controls centrosome positioning during T cell activation. *J Cell Biol* **195**, 781-797 (2011).
82. Kong, L. *et al.* CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res* **35**, W345-349 (2007).
83. Wang, L. *et al.* CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res* **41**, e74 (2013).

bioRxiv preprint doi: https://doi.org/10.1101/709030; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

84. Lin, M.F., Jungreis, I. & Kellis, M. PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. *Bioinformatics* **27**, i275-282 (2011).