Article

Nucleolar RNA polymerase II drives ribosome biogenesis

Check for updates

Karan J. Abraham[1,11], Negin Khosraviani[1,11], Janet N. Y. Chan[1], Aparna Gorthi[2], Anas Samman[1], Dorothy Y. Zhao[3,4], Miling Wang[5], Michael Bokros[5], Elva Vidya[1], Lauren A. Ostrowski[1], Roxanne Oshidari[1], Violena Pietrobon[1], Parasvi S. Patel[6], Arash Algouneh[1,6], Rajat Singhania[6], Yupeng Liu[1], V. Talya Yerlici[1], Daniel D. De Carvalho[6], Michael Ohh[1,7], Brendan C. Dickson[1,8], Razq Hakem[6], Jack F. Greenblatt[3,4], Stephen Lee[5], Alexander J. R. Bishop[2,9] & Karim Mekhail[1,10 ✉]

Proteins are manufactured by ribosomes—macromolecular complexes of protein and RNA molecules that are assembled within major nuclear compartments called nucleoli[1,2]. Existing models suggest that RNA polymerases I and III (Pol I and Pol III) are the only enzymes that directly mediate the expression of the ribosomal RNA (rRNA) components of ribosomes. Here we show, however, that RNA polymerase II (Pol II) inside human nucleoli operates near genes encoding rRNAs to drive their expression. Pol II, assisted by the neurodegeneration-associated enzyme senataxin, generates a shield comprising triplex nucleic acid structures known as R-loops at intergenic spacers flanking nucleolar rRNA genes. The shield prevents Pol I from producing sense intergenic noncoding RNAs (sincRNAs) that can disrupt nucleolar organization and rRNA expression. These disruptive sincRNAs can be unleashed by Pol II inhibition, senataxin loss, Ewing sarcoma or locus-associated R-loop repression through an experimental system involving the proteins RNaseH1, eGFP and dCas9 (which we refer to as 'red laser'). We reveal a nucleolar Pol-II-dependent mechanism that drives ribosome biogenesis, identify disease-associated disruption of nucleoli by noncoding RNAs, and establish locus-targeted R-loop modulation. Our findings revise theories of labour division between the major RNA polymerases, and identify nucleolar Pol II as a major factor in protein synthesis and nuclear organization, with potential implications for health and disease.

Various proteins self-organize via liquid–liquid phase separation (LLPS) into nucleolar subdomains, which are needed for highly stereotyped ribosome assembly[1,2]. At fibrillar centres in the heart of mammalian nucleoli, the major rRNA molecules needed to assemble ribosomes are generated by Pol-I-dependent transcription of rRNA genes within ribosomal DNA (rDNA) repeats[1,3]. Within rDNA, rRNA genes are separated by large intergenic spacers (IGSs) (Extended Data Fig. 1a). At nucleolar rRNA genes, Pol I synthesizes precursor rRNAs (pre-rRNAs) that are processed into mature 28S, 18S and 5.8S rRNA molecules as they migrate to the granular component at the nucleolar periphery. Outside nucleoli, Pol III synthesizes 5S rRNA molecules that are targeted to nucleoli for processing. Mature rRNAs are packaged into 40S and 60S ribosomal subunits for export to the cytoplasm. Traditionally, the nucleolar Pol I and nucleoplasmic Pol III are viewed as the sole mammalian RNA polymerases that directly mediate housekeeping ribosome biogenesis. Interestingly, in the budding yeast *Saccharomyces cerevisiae*, Pol II is

physically enriched at rDNA IGSs, but this phenomenon is deleterious because it drives ageing without affecting rRNA expression[3–5]. It is unclear whether nucleolar Pol II exists in higher organisms or directly promotes ribosome biogenesis in any species.

Active Pol II at rDNA IGSs

To investigate whether Pol II exists within human nucleoli, we first used immunofluorescence coupled to super-resolution microscopy. Within nucleoli, which were outlined by nucleophosmin (NPM), we observed foci corresponding to active Pol II phosphorylated on serine 2 (pS2) (Fig. 1a and Extended Data Fig. 1b, c). Chromatin immunoprecipitation (ChIP) showed that pS2 and another active form of Pol II, phosphorylated on serine 5 (pS5), were enriched across rDNA, with the highest levels—at IGS28 and IGS38—being comparable to those at known Pol-II-transcribed loci (Fig. 1b and Extended Data Fig. 1a, d–f).

[1]Department of Laboratory Medicine and Pathobiology, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada. [2]Greehey Children's Cancer Research Institute, Department of Cell Systems and Anatomy, University of Texas Health at San Antonio, San Antonio, TX, USA. [3]Department of Molecular Genetics, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada. [4]Donnelly Centre, University of Toronto, Toronto, Ontario, Canada. [5]Sylvester Comprehensive Cancer Center, Department of Biochemistry and Molecular Biology, Miller School of Medicine, University of Miami, Miami, FL, USA. [6]Princess Margaret Cancer Research Centre, University Health Network, and Department of Medical Biophysics, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada. [7]Department of Biochemistry, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada. [8]Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada. [9]Mays Cancer Center, University of Texas Health at San Antonio, San Antonio, TX, USA. [10]Canada Research Chairs Program, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada. [11]These authors contributed equally: Karan J. Abraham, Negin Khosraviani. ✉e-mail: karim.mekhail@utoronto.ca
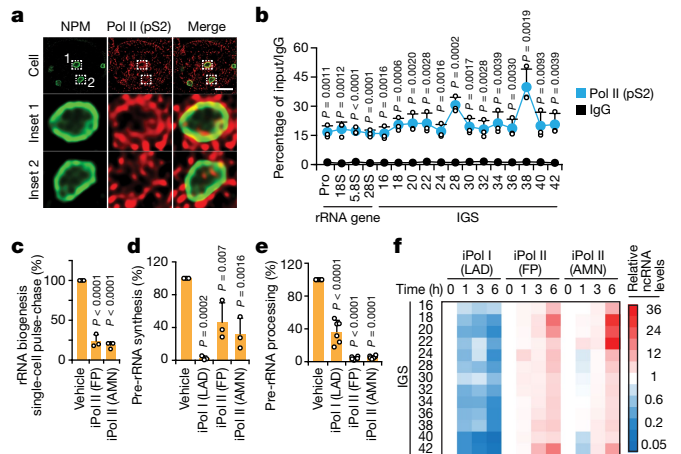
Fig. 1 | Pol I and Pol II localize to rDNA IGSs and compete to modulate IGS ncRNA levels. a, Representative immunofluorescence and super-resolution microscopy images showing the localization of pS2 Pol II within NPM-delineated nucleoli. Scale bar, 5 μm. **b**, Enrichment of pS2 Pol II across rDNA as revealed by chromatin immunoprecipitation (ChIP). Enrichments, accounting for typical background fluctuations across repetitive DNA loci, were calculated as (percentage of input/IgG) = (percentage of input for protein immunoprecipitation)/(percentage of input for mock IgG immunoprecipitation). **c**, Effect of a 3-hour Pol II inhibition (iPol II) using flavopiridol (FP) or α-amanitin (AMN) on rRNA biogenesis as measured in live single-cell pulse-chase assays using 5-fluorouracil (FU)-labelled RNA. **d**, **e**, Cell-population-based RNA pulse-chase assays were used to assess pre-rRNA synthesis (**d**) and processing (**e**) following a 3-hour inhibition of Pol I or Pol II (iPol I/II; low-dose actinomycin-D, LAD). **f**, Pol I promotes, and Pol II represses, IGS ncRNAs, as shown by reverse transcription with quantitative polymerase chain reaction (RT–qPCR). **a**–**f**, Experiments carried out with HEK293T cells; data shown as means ± s.d.; data in **b** and Extended Data Fig. 1d–f, j–l were from large experimental sets sharing immunoglobulin G (IgG) controls; n = 3 biologically independent experiments (**b**); two-tailed t-test (**b**); one-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test (**c**–**e**); image in **a** is representative of two independent experiments.

The Pol II activator cyclin-dependent kinase 9 (CDK9) was similarly enriched across IGSs (Extended Data Fig. 1g). pS2 and CDK9 were also enriched across the IGSs of IMR90 fibroblasts, indicating that enrichments are not limited to tumorigenic cells (Extended Data Fig. 1h, i). Unlike Pol II and CDK9, Pol I and its initiation factor, upstream binding factor 1 (UBF, also known as UBF1), localized primarily to rRNA genes, although low Pol I levels existed across IGSs (Extended Data Fig. 1j, k). Notably, Pol II was overrepresented relative to Pol I only within IGSs (Extended Data Fig. 1l). These data suggest that rDNA loci are cohabited by Pol I and Pol II.

To determine whether rRNA biogenesis is rapidly affected following Pol II perturbation, we conducted a three-hour treatment using the Pol II inhibitors α-amanitin (AMN) or flavopiridol in pulse-chase experiments. Pol II inhibition perturbed global ribosome biogenesis (Fig. 1c). Specifically, unlike Pol I inhibition by low-dose actinomycin-D (LAD), Pol II inhibition almost fully abolished pre-rRNA processing (Fig. 1d, e and Extended Data Fig. 2a–c), indicating a distinct mechanism of ribosome biogenesis arrest. Cell viability and global protein levels were unchanged following Pol II inhibition, arguing against indirect effects (Extended Data Fig. 2d, e). Furthermore, a 30-min Pol II inhibition was sufficient to strongly disrupt rRNA processing, suggesting a direct function for Pol II through its enrichment at rDNA (Extended Data Fig. 2f; Pol II inhibition hereafter was for three hours unless otherwise indicated). These data suggest that Pol II might directly support nucleolar rRNA expression through its association with IGSs.

In different cell types, we detected IGS noncoding RNAs (ncRNAs) that decreased in abundance following Pol I inhibition (Fig. 1f and
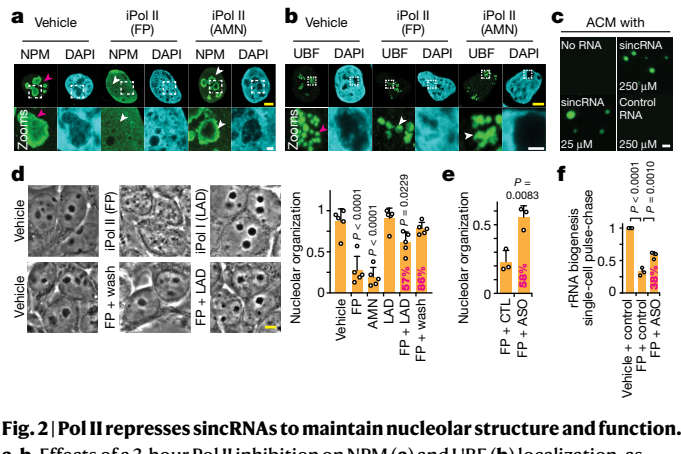
Fig. 2 | Pol II represses sincRNAs to maintain nucleolar structure and function. a, **b**, Effects of a 3-hour Pol II inhibition on NPM (**a**) and UBF (**b**) localization, as shown by immunofluorescence microscopy. Examples of normal and defective phenotypes are respectively marked by magenta and white arrowheads. DAPI, 4′,6-diamidino-2-phenylindole. **c**, Low-complexity sincRNA, not high-complexity control RNA, promoted the formation of liquid droplets in the presence of amyloid-converting motif (ACM) peptides in vitro. Shown on the images is the concentration of ACM peptides incubated with 1 μM of the indicated RNA. **d**–**f**, In cells subjected to Pol II inhibition (FP), nucleolar organization was restored by coinhibition of Pol I (LAD; **d**), removal of FP (wash; **d**), or treatment with sincRNA-repressing ASOs (**e**), which also restored rRNA biogenesis as indicated by live single-cell FU-RNA pulse-chase assays (**f**). Percentages indicating phenotypic rescue relative to FP-treated cells are shown on graphs as applicable. **a**–**f**, Experiments with HEK293T cells; data are shown as means ± s.d.; one-way ANOVA with Dunnett's multiple comparisons test (**d**, **f**); two-tailed t-test (**e**); n = 5 biologically independent experiments (**d**); n = 3 biologically independent experiments (**e**, **f**); images in **a**–**c** are representative of two independent experiments; scale bars, 5 μm (yellow) and 1 μm (white).

Extended Data Fig. 2g, h). Strikingly, IGS ncRNAs were markedly induced and found to be de novo transcribed upon Pol II inhibition (Fig. 1f and Extended Data Fig. 2h, i). Simultaneous inhibition of Pol I abolished the induction of IGS ncRNAs by Pol II inhibition (Extended Data Fig. 2j, k). Thus, Pol II counters Pol-I-dependent synthesis of IGS ncRNAs. Strand-specific transcript analysis of IGSs identified sense intergenic ncRNAs (sincRNAs) and antisense intergenic ncRNAs (asincRNAs) that were transcribed by Pol I and Pol II, respectively (Extended Data Fig. 2l–n). The sincRNA/asincRNA ratio paralleled Pol I/Pol II enrichment across IGSs (Extended Data Fig. 2m, o). The data so far indicate that Pol II operates directly across the IGSs, where it generates asincRNAs and limits the spurious synthesis of sincRNAs by Pol I.

Pol II maintains nucleoli via sincRNA control

Given that nucleolar organization is essential for rRNA synthesis and processing, we characterized disordered proteins at the nucleolar subdomains that are essential for these functions (Extended Data Fig. 3a, b). NPM delineates the granular component of the nucleolus, the LLPS of which is required for rRNA processing[2,6,7]. Pol II inhibition abrogated the phase separation of NPM, which was quickly reorganized into ruffled bodies before undergoing complete mixing with the nucleoplasm (Fig. 2a and Extended Data Fig. 3c–e). At nucleolar fibrillar centres, UBF (which is enriched at the promoters of rRNA genes) forms small foci[6]. Pol II inhibition resulted in UBF relocation to the nucleolar periphery, where UBF formed large spheres, rings or crescent-shaped bodies exhibiting wetting behaviour (Fig. 2b and Extended Data Fig. 3f–h). Changes in NPM and UBF coincided with global nucleolar disorganization (Extended Data Fig. 3i) and matched sincRNA induction kinetics (Fig. 1f). UBF bodies generated upon Pol II inhibition exhibited greater fluorescence recovery after photobleaching

(FRAP; Extended Data Fig. 3j)[6], suggesting decreased UBF–rDNA interactions or rDNA relocation to less viscous environments. Notably, the former nucleolar space that became surrounded with UBF signals following Pol II inhibition showed positive staining with Congo red (Extended Data Fig. 3k), indicating the presence of stress-induced, solid-like nucleolar amyloid bodies[8,9]. The data suggest that Pol II inhibition partly and strongly disrupts the organization of rRNA synthesis and processing sites, respectively. Under these conditions, aberrant liquid-to-solid phase transitions occur within the remnant nucleolar space.

Nucleolar amyloid bodies usually emerge following environmental stresses such as heat shock[6,8,9]. Specifically, heat shock causes proteins with the amyloid-converting motif (ACM) to form nucleolar liquid droplets, which undergo phase transition into solid-like amyloid bodies (Extended Data Fig. 4a, b). Knockdown of different sincRNAs prevented heat-shock-induced formation of ACM-containing nucleolar liquid droplets in vivo (Extended Data Fig. 4c). In a cell-free in vitro system, incubating ACM peptides with a sincRNA segment induced liquid droplet formation (Fig. 2c and Extended Data Fig. 4d)[9]. Moreover, strand-specific RNA sequencing (ss-RNA-seq) revealed that heat shock induced sincRNA and repressed asincRNA levels at IGSs (Extended Data Fig. 4e). Thus, environmental stress represses asincRNA levels and promotes sincRNA-dependent nucleolar remodelling. The results also show that sincRNAs induce liquid droplets in vitro and promote liquid droplets and consequent solid-like amyloid bodies in vivo.

Next, we assessed whether sincRNA repression restores nucleolar organization and function in live cells subjected to Pol II inhibition. Nucleolar organization was restored after Pol II inhibitor wash-off, Pol I co-inhibition, or direct repression of sincRNA levels with antisense oligonucleotides (ASOs) (Fig. 2d, e and Extended Data Fig. 5a). ASOs also partly restored rRNA biogenesis (Fig. 2f). An overexpressed sincRNA localized to nucleoli without decreasing rRNA biogenesis (Extended Data Fig. 5b–d), indicating that nucleolar disruption may depend on specific combinations of sincRNAs or that endogenous sincRNAs have distinctive modifications or interactors. However, cell types with naturally elevated sincRNA levels exhibited more NPM-marked nucleoli (Extended Data Figs. 2g, 5e). Of note, long-term Pol II inhibition may compromise nucleoli indirectly, by limiting the ability of Pol II to synthesize the U8 small nucleolar RNA (snoRNA) or Alu RNA molecules[10–12]. However, following our short-term Pol II inhibition, nucleolar disruption coincided with sincRNA induction in the absence of changes in U8 or Alu levels (Extended Data Fig. 6a, b). Additionally, in contrast with Pol II inhibitors, pharmacological agents[2,12] disrupting nucleolar organization or global protein translation failed to induce sincRNA levels (Extended Data Fig. 6c–e). Thus, sincRNA accumulation drives nucleolar disorganization, and not vice versa. Together, these results show that Pol II constitutively represses different Pol-I-dependent sincRNAs to prevent unscheduled stress-mimicking nucleolar phase transitions, and to maintain endogenous nucleolar condensates that are essential for rRNA biogenesis.

