EPB41L4A-AS1 long noncoding RNA acts in both *cis*- and *trans*-acting transcriptional regulation and controls nucleolar biology

Alan Monziani[1,2], Juan Pablo Unfried[1,2], Todor Cvetanovic[1,2] and Igor Ulitsky[1,2]

[1]Department of Immunology and Regenerative Biology and [2]Department of Molecular Neuroscience, Weizmann Institute of Science, 7610001 Rehovot, Israel

Abstract

Mammalian genomes are pervasively transcribed into long noncoding RNAs (lncRNAs), whose functions and modes of action remain poorly understood. *EPB41L4A-AS1* is an evolutionary conserved, broadly and highly expressed lncRNA that produces the H/ACA snoRNA *SNORA13* from one of its introns. We studied the consequences of *EPB41L4A-AS1* perturbation in breast cancer cells and found that it acts both *in cis*, to enhance transcription of the proximal *EPB41L4A* gene and additional genes in its two flanking topologically associated domains, and in *trans* by broadly regulating gene expression, including expression of snoRNAs, transcription of genes involved in nucleolar biology and the distribution of nucleolar proteins. These effects are phenocopied by the loss of SUB1, an interactor of *EPB41L4A-AS1*, and are observed following transient perturbations of *EPB41L4A-AS1* that do not affect steady-state *SNORA13* levels or the rRNA modification it helps install. Exogenous expression of the full-length *EPB41L4A-AS1* locus but not *SNORA13* expression can rescue the *trans*-acting transcriptional effects of its perturbation. The *EPB41L4A-AS1* gene is thus a versatile locus producing RNA molecules acting on multiple levels for key cellular functions.

Introduction

Long noncoding RNAs (lncRNAs) are a highly heterogeneous group of RNA molecules, defined as being >500 nt long and not displaying a significant protein-coding potential[1–4]. Given this broad definition and the pervasive transcription of the human genome[5–8], lncRNAs are now the RNA class with the largest number of annotated genes[9]. Long noncoding RNAs exhibit a wide range of expression levels, driven by varying transcription rates and half-lives. However, only a limited number of lncRNAs have well-defined mechanisms of action. Although small peptides arising from short and poorly conserved translated open reading frames (ORFs) have been reported in several lncRNAs[10–13] and studies using ultra-deep sequencing of ribosome profiling data and deep proteomic analysis, found evidence of translation of ORFs in hundreds of lncRNAs[14]. Yet it remains unclear how many of these are encoding functional peptides, mainly because these ORFs are typically poorly conserved and the protein products do not appear to be abundant[15]. Similar to protein-coding mRNAs, lncRNAs are transcribed by RNA polymerase II (Pol II), 5' capped with 7-methyl guanosine (m^7G), polyadenylated at their 3' ends, and have a canonical exon–intron structure.

One way to functionally classify lncRNAs is by the nature of the functional elements harbored within their loci. One class comprises genes in which the DNA element has a function unrelated to the transcription of the locus[16]. The other includes cases in which the act of transcription is functionally relevant, while the resulting RNAs are "byproducts" of Pol II activity, which, for example, can tune the expression of another locus in 3D proximity[1,17–19]. This class is potentially broad because it can explain the prevalence of lncRNA loci in which the lncRNAs are expressed

at very few copies and/or have sequences that are evolving without any apparent selection. For the cases where the RNA product is functional, a further distinction can be made between *cis*-acting lncRNAs, the function of which is restricted to the locus from which they are transcribed, and *trans*-acting lncRNAs, which are transcribed, processed, and then move away by diffusion or active transport to exert their function elsewhere. The latter RNAs require a relatively stable RNA molecule because they have to interact with one or more other RNA species and/or proteins, whereas, for *cis*-acting lncRNAs, such an RNA molecule may or may not be necessary, with the act of transcription and concomitant chromatin remodeling reported to be sufficient in several cases[20].

Small nucleolar RNAs (snoRNAs) are intermediate-sized RNAs whose main function is to guide RNA modifications at other RNAs, mostly ribosomal RNAs. In mammalian cells, snoRNAs are produced mainly from introns of polyadenylated transcripts, some of which code for proteins, and others are lncRNAs. Noncoding snoRNA host genes (or ncSHNG) are characterized by several characteristic features, including high and broad expression levels, relatively inefficient splicing, and sensitivity to nonsense-mediated decay [21]. While lncRNA functions have been shown for some ncSNHGs, most remain poorly understood. Here, we describe a bioinformatic screen that converged on *EPB41L4A-AS1*, a very abundant and conserved lncRNA acting as a *cis*-acting regulator, and our observations that this lncRNA also works in *trans* to regulate nucleolar biology.