Pol II sets an R-loop shield for Pol I

Nucleoli are naturally enriched in R-loops, which are triplex nucleic acid structures harbouring a DNA–RNA hybrid and single-stranded DNA[13]. Therefore, we postulated that baseline R-loop levels across IGSs may have beneficial effects through the modulation of Pol I–Pol II crosstalk. DNA–RNA hybrid immunofluorescence (DRIF) revealed nucleolar R-loops that were partly repressed by Pol II inhibition (Fig. 3a) or the recombinant DNA–RNA hybrid repressor RNase H1 (Extended Data Fig. 7a–c). DNA–RNA hybrid immunoprecipitation (DRIP) revealed that several IGS sites exhibited R-loop signals, which peaked at the junctions between rRNA genes and IGSs and were sensitive to RNase H1 (Fig. 3b and Extended Data Fig. 7d)[14]. Despite markedly higher transcription of rRNA genes relative to IGSs, negative GC skews may be one
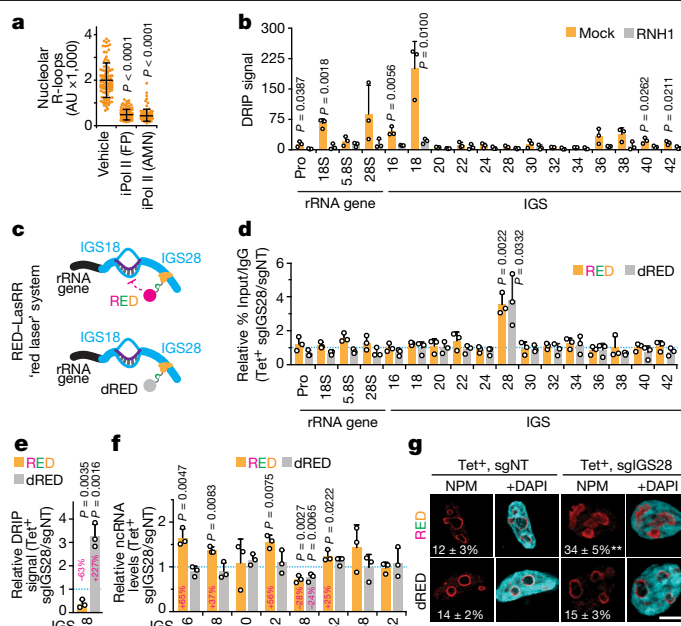
Fig. 3 | Repression of an IGS R-loop shield disrupts nucleoli. a, Pol II inhibition repressed nucleolar R-loops. **b**, DRIP analysis shows RNase H1-sensitive R-loop peaks at rDNA. **c**, The RED–LasRR system created to achieve inducible locus-associated R-loop repression. **d**, The short guide RNA for IGS28 (sgIGS28) enriched the tetracycline (Tet[+])-induced RED or dRED at IGS28 in anti-GFP ChIP, using IgG as control. Enrichments are normalized to a non-targeting control (sgNT). RED and dRED data were from different experiments but are shown on the same graph as a space-saving measure. **e**, Using RED or dRED together with sgIGS28 respectively decreased or increased R-loop levels at IGS18. **f, g**, RED sgIGS28 induced ncRNA levels (**f**) and disrupted NPM localization (**g**). The percentages of cells exhibiting ruffled NPM localization are indicated on the images (**g**). **a–g**, HEK293T cells; data are shown as means ± s.d.; two-tailed Mann–Whitney U-test, $n = 100$ cells (**a**); or two-tailed t-test, $n = 3$ biologically independent experiments (**b, d–f**); scale bar, 5 μm. Percentage changes relative to respective sgNT samples are indicated above or on bars (**e, f**).

of several different factors favouring antisense IGS R-loops (Extended Data Fig. 7e)[15]. Notably, R-loop repression by RNase H1 overexpression partly mimicked Pol II inhibition, increasing sincRNA expression at most IGS sites tested (Extended Data Fig. 7f, g). Together, these findings suggest that R-loops are important molecular mediators of sincRNA repression by Pol II.

RNase H1 overexpression remains the gold-standard method by which to interrogate R-loop function[13]. However, with this approach, RNase H1 is often not enriched at the studied loci, where the observed phenotypic changes may also be due to R-loop repression elsewhere. To specifically interrogate the function of IGS-associated R-loops, we created a tetracycline-inducible RNase H1–eGFP–dCas9 (RED) fusion protein to achieve locus-associated R-loop repression (a process that we abbreviate as 'RED–LasRR', or 'red laser'; Fig. 3c and Extended Data Fig. 7h; eGFP is enhanced green fluorescent protein). As a control, this system uses a similar chimaeric protein that comprises catalytically dead RNase H1 (denoted dRED).

Similar to the RNase H1 protein[16], RED and dRED displayed nucleolar and nucleoplasmic localization in the absence of short guide RNAs (sgRNAs) (Extended Data Fig. 7i, j). Within the IGS, constitutive chromatin looping juxtaposes the IGS27/28 sites with IGS16/18 sites[17,18]. Therefore, we investigated whether a pool of three sgRNAs targeting IGS28 (sgIGS28) can enrich RED at IGS28 and repress the strong R-loop peaks at IGS16/18. ChIP confirmed successful targeting and similar enrichment of RED and dRED at the IGS28 site upon coexpression of sgIGS28 (Fig. 3d). Targeting RED, but not dRED, to IGS28 repressed only

the strong R-loop peak at IGS18, while inducing a subset of sincRNAs across the IGSs (Fig. 3e–f and Extended Data Fig. 7k). Using RED with sgIGS38, which is spatially distal to the IGS18 site[17,18], failed to alter R-loop or ncRNA levels at IGS18 (Extended Data Fig. 7l, m). Targeting dRED to IGS28 stabilized R-loops without decreasing sincRNA levels at IGS18, suggesting that maximal function of IGS18 R-loops is already achieved endogenously (Fig. 3e, f). Of note, ncRNA levels were similarly decreased at the IGS28 site to which RED or dRED was targeted without affecting Pol II enrichments (Fig. 3f and Extended Data Fig. 7n), and the RED–LasRR system can be used to target the fusion proteins to a single-copy locus outside of rDNA (Extended Data Fig. 7o). Using the guide RNAs (gRNAs) targeting RED to IGS28, individually, failed to achieve R-loop repression at the IGSs (Extended Data Fig. 7p). This argues against the possibility that targeting of RED to non-rDNA sites via any single gRNA or the RNase H1 moiety of the fusion protein indirectly represses IGS R-loops. Although the RED/sgIGS28-dependent lowering of R-loops only partially induces sincRNAs, this still mimicked early Pol II inhibition, as shown by the perturbation of NPM architecture into indistinct, ruffled bodies (Fig. 3f, g). This highlights the disruptive impact that even small increases in sincRNA levels can exert on nucleoli. The data show that asincRNAs generated by Pol II form an antisense R-loop shield that limits the synthesis of Pol-I-dependent sincRNAs, which can abrogate nucleolar organization and function. The RED–LasRR system will support studies on the numerous roles of R-loops in genome expression and stability.

Senataxin supports the R-loop shield

We next set out to identify additional factors that may regulate nucleolar Pol II. Senataxin (SETX) is a human neurodegeneration-linked helicase[19]. SETX and its yeast orthologue Sen1 have several transcription-modulatory roles, including Pol II loading and R-loop repression[20–22]. Sen1 associates with rDNA IGSs to promote Pol I transcription termination and to silence lifespan-shortening IGS ncRNAs[16,23,24]. We found that SETX was enriched across human IGSs, especially at IGS28, and exhibited nucleolar localization (Extended Data Fig. 8a, b). The IGS28 SETX peak overlapped one Pol II peak and the intergenic promoter marks H3K27ac, H3K9ac and H3K4me3 in ENCODE ChIP–seq data (Extended Data Fig. 8c). Sequential ChIP revealed that SETX was preferentially coenriched with Pol II compared with Pol I at IGS28 (Fig. 4a). Thus, SETX is coenriched with Pol II at IGSs, especially at a putative intergenic promoter at IGS28.

Notably, SETX knockout decreased the intergenic enrichment of Pol II and its R-loops (Fig. 4b and Extended Data Fig. 8d, e). This change was accompanied by increased intergenic Pol I enrichment (Extended Data Fig. 8f), elevated sincRNA synthesis (Fig. 4c and Extended Data Fig. 8g), and decreased Pol I localization at rRNA genes (Extended Data Fig. 8h). Unlike SETX knockout, the forced release of Pol I from rRNA gene promoters through knockdown of transcription initiation factor 1A (TIF1A) decreased pre-rRNA levels without inducing sincRNA levels (Extended Data Fig. 8i, j). This suggests that SETX loss prevents Pol II from shielding the IGSs from de novo Pol I loading. In addition, northern blotting did not show increases in pre-rRNA length upon Pol II or SETX disruption, arguing against rRNA gene read-through as the basis for increased IGS transcription by Pol I (Extended Data Fig. 8k). Thus, IGS R-loops act more as a shield that prevents Pol I recruitment, rather than a barrier that limits read-through transcription. Increases in sincRNA levels in SETX-knockout cells were associated with nucleolar disorganization and pre-rRNA processing defects, which were partly countered by sincRNA knockdown (Fig. 4d, e and Extended Data Fig. 8l–n). That SETX loss partly mimicked Pol II inhibition probably reflects the partial coenrichment of SETX and Pol II at IGSs. Additionally, SETX knockout did not lower IGS epigenetic silencing marks (Extended Data Fig. 8o), suggesting that SETX loss does not promote sincRNA levels by abrogating epigenetic silencing. In fact, SETX knockout slightly increased
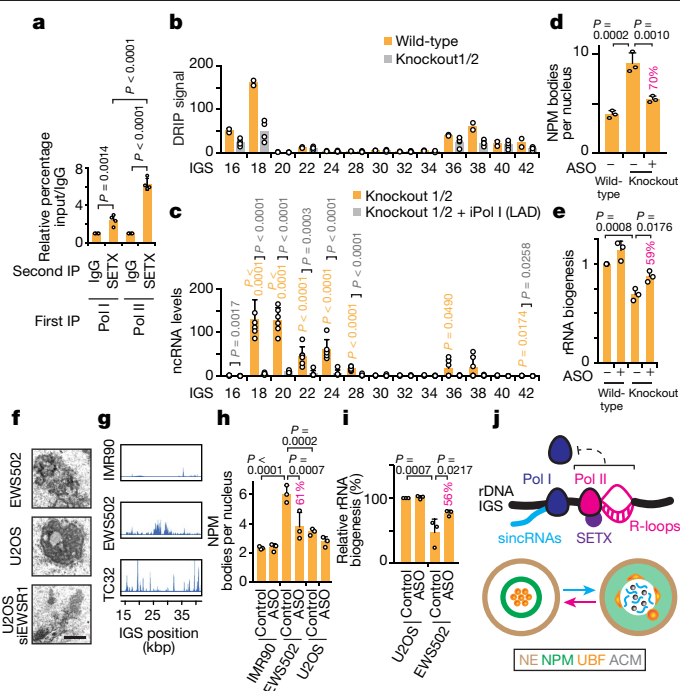
Fig. 4 | Nucleolar Pol II reinforcement by SETX and nucleolus-disrupting sincRNAs in cancer. a, Sequential immunoprecipitations (IPs) revealed preferential coenrichment of SETX with Pol II at IGSs. Signals for Pol I/SETX and Pol II/SETX immunoprecipitations are normalized to signals from Pol I/IgG and Pol II/IgG, respectively. **b**, **c**, SETX knockout in two clones decreased R-loops (**b**) and induced IGS ncRNAs (**c**). **d**, **e**, Single-cell analysis of SETX-knockout cells showed that ASO-mediated repression of Pol-I-dependent sincRNAs partly rescues nucleolar organization (**d**) and rRNA biogenesis (**e**). Percentages indicating the magnitude of ASO-mediated phenotypic rescue are shown above graph bars where applicable. **f**, The patient-derived Ewing sarcoma cell line EWS502 and U2OS osteosarcoma cells with siRNA-mediated depletion of EWS breakpoint region 1 (EWSR1) showed disrupted nucleoli by electron microscopy. **g**, RNA-seq data indicate increased ncRNA levels at the IGSs of EWS502 and TC32 cells, as compared with IMR90 control cells. kbp, kilobasepairs. **h**, **i**, Single-cell analysis showed that sincRNA knockdown partly restores nucleolar organization (**h**) and rRNA biogenesis (**i**) in EWS502 cells. **j**, Model showing how a Pol-II-dependent R-loop shield limits Pol-I-dependent sincRNAs, which compromise nucleolar organization and function. ACM, amyloid-converting motif; NE, nuclear envelope. **a**–**i**, Cells were HEK293T (**a**–**e**) or as indicated (**f**–**i**); data are shown as means ± s.d.; one-way ANOVA with Dunnett's multiple comparisons test (**a**, **d**, **e**, **h**, **i**) and one-way ANOVA with Tukey's multiple comparisons test (**c**); n = 4 biologically independent experiments (**a**), n = 2 biologically independent experiments (**b**, duplicates for each of wild-type, knockout 1 and knockout 2), n = 6 biologically independent experiments (**c**, triplicates for each knockout), n = 3 biologically independent experiments (**d**, **e**, **h**, **i**); images in **f**, **g** are representatives of two independent experiments; scale bar, 1 μm.

silencing marks, possibly reflecting epigenetic compensation constraining the magnitude of sincRNA induction. The data indicate that SETX is coenriched with IGS Pol II and supports it in repressing a subset of Pol-I-dependent sincRNAs that can disrupt nucleolar organization and function. SETX may achieve this effect by promoting the efficient loading and release of Pol II at an IGS28 intergenic promoter.

sincRNAs can disrupt nucleoli in cancer

We then aimed to identify a setting in which naturally elevated sincRNA levels may compromise nucleolar structure and function. Nucleolar organization, which is intimately related to cellular growth and viability, may be an adjunct in the diagnosis and treatment of some cancers[25]. In fact, nucleolar disruption upon Pol II dysregulation is similar to the

Article

constitutive disorganization of nucleoli in human Ewing sarcoma (EWS) tumours, related patient-derived EWS502 or TC32 cells, and U2OS osteosarcoma cells with depletion of EWS breakpoint region 1 (EWSR1) (Fig. 4f and Extended Data Fig. 9a, b)[26]. To determine whether alterations in sincRNA levels could underlie this phenotype, we reanalysed RNA-seq and DRIP–seq data from EWS and healthy IMR90 control cells to include rDNA[17,26]. EWS cells exhibited increased ncRNA and R-loop levels across IGSs (Fig. 4g and Extended Data Fig. 9c–e). Strikingly, in EWS cells, nucleolar disorganization and an rRNA biogenesis defect were countered by sincRNA knockdown (Fig. 4h, i and Extended Data Fig. 9f). These findings suggest that natural increases in sincRNA levels can explain aberrant nucleolar morphologies that are commonly observed in cancer[25]. R-loop increases in this setting may reflect selection for cells that have compensated for the increased sincRNA levels.

Discussion

Our findings indicate that, in mammalian cells, antisense transcription by nucleolar Pol II generates an R-loop shield at rDNA IGSs to block Pol-I-dependent sense intergenic transcripts, which can compromise nucleolar condensates underlying rRNA expression (Fig. 4j and Extended Data Fig. 10). Processes that restrain R-loops at human IGSs probably exist, as unrestricted IGS R-loops destabilize yeast rDNA[4,16,27]. However, our findings differ from those in yeast, where IGS transcription does not regulate rRNA[28,29] and Sen1 limits deleterious IGS ncRNAs by enforcing epigenetic silencing and transcript turnover[21,23]. At the IGSs of human cells under stress[8], protective sense RNAs are likely to be induced through local repression of antisense RNA and R-loops. Nucleolar Pol II at IGSs may also mediate crosstalk with cellular differentiation, which is partly driven by promoter-associated transcripts that are dependent on Pol I or Pol II[30–33]. Future work should explore the potential use of sincRNAs and nucleolar disorganization as cancer biomarkers, and whether tumours exhibiting such features are hypersensitive to Pol-II-inhibiting drugs[34]. Overall, we identify nucleolar Pol II as a new master regulator of ribosome biogenesis, with broad implications for health and disease.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2497-0.

1. Boisvert, F. M., van Koningsbruggen, S., Navascués, J. & Lamond, A. I. The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* **8**, 574–585 (2007).
2. Feric, M. et al. Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686–1697 (2016).
3. Mekhail, K. & Moazed, D. The nuclear envelope in genome organization, expression and stability. *Nat. Rev. Mol. Cell Biol.* **11**, 317–328 (2010).
4. Kobayashi, T. & Ganley, A. R. Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. *Science* **309**, 1581–1584 (2005).
5. Kaeberlein, M., McVey, M. & Guarente, L. The SIR2/3/4 complex and SIR2 alone promote longevity in *Saccharomyces cerevisiae* by two different mechanisms. *Genes Dev.* **13**, 2570–2580 (1999).
6. Mekhail, K. et al. Regulation of ubiquitin ligase dynamics by the nucleolus. *J. Cell Biol.* **170**, 733–744 (2005).
7. Mitrea, D. M. et al. Self-interaction of NPM1 modulates multiple mechanisms of liquid-liquid phase separation. *Nat. Commun.* **9**, 842 (2018).
8. Audas, T. E. et al. Adaptation to stressors by systemic protein amyloidogenesis. *Dev. Cell* **39**, 155–168 (2016).
9. Wang, M. *et al.* Stress-induced low complexity RNA activates physiological amyloidogenesis. *Cell Rep.* **24**, 1713–1721 (2018).
10. Caudron-Herger, M. et al. Alu element-containing RNAs maintain nucleolar structure and function. *EMBO J.* **34**, 2758–2774 (2015).
11. Burger, K. et al. Cyclin-dependent kinase 9 links RNA polymerase II transcription to processing of ribosomal RNA. *J. Biol. Chem.* **288**, 21173–21183 (2013).
12. Burger, K. et al. Chemotherapeutic drugs inhibit ribosome biogenesis at various levels. *J. Biol. Chem.* **285**, 12416–12425 (2010).
13. Santos-Pereira, J. M. & Aguilera, A. R loops: new modulators of genome dynamics and function. *Nat. Rev. Genet.* **16**, 583–597 (2015).
14. Bhatia, V. et al. BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature* **511**, 362–365 (2014).
15. Sanz, L. A. et al. Prevalent, dynamic, and conserved R-loop structures associate with specific epigenomic signatures in mammals. *Mol. Cell* **63**, 167–178 (2016).
16. Abraham, K. J. et al. Intersection of calorie restriction and magnesium in the suppression of genome-destabilizing RNA-DNA hybrids. *Nucleic Acids Res.* **44**, 8870–8884 (2016).
17. Zentner, G. E., Saiakhova, A., Manaenkov, P., Adams, M. D. & Scacheri, P. C. Integrative genomic analysis of human ribosomal DNA. *Nucleic Acids Res.* **39**, 4949–4960 (2011).
18. Shiue, C. N., Nematollahi-Mahani, A. & Wright, A. P. Myc-induced anchorage of the rDNA IGS region to nucleolar matrix modulates growth-stimulated changes in higher-order rDNA architecture. *Nucleic Acids Res.* **42**, 5505–5517 (2014).
19. Groh, M., Albulescu, L. O., Cristini, A. & Gromak, N. Senataxin: genome guardian at the interface of transcription and neurodegeneration. *J. Mol. Biol.* **429**, 3181–3195 (2017).
20. Skourti-Stathaki, K., Proudfoot, N. J. & Gromak, N. Human senataxin resolves RNA/DNA hybrids formed at transcriptional pause sites to promote Xrn2-dependent termination. *Mol. Cell* **42**, 794–805 (2011).
21. Steinmetz, E. J. et al. Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase. *Mol. Cell* **24**, 735–746 (2006).
22. Suraweera, A. et al. Functional role for senataxin, defective in ataxia oculomotor apraxia type 2, in transcriptional regulation. *Hum. Mol. Genet.* **18**, 3384–3396 (2009).
23. Vasiljeva, L., Kim, M., Terzi, N., Soares, L. M. & Buratowski, S. Transcription termination and RNA degradation contribute to silencing of RNA polymerase II transcription within heterochromatin. *Mol. Cell* **29**, 313–323 (2008).
24. Mischo, H. E. et al. Yeast Sen1 helicase protects the genome from transcription-associated instability. *Mol. Cell* **41**, 21–32 (2011).
25. Stamatopoulou, V., Parisot, P., De Vleeschouwer, C. & Lafontaine, D. L. J. Use of the iNo score to discriminate normal from altered nucleolar morphology, with applications in basic cell biology and potential in human disease diagnostics. *Nat. Protoc.* **13**, 2387–2406 (2018).
26. Gorthi, A. et al. EWS-FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. *Nature* **555**, 387–391 (2018); correction **559**, E11 (2018).
27. El Hage, A., French, S. L., Beyer, A. L. & Tollervey, D. Loss of topoisomerase I leads to R-loop-mediated transcriptional blocks during ribosomal RNA synthesis. *Genes Dev.* **24**, 1546–1558 (2010).
28. Kawauchi, J., Mischo, H., Braglia, P., Rondon, A. & Proudfoot, N. J. Budding yeast RNA polymerases I and II employ parallel mechanisms of transcriptional termination. *Genes Dev.* **22**, 1082–1092 (2008).
29. Salvi, J. S. et al. Roles for Pbp1 and caloric restriction in genome and lifespan maintenance via suppression of RNA-DNA hybrids. *Dev. Cell* **30**, 177–191 (2014).
30. Bierhoff, H. et al. Quiescence-induced LncRNAs trigger H4K20 trimethylation and transcriptional silencing. *Mol. Cell* **54**, 675–682 (2014).
31. Savić, N. et al. lncRNA maturation to initiate heterochromatin formation in the nucleolus is required for exit from pluripotency in ESCs. *Cell Stem Cell* **15**, 720–734 (2014).
32. Mayer, C., Schmitz, K. M., Li, J., Grummt, I. & Santoro, R. Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol. Cell* **22**, 351–361 (2006).
33. Zhao, Z., Sentürk, N., Song, C. & Grummt, I. lncRNA PAPAS tethered to the rDNA enhancer recruits hypophosphorylated CHD4/NuRD to repress rRNA synthesis at elevated temperatures. *Genes Dev.* **32**, 836–848 (2018).
34. Zhang, H. et al. Targeting CDK9 reactivates epigenetically silenced genes in cancer. *Cell* **175**, 1244–1258 (2018).

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment, except for the quantification of microscopy images.

Cell culture and general materials

Human HEK293T, HeLa, HAP1 and osteosarcoma (U2OS) cells were cultured in Dulbecco's modified Eagle medium (DMEM, Wisent Bioproducts) with 10% fetal bovine serum (FBS, Wisent). HEK293T T-REx cells (ThermoFisher Scientific) were cultured in DMEM supplemented with 10% tetracycline-free FBS and 1% penicillin/streptomycin. EWS502 and IMR90 cells were cultured in Roswell Park Memorial Institute (RPMI) medium supplemented with 10% FBS. All cell lines were cultured in the presence of 1% (v/v) penicillin/streptomycin (Wisent) at 37 °C in a humidified atmosphere with 5% CO_2. Transfection of cultured cells was achieved using Lipofectamine3000 (Invitrogen, catalogue number L3000008), Lipofectamine RNAiMAX (Invitrogen, catalogue number 13778150) and Polyjet DNA transfection reagent (SignaGen Laboratories, SL100688). For transfections with plasmids encoding GFP–UBF1 or RNaseH1, 70% confluent cells were transfected with 1–3 μg of plasmid per well of a six-well plate; pcDNA3 served as control for RNaseH1 overexpression. For Pol II inhibition, cells were treated either with the reversible inhibitor flavopiridol (2 μM, inhibits Pol II pS2; Santa Cruz catalogue number sc-202157) or with the irreversible inhibitor α-amanitin (AMN, 50 μg ml⁻¹, inhibits translocation; Abcam catalogue number ab144512). Other drug treatments were LAD (50 ng ml⁻¹), MG132 (10 μM), doxorubicin (Dox, 300 nM), camptothecin (CPT, 10 μM), cycloheximide (CHX, 100 μM) or 1,6-hexanediol (HEX, 0.1% v/v). Antibodies, primers, guide RNAs and northern probes are listed in the Supplementary Information (Supplementary Tables 1–4). For Ewing sarcoma analyses, the Ewing sarcoma cell line TC32 was procured from the Children's Oncology group (https://childrensoncologygroup.org/) and EWS502 was a kind gift from S. Lessnick (Nationwide Children's Hospital, OH). Both cell lines were grown in RPMI (Corning). The control cell lines IMR90 (a primary fibroblast cell line) and U2OS (a human osteosarcoma cell line) were purchased from the American Type Culture Collection (ATCC) and grown in DMEM (Corning). Media were supplemented with 10% heat-inactivated FBS (Atlanta Biologicals). Cells were maintained at 37 °C in a humidified atmosphere with 5% CO_2, confirmed using short tandem repeat (STR) profiling and tested for mycoplasma contamination. All siRNA transfections were conducted using Lipofectamine RNAimax (ThermoFisher Scientific) according to the manufacturer's protocols.

Chromatin immunoprecipitation

Cells were grown to 80% confluence in 15 cm plates and crosslinked by adding 1% (v/v) formaldehyde at room temperature for 10 min. The reaction was quenched with 125 mM glycine for 5 min at room temperature. Cells were washed twice with cold phosphate-buffered saline (PBS), lysed with 10 ml lysis buffer (5 mM PIPES, 85 mM KCl, 0.5% (v/v) NP-40, complete protease-inhibitor cocktail (Roche)), scraped into tubes, and incubated for 10 min on ice. Cells were then pelleted at 1,000 r.p.m. for 10 min at 4 °C and resuspended in 500 μl of nuclear lysis buffer (50 mM Tris-HCl, 10 mM EDTA, 10% (w/v) SDS, complete protease-inhibitor cocktail) and incubated on ice for 10 min. Lysates were sonicated eight times for 20 s each at 40% amplitude at 4 °C with intermittent incubations on ice for 2 min. Centrifugation at 12,500g for 10 min at 4 °C clarified lysates. We set aside 10 μl of sheared chromatin for each sample as input. We diluted 50 μl of chromatin at a 1/10 ratio in immunoprecipitation dilution buffer (16.7 mM Tris-HCl pH 8.0, 0.01% (w/v) SDS, 167 mM NaCl, 1.2 mM EDTA, 1.1% (v/v) Triton-X100, complete protease inhibitor) and incubated with 5 μg of antibody on a rotator overnight at 4 °C. Samples were then incubated at constant rotation with 25 μl of prewashed Dynabeads protein G (Life Technology, catalogue number 10004D) for 2 h at 4 °C. Beads were washed once with a low-salt wash buffer (20 mM Tris-HCl, 0.1% (w/v) SDS, 1% (v/v) Triton X-100, 2 mM EDTA, 150 mM NaCl), once with high-salt wash buffer (20 mM Tris-HCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 500 mM NaCl), once with LiCl wash buffer (10 mM Tris-HCl, 1% (v/v) NP-40, 1% (w/v) sodium deoxycholate, 1 mM EDTA, 250 mM LiCl), and twice with TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) before two rounds of incubation with 100 μl of elution buffer (1% SDS, 100 mM NaHCO₃) for 15 min at room temperature. The eluates were incubated with 8 μl of 5 M NaCl on a rotator at 65 °C overnight. We added 3 μl of 10 mg ml⁻¹ RNase A (ThermoFisher Scientific, catalogue number EN0531) and incubated samples first at room temperature for 30 min, and then with 4 μl of 0.5 M EDTA, 8 μl of 1 M Tris-HCl and 1 μl proteinase K (Roche, catalogue number 03115887001) at 45 °C for 2 h. DNA was purified using gel/PCR DNA-fragment extraction (Geneaid, catalogue number DF300) and diluted with 150 μl of TE buffer. Primers are listed in the reagents table included with the Supplementary Information. Following ChIP–qPCR analysis, ChIP enrichments, accounting for typical background fluctuations across repetitive DNA loci, were calculated as (Percentage of input/IgG) = (Percentage of input for protein immunoprecipitation)/ (Percentage of input for mock IgG immunoprecipitation). The mean IgG background is also shown on ChIP graphs (Fig. 1b and Extended Data Figs. 1d–k, 8a).

Sequential chromatin immunoprecipitation

Similar to regular ChIP, for sequential chromatin immunoprecipitation (ChIP–reChIP), cells were grown to 80% confluence, crosslinked and lysed. For the first round of immunoprecipitation, samples were diluted 1/10 in immunoprecipitation dilution buffer (100 μl chromatin plus 900 μl immunoprecipitation dilution buffer) and incubated with 5 μg of anti-Pol-I (anti-RPA135 subunit) or anti-Pol-II (anti-C-terminal domain (CTD)) on a rotator overnight at 4 °C. Samples were then incubated at constant rotation with 25 μl of pre-washed Dynabeads for 2 h at 4 °C. Similar to regular ChIP (see above), beads were washed once with low-salt wash buffer, once with high-salt wash buffer, once with LiCl wash buffer, and twice with TE buffer before one 30-min incubation with 50 μl elution buffer containing 10 mM DTT. Eluates from each of the first immunoprecipitation tubes corresponding to the same antibody were combined, diluted 20-fold in cold immunoprecipitation dilution buffer and incubated overnight at 4 °C with 5 μg of anti-senataxin (anti-SETX) antibody. Once again, beads were incubated at constant rotation with 25 μl of pre-washed Dynabeads for 2 h at 4 °C, washed once with low-salt wash buffer, once with high-salt wash buffer, once with LiCl wash buffer, and twice with TE buffer before two rounds of incubation with 100 μl of elution buffer for 15 min at room temperature, and overnight incubation at 65 °C with 8 μl of 5 M NaCl. Similar to regular ChIP, samples were treated with RNase A/proteinase K (Roche, catalogue number 03115887001) and purified; qPCR was then performed.

Quantitative PCR

Quantitative real-time PCR was performed using a Bio-Rad CFX Connect Real-Time. Ten microlitres of qPCR reactions each containing SensiFAST SYBR No-ROX kit (FroggaBio, catalogue number BIO-98050), 200 nM of each of the forward and reverse primers, and 1 μl of diluted complementary DNA, diluted input, diluted immunoprecipitation ChIP or diluted DRIP DNA depending on the experiment. PCR comprised one cycle of 95 °C for 5 min and 60 °C for 30 s, followed by 39 cycles of 95 °C for 5 s and 60 °C for 30 s, and a final melt curve of 65 °C to 95 °C in 0.5 °C steps at 5 s per step.

RNA extraction

Cells grown to 70–80% confluence were washed with RNase-free PBS before RNA isolation using a Qiagen RNeasy mini Kit (catalogue number 74104).

Article

Reverse transcription

For regular reverse transcription, 1 μg of total RNA was treated with 1 μl of 10× DNase-I reaction buffer and 1 μl of DNase I Amp grade (1 U μl^{-1}; ThermoFisher, catalogue number 18068015), and then incubated for 15 min at room temperature. The reaction was quenched with 1 μl of 25 mM EDTA and incubated for 10 min at 65 °C. We carried out 10 μl reverse-transcription reactions using 10 mM deoxynucleoside trisphosphate (dNTPs), 50 μM random nonamers (Sigma, catalogue number R7647), 500 ng total RNA, 5× first-strand buffer, 100 mM dithiothreitol (DTT), 40 U μl^{-11} RNaseOUT (Invitrogen, catalogue number 10777019) and 200 U μl^{-1} M-MLV reverse transcriptase (Invitrogen, catalogue number 28025013) at 25 °C for 10 min, 37 °C for 60 min and 70 °C for 15 min. For pre-rRNA pulse chase, an additional step comprising 5 min at 85 °C was added to release the RNA from beads. The reverse-transcription reaction was diluted 1:5, and 4 μl were used in qPCR amplification. For strand-specific (ss)RT–qPCR, 30 μg of total RNA was treated with DNase I (10 U DNase I in a 100 μl reaction) for 30 min at 37 °C. The reaction was stopped by adding 2 μl of 250 mM EDTA, pH 8.0, and incubating at 75 °C for 10 min. RNA was precipitated with 25 μl RNA precipitation solution (0.8 M trisodium citrate, 1.2 M NaCl) and 50 μl isopropanol. Samples were incubated for 10 min at room temperature and 20 min at −20 °C, and then centrifuged at 7,500g for 20 min at 4 °C. Supernatants were aspirated and pellets were air-dried for 10 min. Pellets were resuspended in 30 μl deionized, diethylpyrocarbonate-treated (ddDEPC) H$_2$O and incubated at 65 °C for 5 min. Concentrations of purified RNA were measured using NanoDrop.

We designed strand-specific primers to allow the detection of sense and antisense transcripts at the same locus as described previously[35]. Briefly, a primer of roughly 18 bp was designed to recognize the strand of interest (for example, a reverse primer to detect sense transcripts, a forward primer to detect antisense transcripts). A nonsense sequence (CGAGGATCATGGTGGCGAATAA) was added to tag the 5′-end of each strand-specific IGS primer. As a control within each reverse-transcription reaction, we generated a reverse primer to detect 7SK sense transcripts (we added a T7 sequence to the 5′ end of this primer to distinguish it from IGS primers). Separate reverse-transcription reactions were carried out for each transcript of interest. Each 10 μl reverse-transcription reaction contained 200 ng purified RNA, 5 μM strand-specific tagged primer (comprising roughly 18 bp specific to the transcript of interest, with the nonsense sequence CGAGGATCATGGTGGCGAATAA added to the 5′-end), 5 μM control sense primer (for example, 7SK), 1 mM dNTPs, 1× first-strand buffer, 10 mM DTT, 40 U RNaseOUT and 200 U of M-MLV reverse transcriptase. False-prime reactions were also carried out for each RNA sample and were conducted by replacing the transcript-of-interest primers with DEPC ddH$_2$O. Reactions were incubated at 25 °C for 10 min, 37 °C for 60 min, and 70 °C for 15 min. Resulting cDNA was diluted 1 in 10. Each cDNA sample represents one strand-specific transcript of interest and 7SK sense transcripts as a control. Each cDNA sample was amplified using primers directed at the strand-specific transcript of interest (using ss_Tag and hIGS_forward primers for sense transcripts or ss_Tag and hIGS_reverse primers for antisense transcripts), as well as 7SK (using T7 and 7SK forward primers). False-primed cDNA was amplified using all primer sets. qPCR reactions were performed at 95 °C for 5 min and 60 °C for 30 s, followed by 39 cycles of 95 °C for 5 s and 60 °C for 30 s. Results were analysed using the following formula: $\Delta\Delta Ct = 2^\wedge-(\Delta Ct_{Mutant} − \Delta Ct_{WT})$, where $\Delta Ct = Ct_{(transcript of interest)} − Ct_{(control)}$, and Ct is the cycle threshold. Values were normalized to those of false-prime reactions.