Results

A computational screen for putative *cis*-acting lncRNAs

We hypothesized that lncRNAs likely to act in *cis* will be occasionally co-regulated with their target genes and also found in spatial proximity to the target promoters. We sought such lncRNA-target pairs in the GeneHancer database[22], that connects regulatory elements and their targets by combining tissue co-expression data, enhancer-enriched histone modifications, transcription factor binding, Hi-C data, and eQTLs. We complemented this resource with RNA-seq datasets from the breast cancer cell line MCF-7 exposed to different treatments[23–25] (**Fig. 1A** and **Table S1**). This analysis (see Methods) led us to focus on four lncRNA-target pairs connected in GeneHancer and in which both genes were significantly either upregulated or downregulated in the treatment conditions, compared to controls (**Fig. S1A-D**). We knocked down (KD) each candidate with two distinct GapmeRs—antisense oligonucleotides with a DNA core that induces the degradation of target nascent RNA via RNAse H activity. We then tested the expression of the target genes using RT-qPCR. Among the four lncRNAs, KD of *EPB41L4A-AS1* significantly repressed *EPB41L4A* (**Fig. 1B**), whereas the other lncRNAs had no substantial effects on the putative target genes.

EPB41L4A-AS1 is a complex lncRNA gene

The *EPB41L4A-AS1* locus is ~2 kb (kilobases) long, and the main transcribed isoform terminates ~300 nt from the 3' end of the main isoform of the convergently transcribed *EPB41L4A* gene, and >250 kb away from the *EPB41L4A* transcription start site (TSS, **Fig. 1C**). An H/ACA snoRNA, *SNORA13*, a homolog of yeast *snR35*, is embedded within the first intron of *EPB41L4A-AS1*. This snoRNA guides pseudouridylation on position 1248 of the small subunit 18S rRNA[26]. This modification is the first step in the installation of one of the most complex RNA modifications—N1-methyl-N3-aminocarboxypropyl-pseudouridine (m(1)acp(3)Ψ)—that is located next to the tRNA P-site and is often lost in human cancers[26]. Recently, *SNORA13* has been reported to negatively regulate ribosome biogenesis by decreasing the incorporation of RPL23 into maturing 60S subunits, eventually triggering p53-mediated cellular senescence[27].

Several groups reported different functions of *EPB41L4A-AS1*. It was found to be differentially expressed across the cell cycle, and its KD affected cell cycle progression and reduced proliferation in multiple cancer cell lines[28]. Another study implicated *EPB41L4A-AS1* in glucose metabolism through interactions with HDAC2 and NPM1[29]. Silencing of *EPB41L4A-AS1* led to HDAC2 translocation into the nucleoplasm, where it binds the *VDAC1* and *VHL* promoter regions, among others[29], and effects on the expression of additional genes were reported in other systems as well[30–33]. In human cancers, *EPB41L4A-AS1* is also downregulated and correlates with a poor prognosis in a variety of cancers, including breast cancer (BRCA), which we also observed by analyzing the TCGA data (**Fig. S2**), aligning with the widespread loss of the m(1)acp(3)Ψ modification. *EPB41L4A* instead is a poorly studied gene, reported to be involved in predisposition to papillary thyroid carcinoma by modulating the WNT/β-catenin pathway[34,35], possibly playing a role in connecting the cytoskeleton to the plasma membrane, as well as throughout development[36].