Population-level pre-rRNA pulse-chase

Click-iT Nascent RNA capture (Invitrogen, catalogue number C10365) was used. Cells were seeded in six-well dishes at 500,000 cells per well and allowed to grow to 40–50% confluence. Twenty-four hours later, cells were incubated with 0.15 mM ethyl uridine (EU) for 1 h, then with EU-free media for 2.5 h. Total RNA was extracted using Qiagen RNeasy kit (Qiagen, catalogue number 74104), and 1 μg of extracted RNA was incubated with 25 μl Click-iT EU buffer, 4 μl CuSO$_4$, 1.25 μl biotin azide, 1.25 μL Click-iT reaction buffer additive 1 for 3 min before addition of 1.25 μl Click-iT reaction buffer additive 2 and incubation for 30 min. The reaction mix was then incubated with 1 μl of UltraPure Glycogen (Roche, catalogue number 10901393001), 50 μl of 7.5 M ammonium acetate, and 700 μl of chilled 100% ethanol at −80 °C overnight. RNA was then pelleted using centrifugation at 13,000g for 20 min at 4 °C and two rounds of washes with 700 μl of 75% ethanol. We then treated 1 μg of the RNA with 31 μl Click-iT RNA binding buffer and 2 μl RNaseOUT before incubation for 5 min at 68–70 °C. The heated RNA-binding reaction mix was incubated with 12 μl of washed bead suspension at room temperature for 30 min. The beads were washed five times with Click-iT reaction wash buffer 1 and five times with Click-iT reaction wash buffer 2. The beads were then resuspended with 12 μl of Click-iT reaction wash buffer 2 and incubated at 68–70 °C for 5 min before proceeding with reverse transcription and qPCR. Processing was measured by qPCR assessment of the levels of unprocessed pre-rRNA containing the 5′ external transcribed spacer (ETS) compared with the total levels of mature rRNA.

Single-cell rRNA biogenesis assay

On the day before the assay, cells from different experimental conditions were harvested and seeded onto poly-L-lysine (PLL)-coated coverslips in 24-well plates. On the day of the assay, live cells were pulse-labelled with 1 mM 5-fluorouracil (5-FU; Sigma, catalogue number F5130) for 15 min, gently washed with unlabelled media, and chased for 30 min. Cells were then fixed and immunostained as described in the Methods section 'Endogenous protein immunofluorescence'. Double immunofluorescence labelling of nucleolar fibrillar centres or 5-FU-labelled RNA was performed using an anti-ATXN2 or anti-BrdU antibody, respectively. Random single cells were imaged captured at 100× using a Nikon C2+ confocal microscope coupled to NIS-Elements AR software (Nikon). Images were equally and evenly contrasted and ribosome biogenesis was measured as the ratio of ATXN2-marked nucleolar fibrillar centres with surrounding rRNA rings over the total number of nuclear ATXN2 foci.

Nuclear run-on

Click-iT Nascent RNA capture (Invitrogen, catalogue number C10365) was used for nuclear run-on (NRO). The setup was similar to that in the 'Population-level pre-RNA pulse-chase' section above, except that the total RNA was extracted after a 1 h incubation with 0.15 mM EU. Similar to pulse-chase labelling, the extracted RNA was biotinylated, precipitated, washed using Dynabeads, and reverse transcribed; qPCR was performed to measure the synthesis of nascent sincRNAs.

DNA–RNA hybrid immunoprecipitation

For DNA–RNA hybrid immunoprecipitation (DRIP) experiments, cells were first seeded in 60 mm plates at 2.5 × 10^6 cells per ml and allowed to grow to 70% confluence. Cells were then washed twice with ice-cold PBS, scraped, and centrifuged at 253g for 5 min. Cell pellets were resuspended in 1.6 ml TE buffer and incubated with 41.5 μl of 20% SDS and 5 μl of proteinase K overnight at 37 °C. Then, 1.6 ml of phenol-chloroform was added to cells before centrifugation at 466g for 5 min at room temperature. The aqueous layer was transferred and the addition of phenol-chloroform was repeated. The DNA was then precipitated by adding a 1/10 volume of 3 M NaOAc, pH 5.2, and 2.4 volumes of 100% ethanol to the aqueous layer. The DNA fibre was washed five times with 70% ethanol, resuspended in TE buffer and incubated with 3.5 μl spermidine (Bioshop, catalogue number SPR070), 35 μl buffer 2.1 (NEB), 5 μl HindIII (NEB, R01045), 10 μl EcoRI (Thermo Fisher, ERO271), 10 μL BsrGI (NEB, R05755), 5 μl XbaI (NEB, R01455) and 2 μl SspI

(NEB, R0132). We then added 40 μl of 3 M NaOAc, pH 5.2, and one volume of phenol-chloroform to the digested DNA, which was then centrifuged at maximum speed for 5 min. The aqueous layer was transferred, and addition of phenol-chloroform was repeated.

To precipitate the DNA, 2.4 volumes of cold 100% ethanol were added to the aqueous layer, incubated at −20 °C for 15 min, and centrifuged at maximum speed for 30 min at 4 °C. The DNA pellet was washed with 70% ethanol and spun at maximum speed for 5 min at 4 °C. The dry pellet was resuspended in 50 μl TE buffer, and 4.4 μg of the DNA was incubated with 350 μl TE buffer, 50 μl 10× binding buffer (100 mM NaPO$_4$ pH 7.0, 1.4 M NaCl, 0.5% (v/v) Triton X-100) and 10 μg of either mouse IgG or S9.6 antibody at 4 °C overnight. Immunoprecipitation samples were incubated with previously washed Dynabeads for 2 h at 4 °C. Samples were then washed three times with 1× binding buffer and eluted off the beads by incubation with DRIP elution buffer (50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 0.5% (w/v) SDS) and proteinase K for 45 min at 55 °C. The DNA was then purified using gel/PCR DNA fragment extraction (Geneaid, catalogue number DF300) and qPCR of purified DNA was performed. The specificity of the S9.6 antibody for RNA–DNA hybrids was confirmed by in vitro treatment with RNase H1 in all experiments. We also screen all antibodies for specificity by ensuring that signals do not exhibit any statistically significant changes following treatment with RNase III (NEB, catalogue number M0245S). Following ChIP–qPCR analysis, background IgG mock signal was subtracted from S9.6 immunoprecipitation signal to generate a DRIP signal, which was then plotted as a raw DRIP signal or as a relative DRIP signal when normalized to a given site or condition.

Fluorescence recovery after photobleaching
HEK293T cells were transfected with a GFP–UBF1 plasmid 24 h before cell passaging to 2 cm glass-bottomed live microscopy dishes. Next day, the roughly 75% confluent cells were treated with either flavopiridol to a final concentration of 2 μM or dimethylsulfoxide (DMSO) as a control. Cells were incubated for 3.5 h and subjected to fluorescence recovery after photobleaching (FRAP) microscopy. Confocal microscopy was executed using a ×100 oil-immersion lens (numerical aperture 1.47) on a Leica DMi8 motorized inverted microscope (Leica Microsystems) coupled to a VT-iSIM multipoint scanner (VisiTech International) and detected with a Flash 4.0 v3 sCMOS camera (Hamamatsu). FRAP was performed using the iLas FRAP system (Gattaca Systems). Design of the acquisition journals and system integration were by Quorum Technologies. Images were acquired with a 488-nm excitation wavelength laser at 15% intensity. Cells were initially imaged 20 times, and the point of interest was subsequently bleached with a 405-nm laser for 36 ms at a laser intensity of 26%. Cells were then imaged repeatedly for approximately 1 min post-bleach to capture recovery. Signal intensity was measured using MetaMorph analysis software. For analysis, the intensity of the region of interest was normalized to a nucleoplasmic background region at every time point. These background-adjusted values were then normalized to the intensity value from the first time point. The bleach time points (6–6.3 s) display saturated fluorescence as the bleached region of interest and were therefore not included in any quantification. Post-bleach control focus intensity values of greater than 1 are a result of bleach-induced decreases in nucleoplasmic background.

Creation and use of RED–LasRR system
Full-length human RNase H1 was fused to eGFP and the deactivated *Streptococcus pyogenes* Cas9 (with D10A and H840A mutations). The 5,844-nucleotide RNaseH1–SV40NLS–eGFP–SV40NLS–dCas9 gene was synthesized and cloned into the pcDNA4/TO plasmid (Invitrogen) using *Not*I and *Xba*I restriction sites (here, NLS is a nuclear localization sequence, and SV40 is simian virus 40). To ensure protein flexibility, a (GGGS)$_4$ linker was inserted between RNase H1 and the first SV40NLS, and another between eGFP and the second SV40NLS. GGS

linkers were also inserted between the first SV40NLS and eGFP, and between the second SV40NLS and dCas9. dRNaseH1-eGFP-dCas9 was generated by introducing the point mutation D210N to RNase H1 using the Q5 site-directed mutagenesis kit (NEB, catalogue number E055450) according to the manufacturer's protocol, with a modification of a 15-min instead of 5-min incubation with the KLD enzyme mix at room temperature. The oligonucleotide sequences for PCR amplification were 5′-gttctgtatacaaacagtatgttt -3′ and 5′-cagtttattgatgtttgagtctt -3′. The resulting RNaseH1–SV40NLS–eGFP–SV40NLS–dCas9 (RED) or its RNase H1-dead version (dRED) was then integrated into the T-REx (ThermoFisher Scientific) tetracycline-controlled expression system. Inducible expression of the fusion proteins is thus based on the binding of tetracycline to the Tet repressor, thereby derepressing the promoter controlling the expression of the RED and dRED fusion protein. To achieve locus-specific RED and dRED–LasRR enrichment, cells were allowed to reach 70% confluence over a period of roughly 24 h. For the inducible condition, cells were incubated with medium containing tetracycline (1 μg ml^{-1}); for the uninduced condition, cells were incubated with tetracycline-free medium. All cells were transfected with 3 μg of RNH1–eGFP–dCas9 and dRED–LasRR plasmid per 60-mm plate by using Lipofectamine3000 (ThermoFisher) as per the manufacturer's instructions. Induced cells were then cotransfected using RNAiMAX (ThermoFisher) as per the manufacturer's instructions with either 4.5 μl of 10 pmol μl^{-1} nontargeting sgRNA, 1.5 μl of 10 pmol μl^{-1} of each of three sgRNAs for IGS18, 1.5 μl of 10 pmol μl^{-1} for each of three sgRNAs for IGS28, 1.5 μl of 10 pmol μl^{-1} for each of three sgRNAs for IGS38, or 1.5 μl of 10 pmol μl^{-1} for each of three sgRNAs for the *β-actin* 5′ pause element. The cells were incubated for 36 h before further experiments were performed.

CRISPR-mediated genome editing
For CRISPR-mediated gene knockout of SETX, CRISPR/Cas9 plasmids (pCMV–Cas9–GFP) were purchased from Sigma-Aldrich to express with the scrambled guide RNA and guide RNA for SETX (first intron). The transfections of the plasmids into the Flp-In 293 T-REx cell lines were performed with FuGENE transfection reagent (Roche, catalogue number E269A). We transfected 2 μg of the plasmid into HEK293T cells; one day after transfection, we sorted cells by BD FACSAriaTM flow cytometry (Donnelly Centre, Univ. Toronto), and plated single GFP-positive cells into 96-well plates. To confirm SETX knockout, the expression levels of SETX in each clone were detected by qPCR.

Northern blotting
RNA was prepared as described in the 'RNA extraction' and 'Reverse transcription' sections above. We then electrophorized 3.5 μg of RNA, and digoxygenin (DIG)-labelled the DNA probe for northern blotting using the DIG-high prime DNA labelling and detection starter kit I as per the manufacturer's protocol (Roche, catalogue number 11745832910). Northern blots were performed using the DIG northern starter kit as per the manufacturer's protocol (Roche, catalogue number 12039672910), with the following modification: electrophoresis was conducted at 15 V for 24 h at room temperature; RNA was UV-crosslinked (2,400 kJ for 1 min) to a positively charged nylon membrane; gels were blotted by capillary transfer with 20× SSC buffer (3 M NaCl, 0.3 M sodium citrate) overnight; a hybridization temperature of 50 °C was used; blots were hybridized overnight; and 200 μl of NBT/BCIP solution in 10 ml of detection buffer (0.1 M Tris-HCl pH 9.5, 0.1 M NaCl), was used for blot development.

DNA–RNA hybrid immunofluorescence
We seeded 60,000–80,000 cells per PLL-coated coverslip and allowed them to adhere for 24–36 h. Cells were fixed using 1% (v/v) formaldehyde for 15 min at room temperature, washed three times with 1× PBS for 5 min each, permeabilized with 500 μl of 0.3% (v/v) Triton-X100 for 5 min at room temperature, and washed again three times with PBS.

Article

Coverslips were blocked using 500 μl of 5% bovine serum albumin (BSA) for 1 h at room temperature, transferred to humidified chambers and incubated with 60 μl of primary antibody (1:500 of S9.6 antibody, 1% (w/v) BSA, 1× PBS) for 1 h at room temperature. After washing with PBS, cells were incubated with 60 μl of secondary antibody (1% BSA, 1:250 of goat anti-mouse 488 or 1:250 of goat anti-mouse 568) for 1 h at room temperature. The cells were washed again with PBS and incubated with 100 μl of DAPI (0.5 μl of DAPI per ml of PBS) for 2–4 min. The coverslips were then mounted onto microscope slides using DAKO mounting reagent, sealed with nail polish, and allowed to dry for 30 min. Images were acquired using a C2+ confocal microscope with a Plan-Apochromat TIRF ×100 oil objective (numerical aperture 1.45) and NIS-Elements AR software (Nikon). The specificity of the S9.6 antibody for RNA–DNA hybrids was confirmed by in vitro treatment with RNase H1 (NEB, catalogue number M0297S) under the same experimental conditions. Signals were also confirmed to differ from those yielded by immunofluorescence using J2, an antibody against double-stranded (ds)RNA.

Amyloid-body staining with Congo red
We seeded 40,000 cells on PLL-coated coverslips and allowed them to adhere for 24–36 h. Cells were fixed using 1% (v/v) formaldehyde and incubated for 15 min at room temperature, washed three times with 1× PBS for 5 min each, and permeabilized with 500 μl of 0.3% (v/v) Triton-X100 for 5 min at room temperature. The coverslips were then immersed in 250 μl of 0.05% (v/v) Congo red (Sigma, catalogue number C6277) solution for 15 min, followed by four cycles of 2 min rinsing with 500 μl of double-distilled H_2O. The coverslips were then transferred to humidified chambers and nuclear counterstained with 100 μl DAPI, incubated for 4 min, and mounted on microscope slides using DAKO mounting reagent. Images were acquired using a C2+ confocal microscope with a Plan-Apochromat TIRF ×100 oil objective (numerical aperture 1.45) and NIS-Elements AR software (Nikon).

Endogenous protein immunofluorescence
We seeded 40,000 cells onto PLL-coated coverslips. Cells were fixed using 1% formaldehyde for 1 min at room temperature, washed with 1× PBS three times (5 min each wash), permeabilized with 500 μl of 0.3% Triton-X100 for 5 min at room temperature, and washed again three times with 1× PBS. Coverslips were blocked using 500 μl of 5% BSA for 1 h at room temperature, transferred to humidified chambers and incubated with 60 μl of primary antibody (1% BSA and anti-UBF or anti-NPM antibodies) for 1 h at room temperature. After washing with PBS, cells were incubated with 60 μl of secondary antibody (1% BSA, 1:250 of goat anti-mouse 488, 1:250 of goat anti-rabbit 568) for 1 h at room temperature. Coverslips were washed again with PBS, incubated with 100 μl DAPI for 2 min, mounted onto microscope slides using DAKO mounting reagent, and allowed to dry for 30 min. Images were captured at ×100 or ×60 using a Nikon C2+ confocal microscope coupled to NIS-Elements AR software (Nikon). For methanol/acetone-fixation-based immunofluorescence, the protocol was similar except that cells were fixed using ice-cold methanol for 15 min at room temperature, washed once with cold acetone, and washed with 1× PBS (3 × 5 min) before blocking with 5% BSA. Super-resolution microscopy was captured with a Leica DMI6000 SP8 LIGHTNING microscope using the HC PL APO CS2 x93 objective (numerical aperture 1.3, pinhole 110.5 μm). Images were deconvolved using Leica LIGHTNING deconvolution software and processed with Leica LAS software.

Stress-induced droplets and amyloid bodies
The ACM-containing VHL protein was transfected as pFLAG–VHL–GFP using Lipofectamine3000 according to the manufacturer's protocol and treated/visualized 24 h post-transfection. For siRNAs (100 pmol), cells were transfected using RNAiMAX (ThermoFisher Scientific) at 90% confluency, split next day into fresh plates at 70% confluency to allow for subsequent GFP transfection with Lipofectamine3000, and treated/

harvested 48 h post-transfection. The siRNAs (ThermoFischer Scientific) used were siControl (catalogue number 4390843), si-sincRNA16 (catalogue number 4399666) and si-sincRNA22 (catalogue number 4390828). For live-cell imaging of the ACM-containing and GFP-tagged VHL protein, cells were seeded in 145-μm-thick, 35-mm glass-bottom plates. Live-cell images were captured by confocal microscopy (Leica TCS SP5; Leica Microsystems, Mannheim, Germany), fitted with a variable temperature and 5% CO_2 environmental chamber (Okolab), using a ×63 oil-immersion Plan-Apochromat objective (numerical aperture 1.4). Images were uniformly adjusted to increase brightness/contrast in Photoshop (Adobe).

In vitro droplet formation
Peptides were custom synthesized by GenScript (New Jersey, USA) at more than 95% purity. Peptide stock solutions were kept at 50 mM in nuclease-free-water. 5'-carboxyfluorescein (FAM)-labelled RNAs were synthesized by Integrated DNA Technologies (IDT, Coralville, IA) and resuspended in 50 mM NaCl to 100 μM. We mixed 1 μM low- or high-complexity ncRNA with the indicated peptide concentrations in 150 mM NaCl. Droplets were placed on a 1.5 coverslip and imaged after a 10-min incubation on a Zeiss AxioObserver D1 microscope using a ×63 Plan-Apochromat objective (numerical aperture 1.4).

Locked nucleic acid ASO knockdown of sincRNAs
Custom-designed locked nucleic acid (LNA) ASO GapmeRs were ordered from Qiagen. Sequences of 975–1,000 bp corresponding to IGS regions were entered into Qiagen's custom antisense LNA GapmeR design page. The top-ranked ASOs based on Qiagen's optimal design score were selected for each of IGS18, IGS20, IGS22 and IGS24, with standard desalting purification, phosphorothioate backbone modifications, and no-label/ready-to-label design specifications. ASO transfection was performed using RNAiMAX (ThermoFisher Scientific) according to the manufacturer's protocols. ASOs were as follows: antisense LNA GapmeR control negative control B (catalogue number 339515, LG0000001-DDA; gctcccttcaatccaa), IGS18 (LG00210930-DDA; agtgtgtctgtgtgaac), IGS20 (LG00210936-DDA; acgcaagaaaggaaga), IGS22 (LG00210956-DDA; acgtgaccgagagaaa) and IGS24 (LG00210966-DDA; gtgacgtgtagagatt).

Subcellular fractionation by sucrose gradient
Cells were trypsinized, centrifuged at 1,000 r.p.m. for 4 min at 4 °C, washed with PBS and recentrifuged. The pellet was resuspended in osmotic buffer (10 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 10 mM KCl, 0.5 mM DTT). The cells were then lysed using ten strokes of a tight pestle in a dounce homogenizer. Dounced cells were centrifuged at 1,000 r.p.m. for 5 min at 4 °C. The supernatant was retained as the cytosolic fraction. The pellet was resuspended in a 0.25 M sucrose plus 10 mM $MgCl_2$ solution, and deposited over a 0.35 M sucrose plus 0.5 mM $MgCl_2$ layer. The sample was centrifuged at 1,000 r.p.m. for 5 min at 4 °C. The sample was then resuspended in a 0.25 M sucrose plus 10 mM $MgCl_2$ solution and sonicated at 25% power six times for 10 s with intermittent periods of 10 s rest on ice. The sample was deposited over a 0.88 M sucrose plus 0.5 mM $MgCl_2$ layer and centrifuged at 3,500 r.p.m. for 10 min at 4 °C. The supernatant was retained as the nucleoplasmic fraction. The pellet was resuspended in a 0.35 M sucrose plus 0.5 mM $MgCl_2$ solution and centrifuged at 3,500 r.p.m. for 5 min at 4 °C. The pellet was the nucleolar fraction. GAPDH transcripts, which are most abundant in the cytosolic fraction and are depleted from the nucleolar fraction, served as the control.

Aligning sequencing reads to human rDNA IGS
First, we used the Bowtie package to build a version of the human genome assembly hg19 with rDNA sequence; the newly built assembly is 'hg19_plus_rDNA'. The human rDNA sequence FASTA file was obtained as is from the National Center for Biotechnology Information

(NCBI; https://www.ncbi.nlm.nih.gov) under GenBank accession code U13369.1, which refers to the 'Human ribosomal DNA complete repeating unit'. This FASTA file, along with those for chromosomes 1–22, X, Y and M from hg19, obtained from the University of California at Santa Cruz (UCSC) genome browser (https://genome.ucsc.edu), was used to build the new assembly. Next, for testing, we aligned Pol II reads from HeLa cells to this new genome assembly using the Bowtie package aligner. The reads from two replicates were obtained from ENCODE and concatenated. Duplicate reads were removed with the package BBmap and its clumpify tool. Then, the alignment was performed with the parameter '-m 1', which instructs bowtie to refrain from reporting any alignments for reads having more than one reportable alignment. This ensures that only uniquely aligning reads are reported. The alignments were processed further with Samtools to retain only those reads aligning to the rDNA sequence, and to compute the depth/number of reads at each position in the rDNA coordinates. These depths were plotted with an R script.

Calculation of GC skew

Using the roughly 43-kbp rDNA sequence obtained from the rDNA sequence FASTA file, we assessed GC skew, GC observed/expected ratio and GC percentage using (1-bp-at-a-time) sliding windows of size 50, 500 or 1,000 bp. Definitions are as follows: GC skew = (number of Gs − number of Cs)/(number of Gs + number of Cs); CG observed/expected ratio = sliding-window length × number of CpGs/(number of Cs × number of Gs); GC percentage = 100 × (number of Gs + number of Cs)/sliding-window length. To obtain an overall value/quantification and statistic with which to compare coding and IGS regions, we obtained the mean GC skews for the coding and IGS regions with window size 1,000. In the coding region the mean GC skew is 0.02346459; and in IGS regions, it is −0.1541796. Applying a Welch's t-test to the GC skews from these two regions gives a P-value of less than 2.2×10^{-16}. The script for all above analyses is called getGCskewEtc_rRDNA.R and is available upon request.

Sequencing

For Ewing- and osteosarcoma-related analyses, sample preparation and sequencing were carried out as described[26]. RNA-seq and DRIP–seq data sets have been deposited at Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) under accession code GSE68847. Identification of rDNA IGS peaks from RNA-seq and DRIP–seq was conducted as per the section 'Aligning sequencing reads to human rDNA IGS' above, including a normalization of called peaks to the total number of reads per sample. For assessment of signals at non-rDNA loci, aligned .bam files were depth normalized and binned using bamCoverage from deepTools[36]. The resulting bigWig files were loaded into Integrated Genome Viewer (IGV)[37] and the depicted regions were saved. For RNA-seq with/without heat shock, sequencing was performed on a cDNA library of total RNA (non-rRNA-depleted) using stranded paired-end reads. After discarding reads mapped to the rRNA gene (including the 5'-ETS, inverted repeat sequence (IRS)1/2 and 3'-ETS), we mapped the remaining reads to GRCh38. BAM files were separated into forward and reverse strand files (bash script). The remaining reads aligned to supercontig GL000220.1; this is within the latest human genome assembly, which contains a 43-kb rDNA cassette. Signals were normalized to an internal non-stress responsive control site at IGS35. To calculate changes in sincRNA and asincRNA levels following heat shock, we binned the IGSs into 5,000-bp bins, and calculated the change in absolute read counts for each bin. The average of these changes was calculated to obtain a global percentage change across the entire IGS region. The sequencing data are available at GEO under accession code GSE115731. ChIP–seq enrichments were generated by the ENCODE Project Broad Institute for H3K27ac, H3K9ac, H3K4me3 and H3K36me3, and by ENCODE Project SYDH for RNA pol II ChIP–seq. Briefly, bedGraph files previously generated[38] by mapping ChIP–seq

and input data from ENCODE Project Consortium 2012 to the human rDNA sequence from BAC clone GL000220.1 were used to generate IGV genome tracks. We note that qPCR and sequencing analyses of repetitive DNA loci reveal an average profile for the studied repeats and should not be interpreted as an absolute enrichment for any given unit within the repeats.

Transmission electron microscopy

Cell pellets were fixed in phosphate-buffered 4% formaldehyde plus 1% glutaraldehyde fixative for at least 2 h. Samples were subsequently rinsed in 0.1 M phosphate buffer for 5 min and then fixed in 1% Zetterqvist's buffered osmium tetroxide for 1 h. After a short rinse in Zetterqvist's buffer for 1 min, the samples were dehydrated in increasing concentrations of alcohol (70%, 95%, 100%) for 10 min followed by propylene oxide. Finally, pellets were embedded in epoxy resin. Ultrathin sections were contrasted with uranyl acetate and Reynold's lead citrate and observed with a JEOL 1230 TEM equipped with an Advanced Microscopy Techniques (AMT) camera system.

Images of human tumour sections

Images of tumour sections stained with haematoxylin and eosin were obtained through the Sinai Health System (Toronto) without any identifiable personal health information and without personal information, following Institutional Research Ethics Board approval (Sinai Heath Systems, 17-0103-E).

Statistical analysis

GraphPad Prism-based calculations of P-values were carried out via t-test, one-way ANOVA (with Dunnett's or Tukey's multiple comparison test), or Mann–Whitney U-test. Unless otherwise indicated, replicate information is as follows. All data from pulldowns, reverse transcription and viability markers were generated using the indicated number of biological replicates. For blots, images are representative of data obtained from two independent biological replicates. For microscopy, images are representative of phenotypes observed in at least two independent biological replicates, and quantifications are based on at least 100 cells from two technical replicate cultures.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Data are in the paper, Supplementary Fig. 1 (uncropped blots) and the Source Data files related to Figs. 1–4 and Extended Data Figs. 1–3, 5–8. RNA-seq and DRIP–seq data sets have been deposited at GEO under accession codes GSE115731 and GSE68847. All data and materials are available upon reasonable request. In light of the pandemic, shipping of reagents and materials may be slightly delayed. Source data are provided with this paper.

Code availability

All scripts used to analyse data are available upon request.

35. Ho, E. C., Donaldson, M. E. & Saville, B. J. Detection of antisense RNA transcripts by strand-specific RT-PCR. *Methods Mol. Biol.* **630**, 125–138 (2010).
36. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44** (W1), W160–W165 (2016).
37. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
38. Agrawal, S. & Ganley, A. R. D. The conservation landscape of the human ribosomal RNA gene repeats. *PLoS ONE* **13**, e0207531 (2018).

Article

Competing interests D.D.D.C. received research funds, not related to this work, from Pfizer and Nektar Therapeutics. All other authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2497-0.
Correspondence and requests for materials should be addressed to K.M.
Reprints and permissions information is available at http://www.nature.com/reprints.
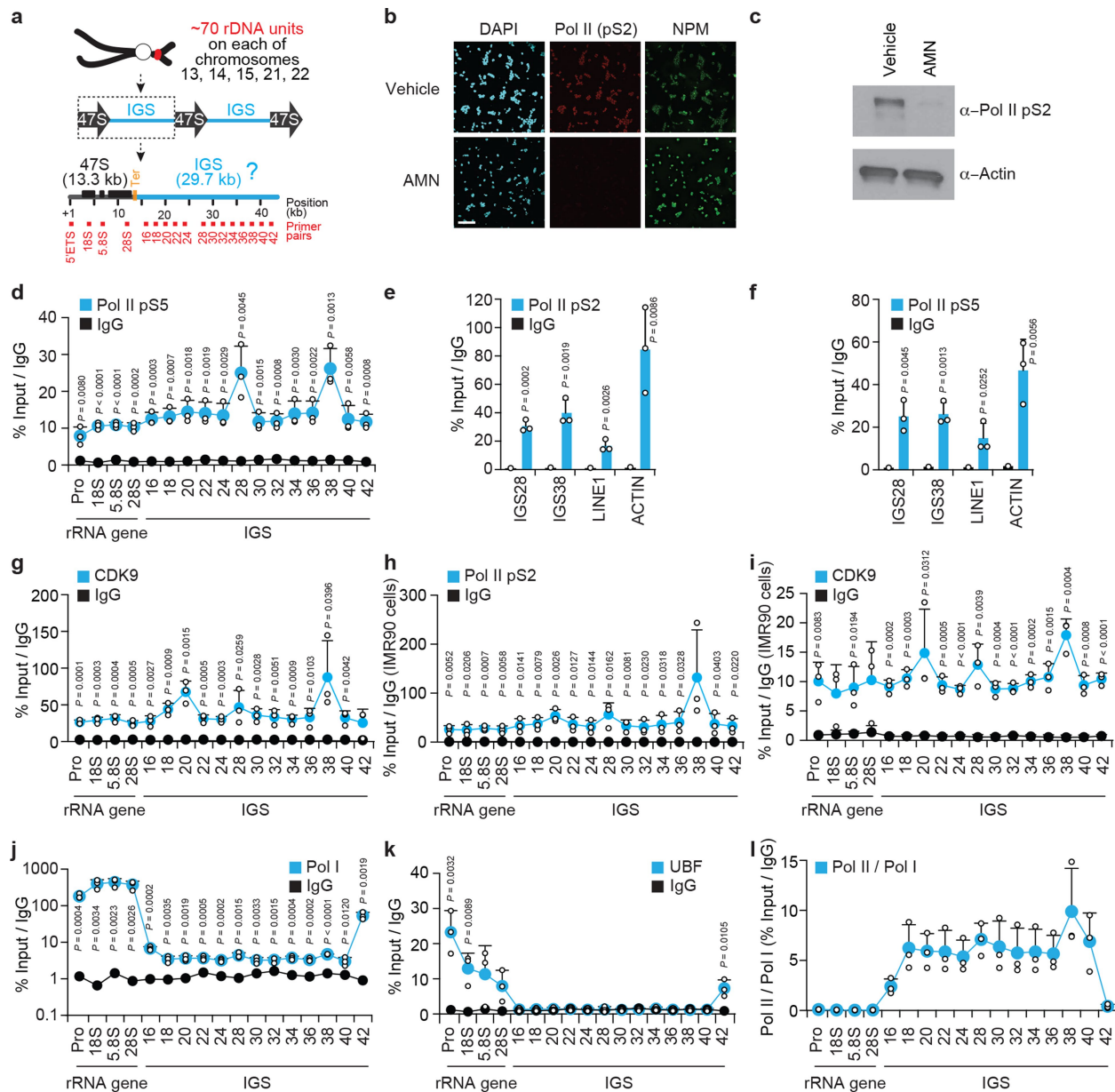
Extended Data Fig. 1 | Additional characterization of Pol I and Pol II occupancy at rDNA IGSs. a, Organization of human rDNA repeats. At each rDNA unit, Pol I transcribes an rRNA gene encoding a 47S pre-rRNA that is processed to remove transcribed spacers, such as the 5′-ETS, and generate 18S, 5.8S and 28S rRNA molecules. The IGS constitutes the bulk of each rDNA unit. Ter, rRNA gene terminator. **b**, **c**, Specificity controls indicating that targeting Pol II for degradation with a 12-hour α-amanitin (AMN) treatment lowers anti (α)-Pol II pS2 signals in both immunofluorescence (**b**) and immunoblotting (**c**). Actin was used as a control for immunoblotting. For gel source data, see Supplementary Fig. 1. **d**, ChIP showing Pol II pS5 enrichment across rDNA.

e, **f**, The enrichment of active Pol II pS2 and pS5 at rDNA IGS sites is higher than at LINE1 but lower than at β-actin sites. **g–k**, ChIP experiments showing the enrichment of the indicated proteins across rDNA. **l**, Comparison of the enrichment of RNA Pol II and Pol I across rDNA reveals the relative overrepresentation of Pol II across IGSs only. **b–l**, HEK293T (**b–g**, **j–l**) or IMR90 (**h**, **i**) cells were used; data shown are means ± s.d.; two-tailed *t*-test, *n* = 3 biologically independent experiments (**d–l**); images in **b**, **c** are representative of two independent experiments. Data in **d–f**, **j–l** and Fig. 1b were from large experimental sets sharing IgG controls. Data in **h**, **i** were from large experimental sets sharing IgG controls.
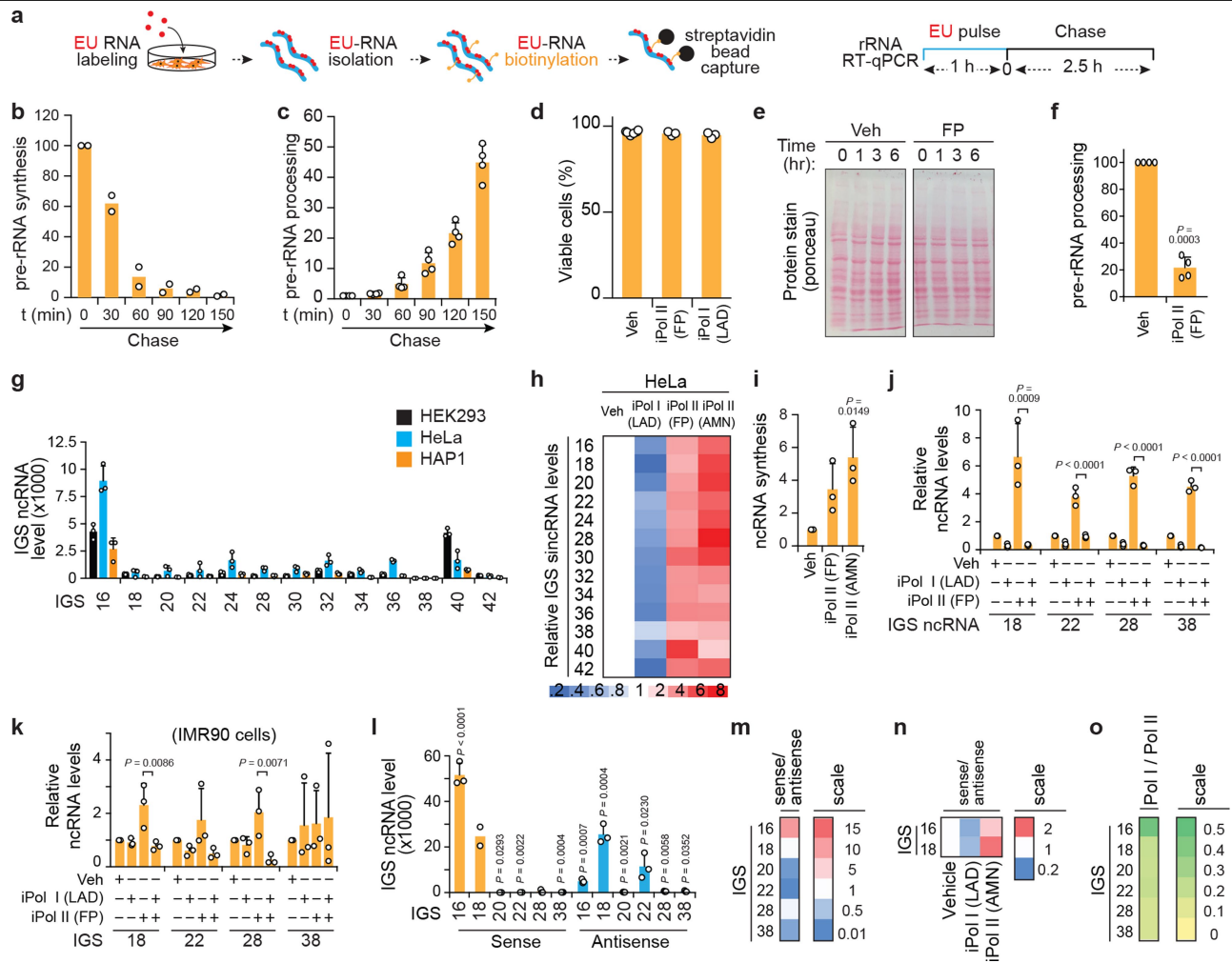
Extended Data Fig. 2 | Impact of Pol I and Pol II on IGS ncRNA levels in various cell types. a, Cell-population-based RNA pulse-chase assay used to assess pre-rRNA synthesis and processing. **b**, **c**, Confirmation of the detection of pre-rRNA synthesis (**b**) and processing (**c**) by EU–RNA pulse-chase assays, as shown in Fig. 1d, e. **d**, Trypan blue exclusion assay confirms that the 3-hour Pol II inhibition (iPol) regimens used in our functional assays do not compromise cell viability. **e**, Ponceau staining shows stable protein levels following Pol II inhibition. Veh, vehicle. **f**, Treatment with the fast-acting RNA Pol II inhibitor flavopiridol (FP) for 30 min is sufficient to abrogate pre-rRNA processing. **g**, Human IGS ncRNAs are also detected across the IGSs of diploid HeLa cells and haploid HAP1 cells. **h**, Pol I promotes and Pol II represses IGS ncRNAs in HeLa cells. **i**, Nuclear run-on assay showing de novo IGS ncRNA synthesis mediated by Pol II inhibition. **j**, **k**, Reverse-transcription experiments showing the effect of combining Pol I and Pol II inhibition on IGS ncRNAs in HEK293T cells (**j**) and