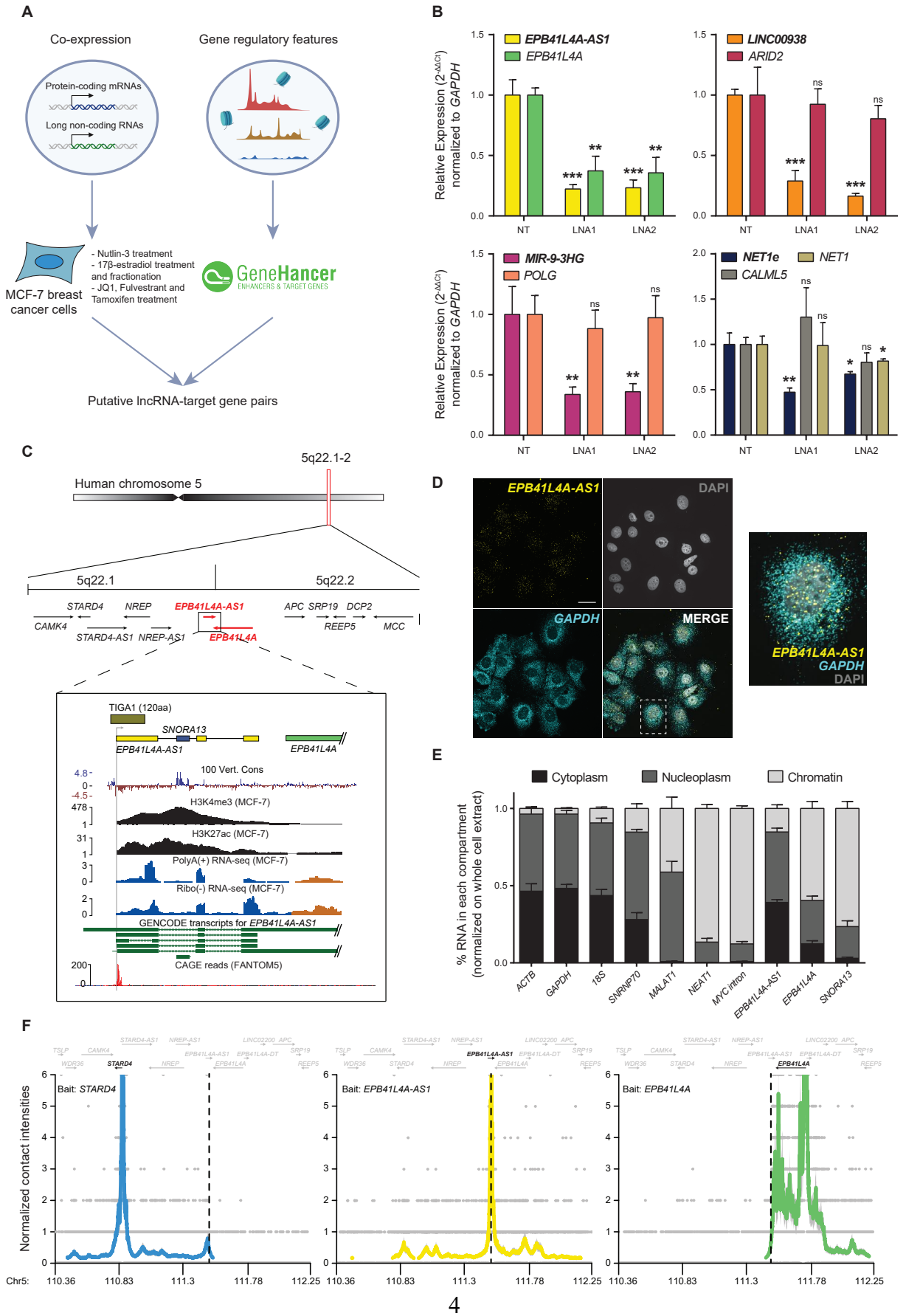

A
Co-expression
Gene regulatory features
Protein-coding mRNAs
Long non-coding RNAs
MCF-7 breast cancer cells
- Nutlin-3 treatment
- 17β-estradiol treatment and fractionation
- JQ1, Fulvestrant and Tamoxifen treatment
GeneHancer
ENHANCERS & TARGET GENES
Putative lncRNA-target gene pairs
B
Relative Expression (2^-ΔΔCt) normalized to GAPDH
EPB41L4A-AS1
EPB41L4A
LINC00938
ARID2
MIR-9-3HG
POLG
NET1e
NET1
CALML5
NT
LNA1
LNA2
ns
C
Human chromosome 5
5q22.1-2
5q22.1
5q22.2
STARD4
NREP
EPB41L4A-AS1
APC
SRP19
DCP2
CAMK4
STARD4-AS1
NREP-AS1
EPB41L4A
REEP5
MCC
TIGA1 (120aa)
SNORA13
100 Vert. Cons
H3K4me3 (MCF-7)
H3K27ac (MCF-7)
PolyA(+) RNA-seq (MCF-7)
Ribo(-) RNA-seq (MCF-7)
GENCODE transcripts for EPB41L4A-AS1
CAGE reads (FANTOM5)
D
DAPI
GAPDH
MERGE
E
Cytoplasm
Nucleoplasm
Chromatin
% RNA in each compartment (normalized on whole cell extract)
ACTB
GAPDH
18S
SNRNP70
MALAT1
NEAT1
MYC intron
EPB41L4A-AS1
EPB41L4A
SNORA13
F
Bait: STARD4
Bait: EPB41L4A-AS1
Bait: EPB41L4A
Normalized contact intensities
Chr5:
110.36
110.83
111.3
111.78
112.25

Figure 1. *EPB41L4A-AS1* is a highly expressed lncRNA with a widespread cellular distribution. (A) Graphical overview of the screen to identify candidate *cis*-acting lncRNAs. **(B)** RT-qPCR to assess the expression of each candidate *cis*-acting lncRNA and their predicted *cis*-regulated target genes, following KD with two unique GapmeRs. **(C)** (Top) Overview of the 5q22.1-2 locus, with the bars highlighting the two TADs. (Bottom) Zoom in on the *EPB41L4A-AS1* locus, with tracks for the PhyloP conservation score (UCSC genome browser), H3K4me3 and H3K27ac histone modifications (ENCODE), PolyA+ and Ribo(-) RNA-seq coverage (this study), GENCODE transcripts and CAGE reads (FANTOM5). **(D)** A representative single-molecule RNA FISH (smFISH) image shows *EPB41L4A-AS1* intra-cellular localization. *GAPDH* and DAPI were used to label the cytoplasm and the nucleus, respectively (scale bar = 20 µm, 100X magnification). **(E)** RT-qPCR analysis of subcellular fractionation experiments. The percentages of RNA in each compartment were obtained by normalizing the expression in the different fractions to that in whole cells. **(F)** UMI-4C contact profiles using baits targeting the TSSes of *STARD4* (left), *EPB41L4A-AS1* (center), or *EPB41L4A* (right). The dotted line represents the center of the *EPB41L4A-AS1* locus. All experiments were performed in n=3 biological replicates, except UMI-4C with n=2, with the error bars in the barplots representing the standard deviation. ns = P>0.05; * = P<0.05,**=P<0.01 = **; ***=P<0.001 (two-sided Student's t-test).