IMR90 cells (**k**). **l**, **m**, Strand-specific RT–qPCR (ss-RT) showing the levels of sense and antisense intergenic ncRNAs (**l**) and their derived sense/antisense ratios (**m**) at various IGS sites. **n**, ss-RT shows that Pol I inhibition decreases and Pol II inhibition increases the sense/antisense ratio of the most abundant IGS ncRNAs. **o**, Despite the preferential enrichment of Pol II over Pol I across IGSs, Pol II is the least overrepresented relative to Pol I at IGS16 compared with all other IGSs tested. **a**–**o**, HEK293T cells were used unless otherwise indicated; data are shown as means ± s.d.; two-tailed *t*-test (**b**–**d**, **f**) or one-way ANOVA with Dunnett's multiple comparison test (**g**, **i**, **k**); $n = 2$ biologically independent experiments (**b**), $n = 4$ biologically independent experiments (**c**, **f**), and $n = 3$ biologically independent experiments (**d**, **g**–**o**), except in the case of sense IGS18, for which $n = 2$ biologically independent samples (**l**, **m**); image in **e** is representative of two independent experiments.
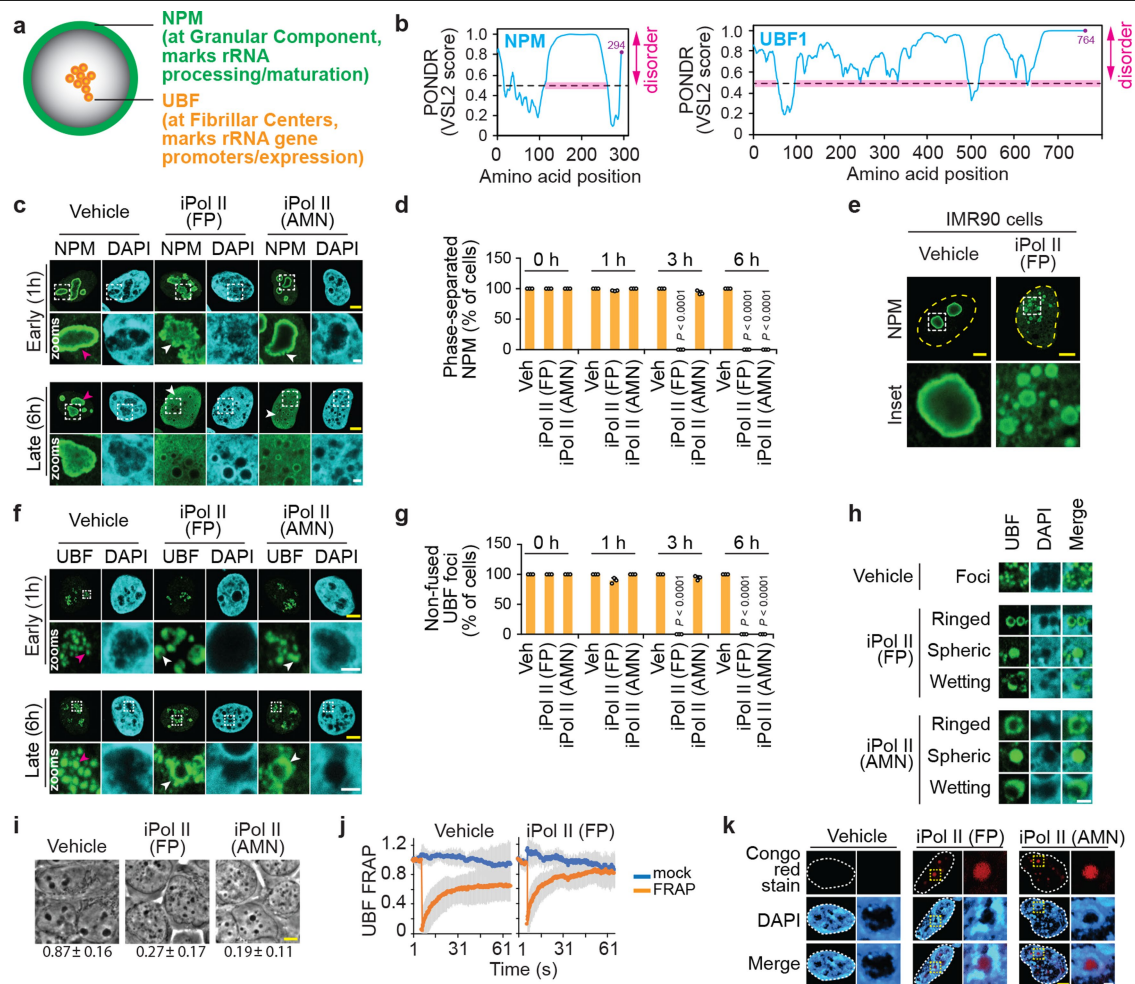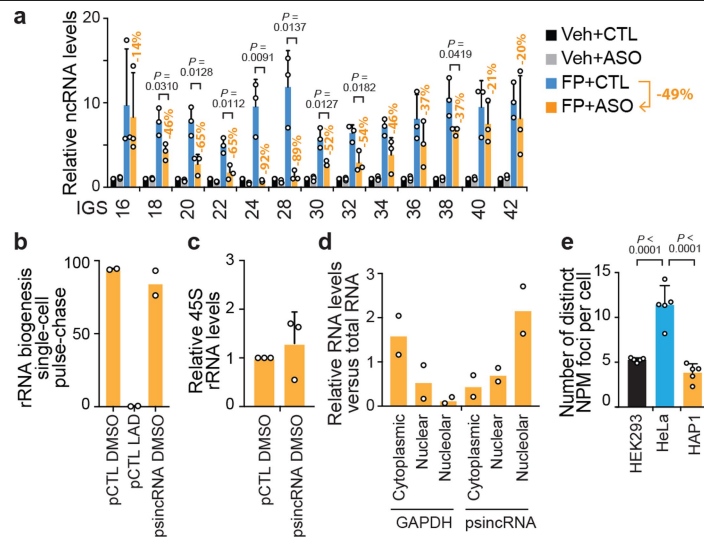
Extended Data Fig. 3 | Characteristics of nucleoli and nucleolar proteins in the presence or absence of Pol II inhibition. a, b, Schematic of a nucleolus, illustrating the localization of LLPS nucleolar subcompartments marked by the resident proteins NPM and UBF (**a**), which are highly disordered, as revealed using the various short long 2 (VSL2) predictor of natural disordered regions (PONDR) algorithm (**b**). **c**, Effects of Pol II inhibition (iPol II) on NPM localization, as shown by immunofluorescence microscopy. Examples of normal and defective phenotypes are respectively marked by magenta and white arrowheads. **d**, Quantification of the percentage of cells that have any NPM phase-separated body reveals that the fast-acting Pol II inhibitor FP completely disrupts nucleoli before the slower-acting Pol II inhibitor AMN can take effect. Not depicted is the percentage of cells with perturbed nucleolar architecture as evidenced by NPM1 ruffling, which increased from $0.6 \pm 4.6\%$ to $63.3 \pm 5.7\%$ following the 1-hour FP treatment. **e**, Pol II inhibition also disrupts NPM localization in IMR90 cells. **f**, Effects of Pol II inhibition on UBF localization, as shown by immunofluorescence microscopy. Examples of normal and defective phenotypes are respectively marked by magenta and white arrowheads.

g, Quantification of the percentage of cells that have any punctate UBF localization confirmed that the fast-acting FP completely disrupts nucleoli before the slower-acting AMN. **h**, Pol II inhibition triggers various aberrant UBF localization phenotypes, as shown in representative images. **i**, Global nucleolar disruption following Pol II inhibition, as revealed by phase-contrast microscopy. The fraction of cells with more than three black nucleolar bodies is indicated. **j**, Live-cell UBF fluorescence recovery after photobleaching (FRAP). Mock control cells were continuously imaged without a photobleaching step. FRAP FP/vehicle rate-constant ratio = 2.3. **k**, Formerly nucleolar space became Congo red positive after Pol II inhibition. **c**–**k**, HEK293T cells were used unless otherwise indicated; data are means ± s.d.; one-way ANOVA with Dunnett's multiple comparisons test, $n = 3$ biologically independent experiments (**d**, **g**) or $n = 5$ biologically independent experiments (**i**); for **j**, vehicle FRAP cells $n = 30$, vehicle control cells $n = 4$, FP FRAP cells $n = 15$, and FP control cells $n = 6$; images in **e**, **h**, **k** are representative of two independent experiments. Scale bars, 5 μm (yellow) or 1 μm (white).
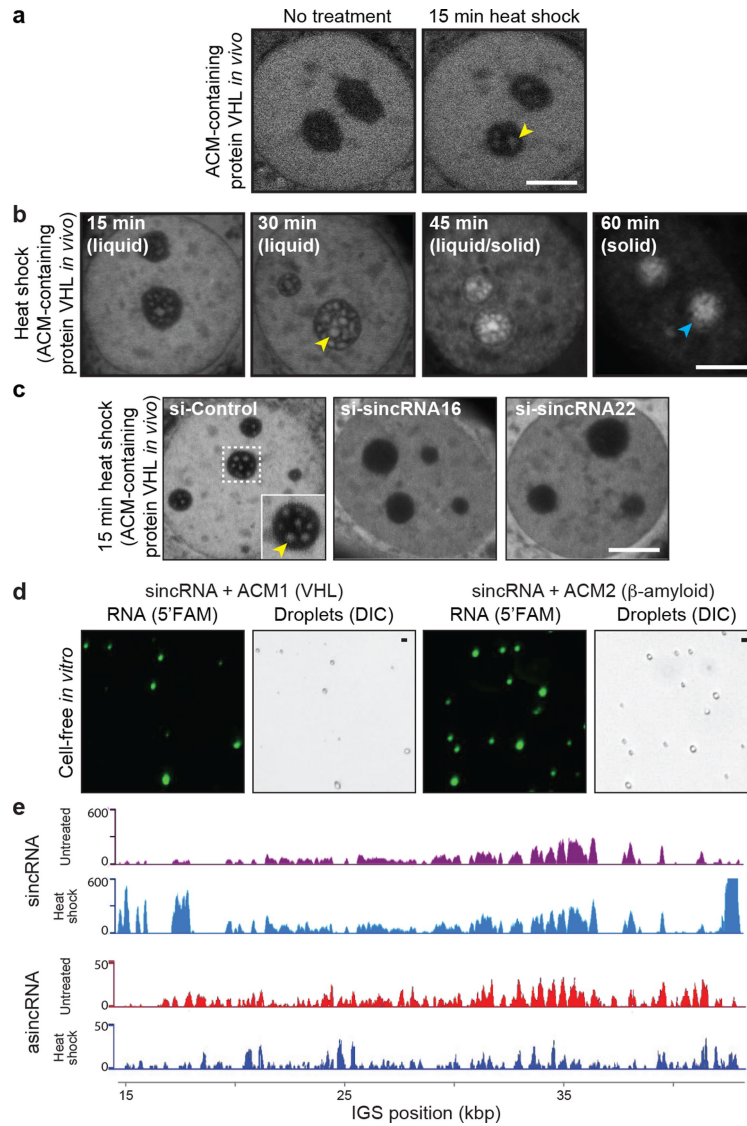
Extended Data Fig. 4 | Heat shock limits asincRNAs and triggers sincRNA-dependent nucleolar phase transitions. a, Heat shock (43 °C) rapidly induces the formation of intranucleolar liquid droplets harbouring the ACM-containing VHL protein. **b**, Gradual amyloid-body (A-body) formation. The stress-induced, mobile and spherical liquid-like foci (yellow arrowhead) gradually transition into irregularly shaped, solid-like amyloid bodies (cyan arrowhead) in cells subjected to heat shock (43 °C)[19]. **c**, The appearance of early-stage, ACM-marked, liquid-like foci[18,19] in cells subjected to a 15-min heat-shock treatment is abrogated upon siRNA-mediated knockdown of either sincRNA16 or sincRNA22. **d**, In a cell-free in vitro system, the low-complexity sincRNA (1 μM) forms liquid droplets when mixed with the ACM of human VHL or β-amyloid proteins (25 μM). Droplets were detected using fluorescently labelled RNA (5′FAM) and differential interference contrast (DIC). **e**, ss-RNA-seq reveals that sincRNA levels increase while asincRNA levels decrease across the IGS following a 30-min heat shock. Heat shock increases sincRNA levels by 607% and decreases asincRNA levels by 38%. **a**–**e**, Nucleolar-stress hyperresponsive MCF7 cells were used where applicable; images are representative of two independent experiments; scale bars, 5 μm.

a, In HEK293T cells treated with the Pol II inhibitor FP, introduction of ASOs targeting sincRNAs lowers IGS ncRNA levels relative to ASO control-treated cells (CTL). ASO-dependent percentage decreases in sincRNA levels are indicated for each IGS site; the average decrease in total sincRNA levels is 49%. Data are means ± s.d.; two-tailed t-test, $n = 3$ biologically independent experiments. **b**–**d**, In the absence of heat shock, artificial overexpression of sincRNA22 (psincRNA) in nucleolar-stress hyperresponsive MCF7 cells failed to repress rRNA biogenesis (**b**) or rRNA levels (**c**), despite the enrichment of sincRNA22 in the nucleolar fraction (**d**). Plasmid (pCTL), iPol I (LAD), vehicle (DMSO) and GAPDH cell fractionation controls were included. Data are means ± s.d.; $n = 2$ biologically independent experiments (**b**, **d**); two-tailed t-test, $n = 3$ biologically independent experiments (**c**); **e**, Quantification of the number of distinct NPM foci per cell in different cell types. Data are means ± s.d.; one-way ANOVA with Tukey's multiple comparisons test, $n = 5$ biologically independent experiments.

Extended Data Fig. 5 | Artificial and natural modulation of sincRNA levels.

Article
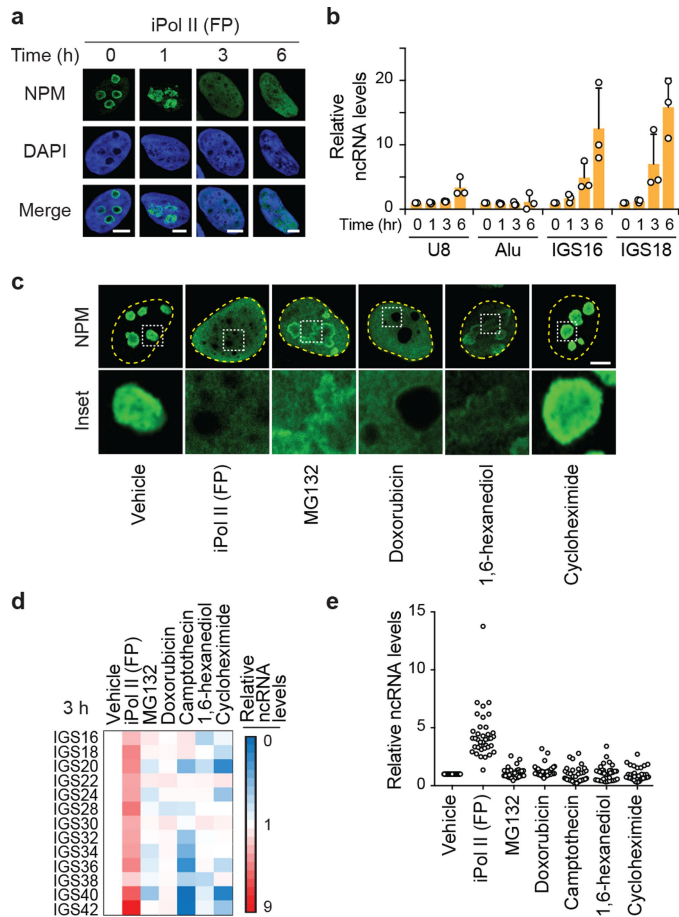
Extended Data Fig. 6 | Controls related to the disruption of nucleolar structure following Pol II inhibition. a, b, The disruption of NPM phase separation following Pol II inhibition (**a**) coincides with time points at which the levels of IGS ncRNAs greatly increase (**b**; means ± s.d., $n = 3$ biologically independent experiments). At these time points, no reductions in the levels of the small nucleolar (sno)RNA U8 or Alu RNA were observed. **c–e,** Treating cells with the Pol II inhibitor FP, with various drugs that disrupt nucleolar morphology through unclear mechanisms (MG132, doxorubicin), with the LLPS/nucleolus disruptor 1,6-hexanediol, or with the global translation inhibitor cycloheximide reveals that only Pol II inhibition simultaneously disrupted NPM phase separation (**c**) and induced IGS ncRNA levels (**d**, **e**). Shown are representative anti-NPM immunofluorescence images (**c**) and two different visual representations of ncRNA levels as detected by RT–qPCR (**d**, **e**); $n = 3$ biologically independent experiments. In the scatter plot (**e**), each circle represents the value of one IGS site from one of three biological replicates. Scale bars, 5 μm.
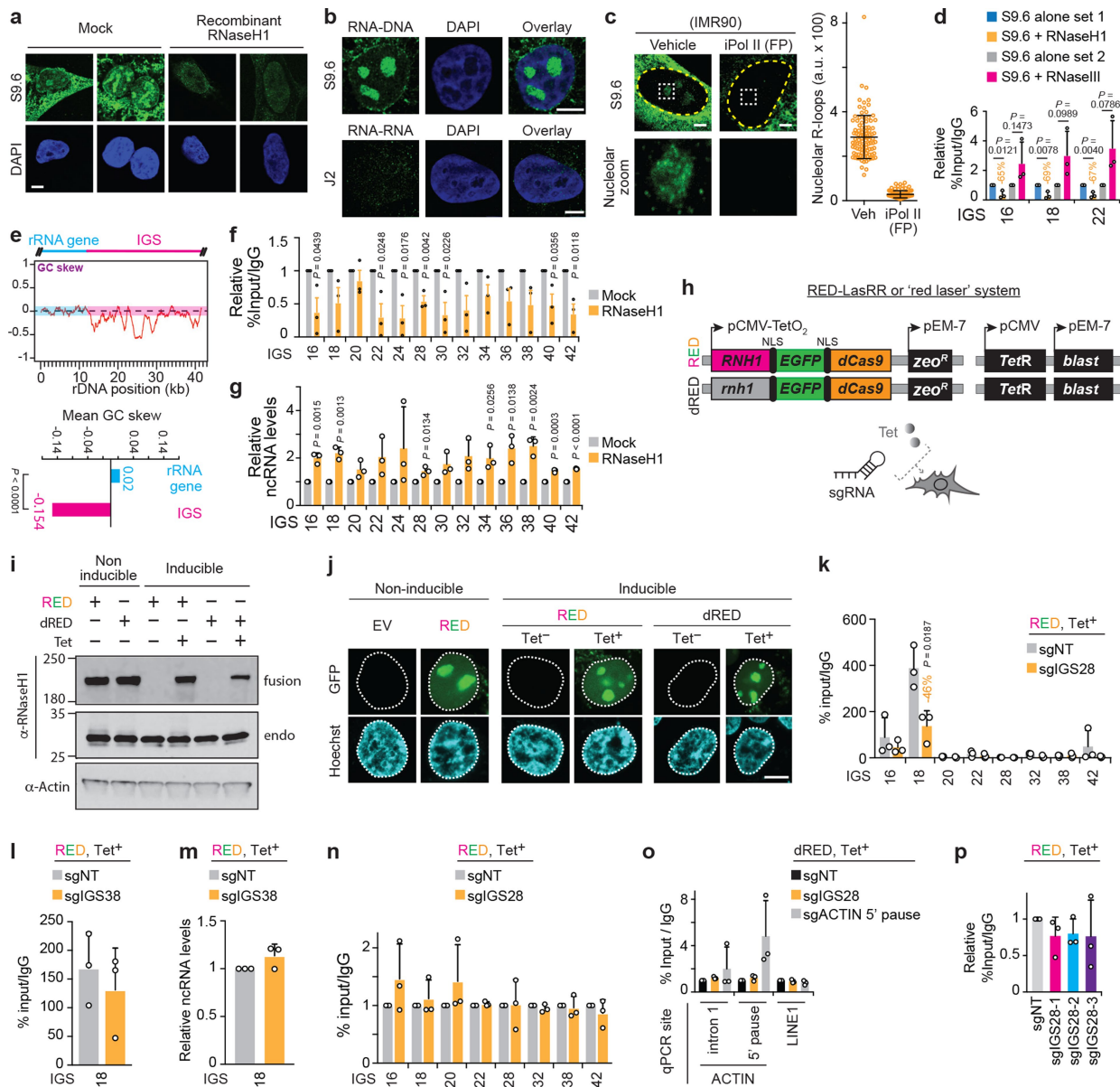
Extended Data Fig. 7 | Nucleolar R-loops and their modulation. a, In vitro treatment with recombinant RNase H1 greatly decreases the nuclear immunofluorescence signals obtained with S9.6, an antibody against DNA–RNA hybrids. Signals remaining following RNase H1 treatment may reflect resistant hybrid structures or other nucleic acid structures. **b**, Immunofluorescence using S9.6, but not the anti-dsRNA antibody J2, reveals a nucleolar signal under standard cell culture conditions. **c**, Immunofluorescence using S9.6 with IMR90 cells also shows nucleolar signals that are repressed upon Pol II inhibition ($n = 100$ cells). **d**, In our DRIP assays, in vitro treatment with RNase H1, but not RNase III, consistently lowers DRIP signals. **e**, Bioinformatic analysis of the rDNA GC skew distribution and mean shows that the IGSs, but not rRNA genes, display a strongly negative GC skew; Welch's two-tailed t-test, $n = 14$ (rRNA gene) and $n = 30$ (IGS). **f, g**, RNase H1 overexpression partly lowers R-loop levels (**f**) and increases ncRNA levels (**g**) at the IGS. **h**, Design details for the RED/dRED–LaSRR systems created to achieve inducible locus-associated R-loop repression. The zeocin resistance gene (zeo^R) was used for stable cell line generation, and the blasticidin-resistance gene (*blast*) for selection of the tetracycline repressor

(TetR). **i, j**, Validation of noninducible and tetracycline-inducible RED and dRED protein expression using immunoblotting (**i**) and microscopy (**j**). For gel source data, see Supplementary Fig. 1. **k**, Using RED together with sgIGS28 decreases R-loop levels at IGS18. **l, m**, Using RED together with sgIGS38 fails to alter R-loop levels (**l**) or ncRNA levels (**m**) at IGS18. **n**, Using RED together with sgIGS28 does not alter Pol II enrichments across the IGS. **o**, The fusion-protein system can be used to preferentially enrich the dRED fusion protein at the 5′ pause site of the *ACTIN* locus. **p**, Use of the nonoverlapping sgRNAs targeting IGS28, individually instead of as a pool, failed to significantly repress R-loop levels at IGS18, arguing against nonspecific effects related to the RNase H1 moiety of RED or any of the gRNAs used. **a**–**p**, HEK293T cells were used unless otherwise indicated. Data are means ± s.d.; one-way ANOVA with Dunnett's multiple comparisons test (**p**; $n = 3$ biologically independent experiments) or two-tailed t-test (**d**, **f**–**g**, **k**, **l**–**n**; $n = 3$ biologically independent experiments); $n = 3$ biologically independent experiments (**o**); images in **a**, **b**, **i**–**j** are representative of two independent experiments. Scale bars, 5 μm.
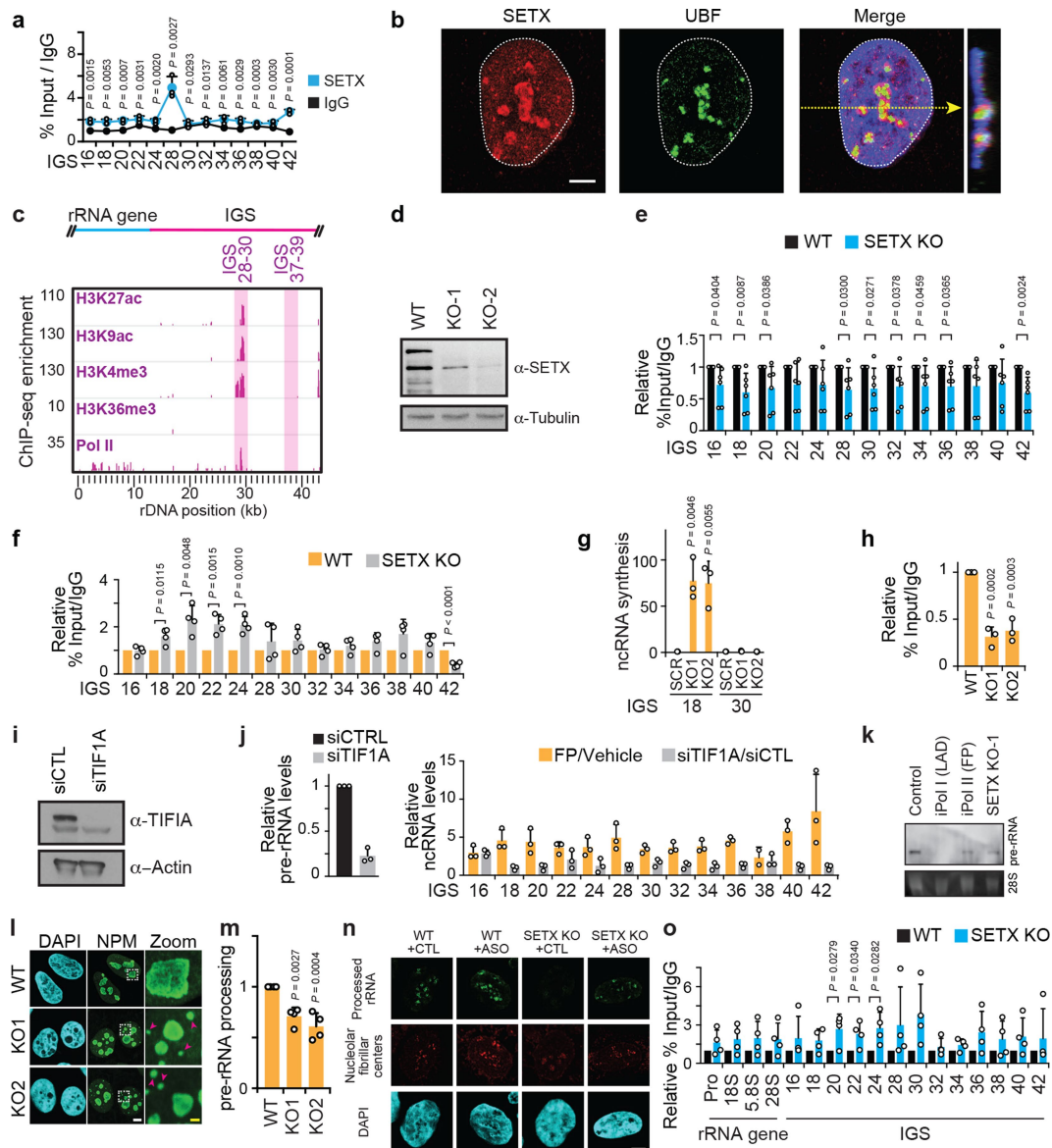
Extended Data Fig. 8 | Nucleolar and IGS features of wild-type and SETX-knockout cells. a, ChIP showing SETX enrichment at the IGS. **b**, SETX has a nucleolar/nucleoplasmic localization. **c**, Bioinformatic analysis of ENCODE-K562 data, showing coenrichment of epigenetic marks consistent with transcriptional activation near IGS28. **d**, Immunoblot showing CRISPR/ Cas9-mediated SETX knockout (KO). **e**, ChIP showing Pol II enrichment across rDNA in wild-type and SETX-KO cells. **f**, ChIP reveals that SETX KO, in two clones, enriches RNA Pol I at the IGSs. **g**, **h**, SETX KO induces IGS ncRNA synthesis (**g**) and decreases Pol I enrichment at the rRNA gene (5'-ETS region) (**h**). **i**, **j**, siRNA-mediated knockdown of TIF1A lowers Pol-I-dependent pre-rRNA levels but fails to induce IGS ncRNAs. Because of differences in experimental design, FP/vehicle data (**j**) were from a different experiment (Extended Data Fig. 6d) but are shown here for better visual comparison. **k**, Northern blotting reveals that Pol II or SETX disruption does not induce rRNA gene read-through transcripts. A probe for the 5'-ETS of pre-rRNA was used. **l**, **m**, SETX KO disrupts

nucleolar organization as indicated by NPM immunofluorescence (**l**), and decreases pre-rRNA processing in pulse-chase assays (**m**). **n**, ASO-mediated knockdown of sincRNAs increases rRNA biogenesis, as indicated by single-cell rRNA biogenesis assays. Shown are nucleolar fibrillar-centre-associated RNA rings revealed by single-cell FU-RNA pulse-chase immunofluorescence. Quantification shown in Fig. 4e. **o**, ChIP showing H3K9me2 enrichment across rDNA in wild-type and SETX-KO cells. **a**–**o**, HEK293T cells were used unless otherwise indicated. Data in **e**, **o** were from large experimental sets sharing IgG controls. Data are means ± s.d.; two-tailed t-test, $n = 3$ biologically independent experiments (**a**, **j**), $n = 6$ biologically independent experiments (**e**), and $n = 4$ biologically independent experiments (**f**, **o**); one-way ANOVA with Dunnett's multiple comparisons test, $n = 3$ biologically independent experiments (**g**, **h**) and $n = 4$ biologically independent experiments (**m**); images in **b**–**d**, **k** are representative of two independent experiments. Scale bars, 5 μm. For gel source data (**d**, **i**, **k**), see Supplementary Fig. 1.
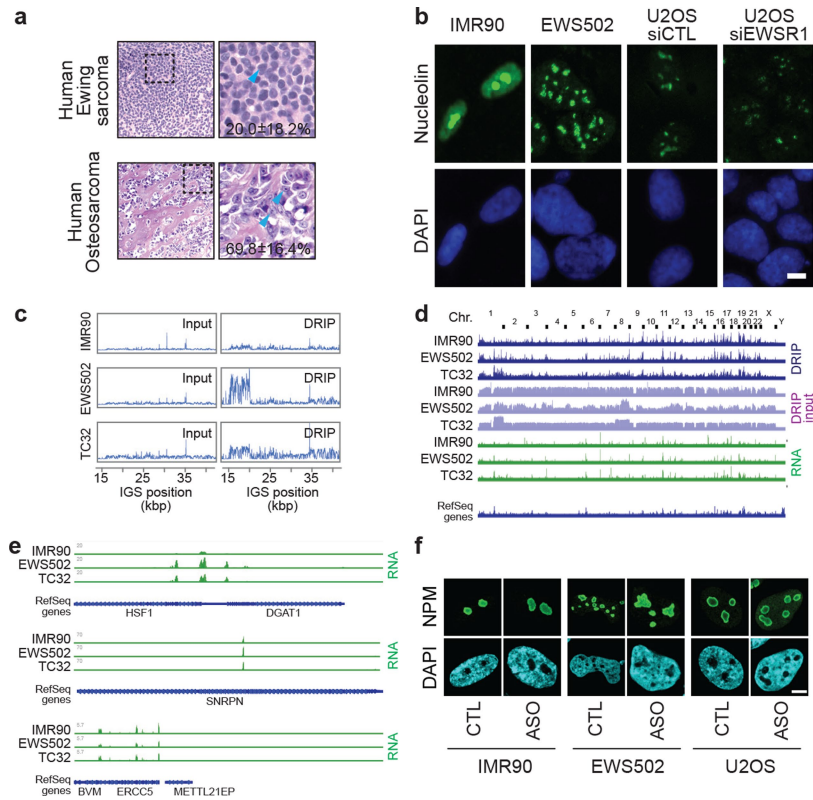
Extended Data Fig. 9 | Additional nucleolar organization and sequencing analyses related to Ewing sarcoma. a, Representative tissue sections of human Ewing sarcoma and osteosarcoma (haematoxylin and eosin staining; magnification ×400). Materials were obtained following Institutional Research Ethics Board approval (Sinai Heath Systems, 17-0103-E). The percentages of cells with one or two distinct nucleoli per nucleus are shown. Data are means ± s.d.; per cancer type, *n* = 5 cases (100 cells each); two-tailed *t*-test *P* = 0.0019. **b**, Ewing sarcoma cells (EWS502 cells) and U2OS cells with siEWSR1 display disrupted nucleoli, as indicated by the nucleolin protein, compared to their respective control IMR90 and U2OS siControl (siCTL) cells. Scale bar, 5 μm. **c**, Ewing sarcoma (EWS502 and TC32) cells showed increased R-loop levels across IGSs in DRIP–seq. **d**, Genome-wide view of sequence read alignments for DRIP–seq and RNA-seq. Chr., chromosome. **e**, IMR90, EWS502 and TC32 cells can exhibit similarities and differences at non-rDNA loci in sequencing read alignments from RNA-seq. **f**, ASOs targeting sincRNAs ameliorate nucleolar organization. Shown are representative images related to the quantifications in Fig. 4h. Images are representative of two independent experiments. Scale bar, 5 μm.
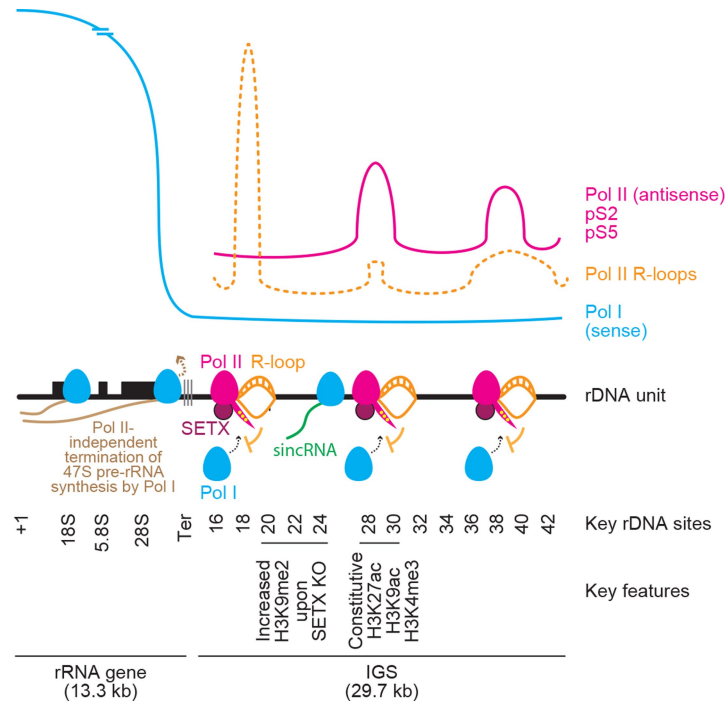
Extended Data Fig. 10 | Detailed model illustrating how nucleolar Pol-II-dependent R-loops shield the IGS from sincRNA synthesis by Pol I. Top and centre, Pol II at rDNA intergenic spacers (IGSs) synthesizes antisense intergenic ncRNAs (asincRNAs) that constitutively engage in R-loops containing DNA–RNA hybrids (orange). Centre, nucleolar Pol II function is promoted by the neurodegeneration-linked SETX protein (purple). Within rRNA genes, the formation of R-loops usually inhibits the function of Pol I, which is subject to Pol II-independent termination. However, disruption of nucleolar Pol II or its R-loops enables the recruitment of Pol I to IGSs. There, Pol I synthesizes sense intergenic ncRNAs (sincRNAs; green) that mimic environmental stress, disrupting nucleolar liquid–liquid phase separation and triggering an aberrant nucleolar liquid-to-solid phase transition. This unscheduled activation of nucleolar stress responses compromises the natural organization of nucleoli, leading to defects in pre-rRNA biogenesis, especially at the processing level. Nucleolar sincRNA levels are naturally elevated in Ewing sarcoma cells, explaining the indistinct nucleoli often seen in this cancer. In the context of Pol II inhibition, SETX loss or Ewing sarcoma, sincRNA repression ameliorates nucleolar organization and rRNA biogenesis.