EPB41L4A-AS1 is highly expressed in both the nucleus and cytoplasm and is regulated in multiple conditions

Using single-molecule fluorescence *in-situ* hybridization (smFISH) and subcellular fractionation we found that *EPB41L4A-AS1* displays both nuclear and cytoplasmic localization in MCF-7 cells (**Fig. 1D**), with a minor fraction associated with chromatin as well (**Fig. 1E**). In contrast, both *EPB41L4A* and *SNORA13* were mostly found in the chromatin fraction (**Fig. 1E**), the former possibly due to the length of its pre-mRNA (>250 kb), which would require substantial time to transcribe. Coding-Potential Assessment (CPAT)[37] analysis showed the sequence of *EPB41L4A-AS1* has little coding potential. A translated ORF in the first exon of *EPB41L4A-AS1* has been reported to produce a small peptide (TIGA1), which localizes to the mitochondria and induces growth arrest[38] (**Fig. 1C**). Applying CPAT to each exon separately, found evidence for some coding potential only in the first exon that harbors the TIGA1 peptide (**Fig. S3A**). Importantly, analysis of the RNA-seq signal coverage shows that the common *EPB41L4A-AS1* isoform in MCF-7 cells begins downstream of the AUG codon at the start of TIGA1, and TIGA1 doesn't have another in-frame AUG in its sequence. Consistently, the vast majority of CAGE tags at the *EPB41L4A-AS1* are also downstream of the AUG of TIGA1, consistent with the start sites of most isoforms of *EPB41L4A-AS1* annotated in GENCODE (**Fig. 1E**), We conclude that while some fraction of *EPB41L4A-AS1* might produce TIGA1 peptides, most RNA products of this gene in MCF-7 do not produce TIGA1 or other peptides predicted to be functional.

Consistently with a previous study[38], we found a strong upregulation of *EPB41L4A-AS1, EPB41L4A*, and *SNORA13* in MCF-7 cells upon serum starvation (**Fig. S3B**). In contrast with a previous report[28], we did not detect *EPB41L4A-AS1* to be differentially expressed across the cell cycle in MCF-7 cells (**Fig. S3C**). We further explored the effect of different stresses and environmental conditions on *EPB41L4A-AS1* by analyzing its expression in a published dataset comprising five different cell types exposed to a panel of 50 different substances[39] (**Fig. S3D**), as well as exposing MCF-7 cells to LPS (bacterial infection), H_2O_2 (oxidative stress), thapsigargin (ER stress and unfolded protein response) and etoposide (DNA damage) (**Fig. S3E**). Strikingly, *EPB41L4A-AS1* expression was induced in nearly all conditions, often together with *EPB41L4A* but not *SNORA13*, suggesting the levels of the snoRNA are typically detached from that of the lncRNA[40], likely because the latter is substantially more stable (see below).

EPB41L4A-AS1 regulates *EPB41L4A* transcription

We assessed the 3D chromatin architecture around the *EPB41L4A-AS1* locus by performing UMI-4C-seq[41] with baits targeting the lncRNA, *EPB41L4A* and the promoter of *STARD4*—a flanking gene found ~650 Kb upstream of the *EPB41L4A-AS1* promoter. *EPB41L4A-AS1* formed contacts with both genes (**Fig. 1F**, middle), with prominent peaks corresponding to the promoter of *EPB41L4A*. Interestingly, the promoter of *EPB41L4A* displays several contacts with its own gene body (**Fig. 1F**, right panel), consistent with a "stripe" in HiC data where a single anchor interacts with the whole topologically associated domain (TAD) at high frequency[42], with those frequencies reduced to below background levels upstream of the *EPB41L4A-AS1* locus. Conversely, contacts from the *STARD4* promoter also show a peak near *EPB41L4A-AS1* and then reduced below the background (**Fig. 1F**, left panel). These data indicate that *EPB41L4A-AS1* is located at a TAD boundary region, forming strong spatial contacts with sites upstream and downstream, including the promoter of *EPB41L4A*. In agreement with this, *EPB41L4A-AS1* is located between the 5q22.1 and 5q22.2 chromosomal bands, delimiting two TADs as evident in both our 4C-seq and Micro-C data from H1 cells (**Fig. 1F** and **2A**). Both *EPB41L4A* and *EPB41L4A-AS1* are characterized by CTCF binding near their promoters and their chromatin is extensively demarcated with H3K27ac (**Fig. 2A**). A recent study identified this boundary as one of the most insulated in human embryonic stem cells (hESCs)[43].

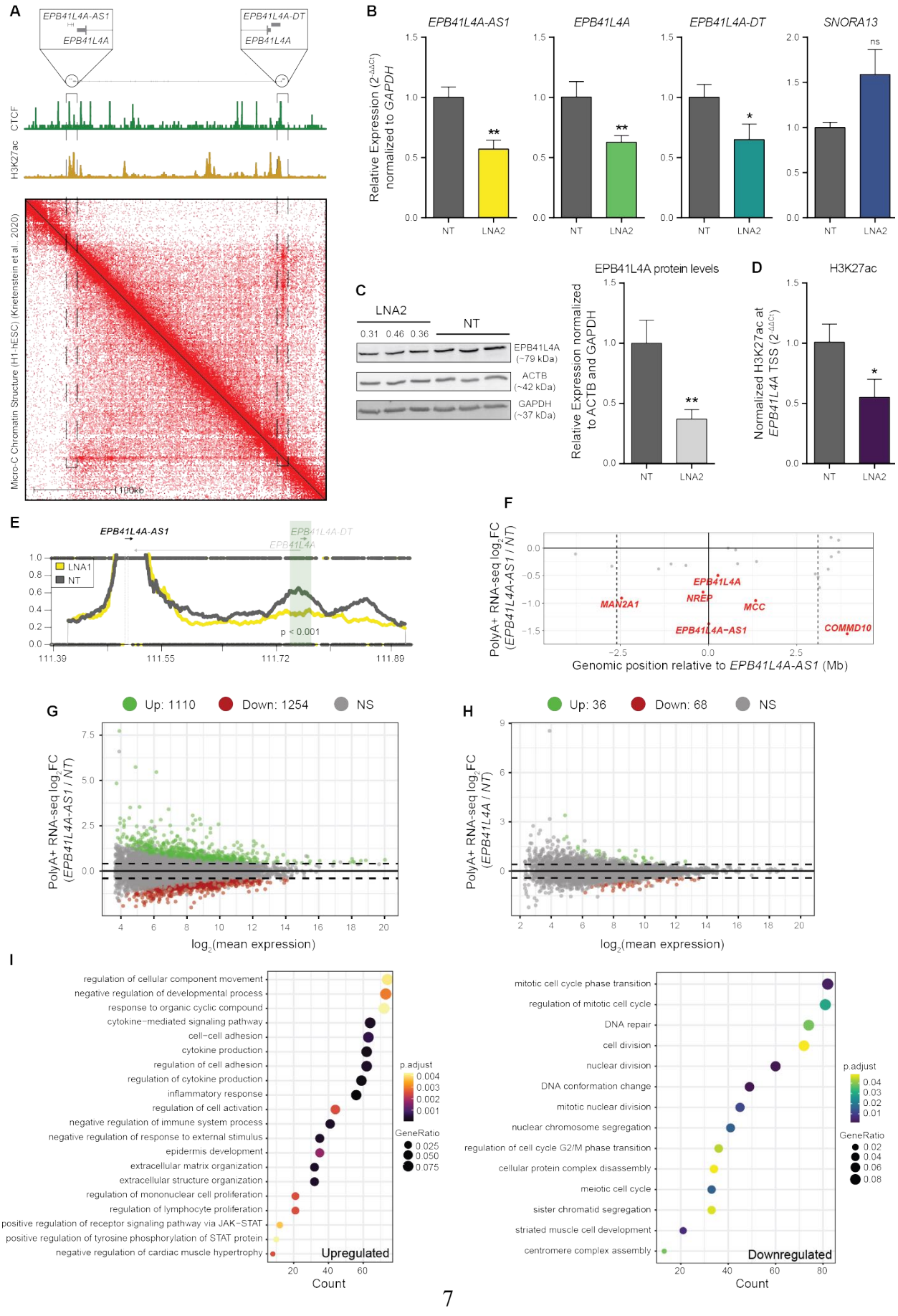

A
EPB41L4A-AS1
EPB41L4A
EPB41L4A-DT
EPB41L4A
CTCF
H3K27ac
Micro-C Chromatin Structure (H1-hESC) (Krietenstein et al., 2020)
100kb
B
EPB41L4A-AS1
EPB41L4A
EPB41L4A-DT
SNORA13
Relative Expression (2^-ΔΔCt) normalized to GAPDH
NT
LNA2
ns
C
LNA2
NT
0.31 0.46 0.36
EPB41L4A (~79 kDa)
ACTB (~42 kDa)
GAPDH (~37 kDa)
EPB41L4A protein levels
Relative Expression normalized to ACTB and GAPDH
D
H3K27ac
Normalized H3K27ac at EPB41L4A TSS (2^-ΔΔCt)
E
EPB41L4A-AS1
EPB41L4A-DT
EPB41L4A
LNA1
NT
p < 0.001
111.39
111.55
111.72
111.89
F
PolyA+ RNA-seq log2FC (EPB41L4A-AS1 / NT)
EPB41L4A
NREP
MAN2A1
MCC
EPB41L4A-AS1
COMMD10
Genomic position relative to EPB41L4A-AS1 (Mb)
G
Up: 1110
Down: 1254
NS
PolyA+ RNA-seq log2FC (EPB41L4A-AS1 / NT)
log2(mean expression)
H
Up: 36
Down: 68
NS
PolyA+ RNA-seq log2FC (EPB41L4A / NT)
log2(mean expression)
I
regulation of cellular component movement
negative regulation of developmental process
response to organic cyclic compound
cytokine-mediated signaling pathway
cell-cell adhesion
cytokine production
regulation of cell adhesion
regulation of cytokine production
inflammatory response
regulation of cell activation
negative regulation of immune system process
negative regulation of response to external stimulus
epidermis development
extracellular matrix organization
extracellular structure organization
regulation of mononuclear cell proliferation
regulation of lymphocyte proliferation
positive regulation of receptor signaling pathway via JAK-STAT
positive regulation of tyrosine phosphorylation of STAT protein
negative regulation of cardiac muscle hypertrophy
Upregulated
Count
p.adjust
GeneRatio
mitotic cell cycle phase transition
regulation of mitotic cell cycle
DNA repair
cell division
nuclear division
DNA conformation change
mitotic nuclear division
nuclear chromosome segregation
regulation of cell cycle G2/M phase transition
cellular protein complex disassembly
meiotic cell cycle
sister chromatid segregation
striated muscle cell development
centromere complex assembly
Downregulated
Count
p.adjust
GeneRatio