| Corresponding author(s): | Karim Mekhail, Ph.D. |
| Last updated by author(s): | Apr 7, 2020 |

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist .

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

| Data collection | 1) We used a pipeline to align sequencing reads to human rDNA IGS, as described [PMID: 21355038]. A version of the human genome build hg19 with rDNA sequence is built using the Bowtie package (version 1.2.2). The newly built genome assembly is called "hg19_plus_rDNA". The human rDNA sequence FASTA file is obtained as is from NCBI. U13369.1 is the GenBank Accession ID and refers to the "Human ribosomal DNA complete repeating unit" as can be seen on NCBI. This FASTA file along with those from Chr 1-22, X, Y & M from hg19 obtained from UCSC are used to build the new assembly. Next, as control, HeLa Pol II reads are first aligned to this new genome assembly using the Bowtie package aligner. The reads from two replicates are obtained from ENCODE and concatenated. Duplicate reads are removed with the package BBmap (version 37.80) and its clumpify tool. Then, the alignment is performed with the parameter "-m 1" that instructs bowtie to refrain from reporting any alignments for reads having more than 1 reportable alignment. This ensures that only uniquely aligning reads are reported. The alignments are processed further with Samtools (version 0.1.19-44428cd) to retain only the reads aligning to the rDNA sequence, and to compute the depth/number of reads at each position in the rDNA coordinates. These depths are plotted with an R script. The R Project for Statistical Computing (version 3.6.1) from CRAN was used for analysis of IGS read counts across samples.

2) Next, we conducted GC skew calculations across rDNA using R software (version 3.4). Using the ~43K bp rDNA sequence obtained from the rDNA sequence FASTA file, GC skew, CG observed/expected ratio and GC% were assessed using (1 bp-at-a-time) sliding windows of size 50, 500 or 1000 bp. Definitions were as follows: GC skew = (number of Gs - number of Cs)/(number of Gs + number of Cs); CG observed/expected ratio = sliding window length * number of CpGs / (number of Cs X number of Gs); GC % = 100*(number of Gs + number of Cs)/sliding window length. To obtain an overall value/quantification and statistic to compare coding and IGS region, the mean GC skews for the coding and IGS regions with window size 1000 were obtained. In the coding region, the mean GC skew is 0.02346459 and in IGS it is -0.1541796. Doing a Welch Two Sample t-test on the GC skews from these two regions gives a p-value < 2.2x10-16. Script for all above analyses is called getGCskewEtc_rRDNA.R and is available upon request.

3) For Ewing sarcoma-related analyses, sample preparation and sequencing were done as described [PMID: 29513652]. RNA-sequencing and DRIP-sequencing data sources were as described in the Methods section. Identification of rDNA IGS peaks from RNA-seq and DRIP-seq were conducted as per the above described pipeline including a normalization of called peaks to the total number of reads per sample. Tool part of deepTools (version 3.3.0) was used to normalize the .bam files for comparison across samples and Integrated Genome Viewer (version 2.8.0) was used for visualization of .bam files (aligned sequence files). |

4) For strand-specific RNA-seq with/without heat shock, sequencing was performed on a non-ribosomal RNA depleted cDNA library of total RNA using stranded paired-end reads as described [PMID: 30110628]. After discarding reads mapped to the rRNA gene (including 5'ETS, IRS1/2, and 3'ETS), the remaining reads were mapped to GRCh38. BAM files were separated into forward and reverse strand files (bash script). The remaining reads aligned to supercontig GL000220.1 that is within the latest human genome assembly that contains a 43 kb ribosomal DNA cassette. Signals are normalized to an internal non-stress responsive control site at IGS35 as described [PMID: 30110628]. The sequencing data source was as described in the Methods section.

5) BioRad CFX Manager (version 3.1) was used for the collection of qPCR data for ChIP and DRIP experiments

6) NIS-Elements AR (version 4.10) was used to acquire microscopy images in DRIF, A-body staining, and endogenous protein immunofluorescence experiments

7) Leica TCS SP5 confocal laser scanning microscope, which uses software platform Leica Application Suite AF (advanced fluorescence) version 2.0.2 was used to collect images in stress induced nucleolar droplets and amyloid bodies experiments. Photoshop (version 20.0.4 CC2018) was used uniformly to adjust brightness and contrast of the images.

8) Zeiss AxioObserver D1 microscope, which uses software platform Zen Blue 2.3 was used for in vitro droplet formation experiments

Data analysis	GraphPad Prism (version 7.0e) was used to display data and perform statistical analyses. ImageJ (version 1.52a) was used to quantify single cell RNA pulse chase and phase contrast imaging experiments. MetaMorph analysis software (version 7.10.3) was used to measure signal intensities in FRAP experiments. Adobe Photoshop CS6 (version 13.0 x64) and Adobe Illustrator CS6 (version 16.0.4) were used to prepare figures for publication.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Data are in the Article, Supplementary Fig. 1 (uncropped blots), and the Source Data files related to Figs. 1-4 and extended data Figs. 1-3 and 5-8. All data and materials are available upon reasonable request. In light of the pandemic, shipping of reagents and materials may be slighted delayed.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

[✗] Life sciences   [ ] Behavioural & social sciences   [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to predetermine sample size. All experiments were conducted with cell lines with multiple available biological or technical replicates as specified in the manuscript, based on previous experience with specific experimental setups, and conforming to field standards. For single cell microscopy experiments, cell counts used per experiment reflect numbers routinely used in stringent quantitative cell biological experiments.
Data exclusions	Exclusion criteria were pre-determined based on internal controls and quality control indicators. For example, any experiment requiring transfection was assessed for successful transfection in parallel before inclusion in data analysis.
Replication	Several observations were tested for their generalizabiliy by 1) assessing multiple cell lines and 2) where applicable multiple knockout clones or chemical inhibitors to rule out clone-specific, cell line-specific, and reagent-specific artifacts. Reproducibility was confirmed by using suitable internal controls to establish validity and replication of findings in biological and technical replicates as indicated in the manuscript.
Randomization	Randomization was not part of the experimental design.
Blinding	Blinding was used for the quantification of microscopy images.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	
	1) Mouse IgG monoclonal Concentration: 1 mg/mL Supplier: Sigma-Aldrich/Millipore Cat #: 12-371 Application: ChIP, DRIP Dilution/Usage: ChIP (5 ug per sample), DRIP (10 ug per sample) Lot#: 3307779, 3267938 2) Rabbit IgG polyclonal Concentration: 1 mg/mL Supplier: Abcam Cat#: Ab171870 Application: ChIP Dilution/Usage: 5 ug per sample Lot#: GR3228514 3) H3K9me monoclonal Concentration: 1 mg/mL Supplier: Abcam Cat#: mAbcam 1220 Application: ChIP Dilution/usage: 5 ug per sample Lot# GR3228498 4) RNase H1 polyclonal Concentration: 24 ug/150 ul Supplier: Proteintech Cat#: 15606-1-AP Application: ChIP Dilution/usage: 5 ug per sample Lot#: 00043690 5) UBF (F-9) monoclonal Concentration: 200 ug/mL Supplier: Santa Cruz Cat#: sc-13125 Application: IF Dilution/usage: 1:10 Lot#: H0715, 12413 6) Pol I/RPA135 (N-17) Concentration: 200 ug/mL Supplier: Santa Cruz Cat#: sc-17913 Application: ChIP Dilution/usage: 5 ug per sample Lot#: F1714 7) NPM/B23 monoclonal Clone: FC82291 Concentration: 0.5-0.6 mg/mL

Supplier: Sigma-Aldrich
Cat#: B0556
Application: IF
Dilution/usage: 1:250
Lot#: IC52771

8) Senataxin polyclonal A
Concentration: 1mg/mL
Supplier: Bethyl Laboratories
Cat#: A301-104A
Application: ChIP, ChIP-Re-ChIP, WB
Dilution/usage: ChIP (5 ug per sample), ChIP-re-ChIP (5 ug per sample), WB (1:1000)

9) Senataxin polyclonal B
Concentration: 0.54 mg/mL
Supplier: Novus Bio
Cat#: NBP1-94712
Application: IF
Dilution/usage: 1:250
Lot#: A-1

10) GFP
Concentration: 5 mg/mL
Supplier: Abcam
Cat#: Ab290
Application: WB, ChIP
Dilution/usage: WB (1:1000), ChIP (5 ug per sample)
Lot#: GR3196305

11) RNA-DNA hybrid
Clone: S9.6
Concentration: 1 mg/mL
Supplier: prepared in house by Mekhail lab
Source: ATCC hybridoma (Cat# HB-8730, lot#62851141)
Application: DRIF, DRIP, IF
Dilution/usage: DRIF (1:500), DRIP (10 ug per sample), IF (1:500)

12) BrdU
Clone: BU-33
Concentration: 1 mg/mL
Supplier: Sigma-Aldrich
Cat# B2531
Application: single cell pulse chase
Dilution/usage: 1:250
Lot# 038M4861V

13) RNA polymerase II CTD repeat YSPTSPS (pS2)
Concentration: 1 mg/mL
Supplier: Abcam
Cat#: Ab5095
Application: ChIP, WB, IF
Dilution/usage: ChIP (5 ug per sample), WB: (1:1000), IF (1:600)
Lot#: GR3278442, GR3225147, GR3172948, GR231750

14) RNA polymerase II CTD repeat YSPTSPS (pS5)
Concentration: 1 mg/mL
Supplier: Abcam
Cat#: Ab5048
Application: ChIP
Dilution/usage: 5 ug per sample
Lot#: GR205997

15) ATXN2 polyclonal
Concentration: 200 ug/mL

Supplier: Sigma-Aldrich
Cat# HPA021146
Application: IF
Dilution/usage: 1:250
Lot#: A113803

16) dsRNA J2
Clone: rJ2
Concentration: 1 mg/mL
Supplier: Sigma-Aldrich/Millipore
Cat# MABE1134-100UL
Application: DRIF
Dilution/usage: 1:600
Lot#: 3170762

17) RRN3/Tif1A polyclonal
Concentration: 1 mg/mL
Supplier: Abcam
Cat# ab112052
Application: WB
Dilution/usage: 1:1000
Lot#: GR251820

18) Beta-Actin
Clone: mAbGEa
Concentration: 1 mg/mL
Supplier: Invitrogen/Thermo Fisher
Cat# MA1-744
Application: WB
Dilution/usage: 1:1000
Lot#: UB272750

19) CDK9 polyclonal
Supplier: Proteintech
Cat# 11705-1-AP
Application: ChIP
Dilution/usage: 5 ug per sample
Lot# 00047991

Validation	Commercially available antibodies were validated for specificity by the manufacturer using knockdown or knockout of cognate transcript/gene. The SETX antibody was additionally validated for specificity using CRISPR/Cas-mediated knockout of SETX. In addition, the specificity of our S9.6 antibody for RNA-DNA hybrids was validated using in vitro treatment with RNase H1, in vivo over-expression of RNase H1, RED-mediated signal repression, and dRED-mediated signal amplification. RNase H1 controls are also included in individual experiments to ensure that signals reflect RNA-DNA hybrids.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HEK293T, HeLa, HAP1, IMR90, MCF7 and U2OS cell lines were purchased from ATCC. HEK293T T-RExTM cells were purchase from ThermoFisher Scientific. EWS502 cells were from Dr. A. J. R. Bishop, who had previously obtained the cells from Dr. S. Lessnick. TC32 cells were also from Dr. A. J. R. Bishop, who had previously obtained them from the Children's Oncology Group.
Authentication	Purchased cell lines were commercially authenticated by ATCC or ThermoFisher Scientific. Cells obtained from Dr. A. J. R. were previously authenticated [PMID: 29513652]. Specifically, Following sequencing, identity was confirmed using known mutations in the cell lines, in addition to performing STR profiling on TC32 and U2OS. For all cell lines, cutlures were not maintained for more than 6 months prior to returning to low passage stocks.
Mycoplasma contamination	The cell lines used tested negative for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	DRIP-seq, which is similar to ChIP-seq, was used. Raw and processed files used for alignment to rDNA IGS are previously reported [PMID: 29513652] and described in the Methods section.
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	Experiments were previously done with biological replicates (EWS502 and TC32 Ewing cell lines) [PMID: 29513652].
Sequencing depth	The analysis conducted in our paper is based on published data that were obtained from samples that were amplified by PCR through 40 cycles. These samples were then processed as 50bp single-end sequencing and sequenced with 30-50 million reads for each sample.
Antibodies	Antibody against: RNA:DNA hybrids Supplier: Kerafast Catalog Number: ENH001 Clone number: S9.6 Validation: The specificity of the antibody was validated by: (1) References listed on the company website (https://www.kerafast.com/product/1552/anti-dna-rna-hybrid-s96-antibody), (2) Using samples treated with RNaseH1 to demonstrate specificity to mark RNA-DNA hybrids, and (3) qPCR on known R-loop sites as well as sites that are known to not have R-loops.
Peak calling parameters	The alignment is performed with the parameter "-m 1" that instructs bowtie to refrain from reporting any alignments for reads having more than 1 reportable alignment.
Data quality	Data quality was controlled with FDR < 0.05.
Software	We used a published pipeline for aligning sequencing reads to human rDNA IGS [PMID: 21355038]. A version of the human genome build hg19 with rDNA sequence is built using the Bowtie package. The resulting genome assembly is called "hg19_plus_rDNA". The human rDNA sequence FASTA file is obtained as is from NCBI. U13369.1 is the GenBank Accession ID and refers to the "Human ribosomal DNA complete repeating unit" as can be seen on NCBI. This FASTA file along with those from Chr 1-22, X, Y & M from hg19 obtained from UCSC are used to build the new assembly. Next, sequencing reads are aligned to this new genome assembly using the Bowtie package aligner. Duplicate reads are removed with the package BBmap and its clumpify tool. Then, the alignment is performed with the parameter "-m 1" that instructs bowtie to refrain from reporting any alignments for reads having more than 1 reportable alignment. This ensures that only uniquely aligning reads are reported. The alignments are processed further with Samtools to retain only the reads aligning to the rDNA sequence, and to compute the depth/number of reads at each position in the rDNA coordinates. Signals are normalized to the total number of reads per sample. The normalized depths are plotted with an R script.