Figure 2. *EPB41L4A-AS1* is a *cis*-acting lncRNA affecting genome-wide gene expression. **(A)** (Top) In scale view of the *EPB41L4A-AS1-EPB41L4A* locus, with zoomed areas corresponding to the two TSS. (Middle) CTCF and H3K27ac coverage across this region. (Bottom) Micro-C data in H1-hESCs[44] show continuous contacts throughout the *EPB41L4A* gene body. **(B)** RT-qPCR and **(C)** Western blot for the indicated genes and proteins following *EPB41L4A-AS1* KD.**(D)** H3K27ac CUT&RUN-qPCR following *EPB41L4A-AS1* KD with GapmeRs using primers targeting the promoter of *EPB41L4A*. **(E)** UMI-4C contact profiles in control and LNA1-transfected cells using baits targeting the TSS of *EPB41L4A-AS1*. The green area represents the quantified genomic interval, and the p-value was calculated using a Chi-squared test. **(F)** Changes in gene expression for the genes in the two flaking TADs of the lncRNA in cells transfected with GapmeRs targeting *EPB41L4A-AS1*. The vertical dotted lines represent the TAD boundaries (as assessed by TADmap), the continuous vertical line the lncRNA locus and inter-TAD boundary, and the horizontal continuous line a $\log_2$Fold-change equal to 0. The dots represent individual genes, with the significant ones highlighted. **(G)** MA plot showing the changes in genome-wide gene expression in cells transfected with GapmeRs targeting *EPB41L4A-AS1*. **(H)** Same as (G), but with GapmeRs targeting *EPB41L4A*. **(I)** GO enrichment analysis for the upregulated (left) and downregulated (right) genes after *EPB41L4A-AS1* KD. All experiments were performed in n=3 biological replicates, except UMI-4C with n=2, with the error bars in the barplots representing the standard deviation. ns - P>0.05; * - P<0.05; ** - P<0.01; *** - P<0.001 (two-sided Student's t-test). A gene was considered to be differentially expressed if both adjusted P<0.05 and |$\log_2$Fold-change| >0.41 (corresponding to a change of 33%).

We then focused on a possible reciprocal regulation between *EPB41L4A-AS1* and *EPB41L4A*. *EPB41L4A-AS1* KD with antisense GapmerRs reduced the expression of *EPB41L4A* mRNA and its promoter-associated divergent lncRNA *EPB41L4A-DT* (**Fig. 2B**), as well as the EPB41L4A protein levels (**Fig. 2C**). CUT&RUN-qPCR of the H3K27ac mark—which is associated with active promoters and enhancers—showed a reduction of H3K27ac at the *EPB41L4A* promoter (**Fig. 2D**). In contrast to the effects of GapmeRs, which mostly act in the nucleus and on the nascent RNA[45,46], KD of *EPB41L4A-AS1* with siRNAs, which instead mainly act on the mature RNA in the cytoplasm, did not affect *EPB41L4A* levels (**Fig. S4A**). Levels of *EPB41L4A* were not affected by increased expression of *EPB41L4A-AS1* from the endogenous locus by CRISPR activation (CRISPRa), nor by its exogenous expression from a plasmid (**Fig. S4B** and **S4C**). Moreover, downregulation (**Fig. S4D** and **S4E**) or upregulation (**Fig. S4F** and **S4G**) of *EPB41L4A* did not have any consistent effects on *EPB41L4A-AS1* expression, suggesting a unidirectional effect of *EPB41L4A-AS1* transcription of its nascent RNA product on *EPB41L4A*. Consistently with *cis*-acting regulation, the reduced *EPB41L4A* mRNA levels after *EPB41L4A-AS1* knockdown were rescued only after CRISPRa of the endogenous *EPB41L4A-AS1* locus (**Fig. S4H** and **S4I**). To study the potential mode of action of this regulation, we inspected UMI-4C data anchored at the *EPB41L4A-AS1* promoter and found a reduction in the contact frequency with the *EPB41L4A* promoter upon treatment with *EPB41L4A-AS1*-targeting GapmeRs, that was statistically significant for LNA1 (**Fig. 2E** and **S4J**). We conclude that the transcription of the RNA product of *EPB41L4A-AS1* acts unidirectionally and in *cis* to promote the expression of *EPB41L4A*, possibly by ensuring spatial proximity between the TAD boundary and the *EPB41L4A* promoter.

Genome-wide transcriptional alterations upon depletion of *EPB41L4A-AS1*

To evaluate the global transcriptional response to *EPB41L4A-AS1* levels, we sequenced polyadenylated RNA from cells transfected with *EPB41L4A-AS1*-, *EPB41L4A*-targeting, or control GapmeRs by RNA-seq (**Fig. S5A** and **Table S2**). Consistently with the qRT-PCR data, KD of *EPB41L4A-AS1* reduced *EPB41L4A* expression, and also reduced expression of several,

but not all other genes in the TADs flanking the lncRNA (**Fig. 2F**). Based on these data, *EPB41L4A-AS1* is a significant *cis*-acting activator according to TransCistor[47] (P=0.005 using the digital mode). The *cis*-regulated genes reduced by *EPB41L4A-AS1* KD included *NREP*, a gene important for brain development, whose homolog was downregulated by genetic manipulations of regions homologous to the lncRNA locus in mice[43]. Depletion of *EPB41L4A-AS1* thus affects several genes in its vicinity.

Downregulation of *EPB41L4A-AS1* led to broad changes in gene expression beyond the proximal TADs. A total of 2,364 genes were differentially expressed (1,110 up, 1,254 down, adjusted $P<0.05$ and $|\log_2\text{Fold-change}| >0.41$) (**Fig. 2G**), in contrast to only 104 (36 up, 68 down) when *EPB41L4A* was targeted (**Fig. 2H**). A total of 25 (69%) of the upregulated, and 22 (32%) of the downregulated genes after *EPB41L4A* KD were also regulated in the same direction by the KD of *EPB41L4A-AS1* (**Fig. S5B,** $P<10^{-15}$ for both). In agreement with studies on *EPB41L4A* function[34,35], these genes were enriched with those functioning in cell-cell adhesion and cell surface receptor signaling, although with a borderline significance (**Fig. S4C**). Most of the consequences of the depletion of *EPB41L4A-AS1* are thus not directly explained by changes in *EPB41L4A* levels. An additional *trans*-acting function for *EPB41L4A-AS1* would be consistent with its high expression levels compared to most lncRNAs detected in MCF-7 (**Fig. S4D**). Several GO terms were found to be enriched among both up- and down-regulated genes following *EPB41L4A-AS1* KD. Upregulated genes were linked to cell-cell adhesion, inflammatory response and JAK/STAT/ERK pathway, and the downregulated ones to cell cycle, division and DNA repair, among others (**Fig. 2I**). When considering cellular components, the downregulated genes were associated with ribonucleoprotein and spliceosomal complexes, as well as to ribosome subunit and, interestingly, the nucleolus (**Fig. S6A** and **S6B**). Thus, since *EPB41L4A-AS1* is a snoRNA-host gene (SNHG) and snoRNAs typically localize to and act in the nucleolus, we aimed to investigate whether this lncRNA might interact with some factors that can explain the observed enrichment in nucleolar-associated genes.

***EPB41L4A-AS1* binds SUB1 and affects the expression of its target genes**

In order to identify potential factors that might be associated with *EPB41L4A-AS1*, we inspected protein-RNA binding data from the ENCODE eCLIP dataset[48]. The exons of *EPB41L4A-AS1* lncRNA were densely and strongly bound by SUB1 (also known as PC4) (**Fig. 3A**). SUB1 interacts with all three RNA polymerases and was reported to be involved in transcription initiation and elongation, response to DNA damage, chromatin condensation[49–52] and telomere maintenance[53,54]. SUB1 normally localizes throughout the nucleus in various cell lines, yet staining experiments show a moderate enrichment for the nucleolus (source: Human Protein Atlas; https://www.proteinatlas.org/ENSG00000113387-SUB1/subcellular). Another nucleolus-related protein, NPM1, which binds C/D-box snoRNAs and directs 2'-O-methylation of rRNAs[55], had a single moderately confident binding site within *EPB41L4A-AS1* intron overlapping *SNORA13* (**Fig. 3A**). Interestingly, both *SUB1* and *NPM1* mRNAs were significantly downregulated following *EPB41L4A-AS1* KD (**Fig. 3B**).

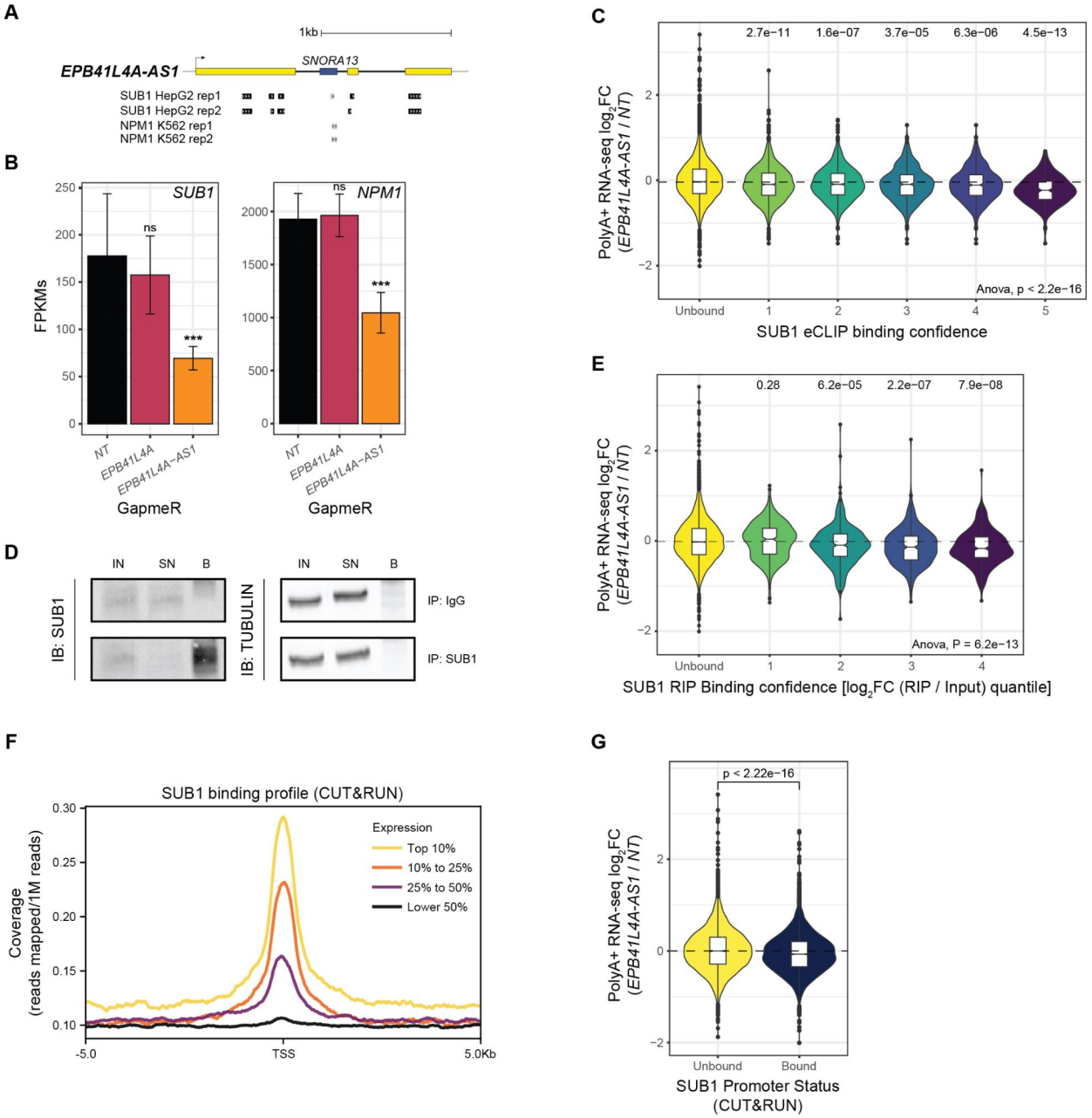


Figure 3. SUB1 interacts with *EPB41L4A-AS1* and affects gene expression at the DNA and RNA levels. **(A)** Schematics of the *EPB41L4A-AS1* locus with tracks depicting the eCLIP peaks for both SUB1 and NPM1 (source: ENCODE). **(B)** Average expression levels (in FPKM) of *SUB1* and *NPM1* in cells transfected with GapmeRs targeting either *EPB41L4A-AS1* or *EPB41L4A*, with the error bars representing the standard deviation across the n = 3 replicates. DESeq2 adjusted P-values compared to control GapmeR are also reported. **(C)** Changes in gene expression upon *EPB41L4A-AS1* KD of the genes ranked by the SUB1 eCLIP binding confidence. **(D)** Western blot following RIP using either a SUB1 or IgG antibody. TUBULIN was used as a negative control, IN - input, SN - supernatant/unbound, B - bound. **(E)** Same as (C), but with genes ranked by their enrichment in the RIP data ($\log_2$FC RIP / Input). **(F)** Metagene profile around TSS of the normalized SUB1 CUT&RUN signal, stratified by gene expression levels in MCF-7 cells. **(G)** Changes in gene expression upon *EPB41L4A-AS1* KD with GapmeRs for genes with and without a high-confidence SUB1 peak in their TSS. All experiments were performed in n = 3 biological replicates. In the boxplots, the thick line, edges of the box, and

whiskers represent the median, first and third quartiles, and the upper and lower 1.5 interquartile ranges (IQRs), respectively. Outliers (observations outside the 1.5 IQRs) are drawn as single points, the significance of the different comparisons was computed by a Mann-Whitney test, and a global ANOVA p-value is also reported. In all cases, ns - P>0.05; * - P<0.05; ** - P<0.01; *** - P<0.001.

We hypothesized that loss of *EPB41L4A-AS1* might affect SUB1, either via the reduction in its expression or by affecting its functions. We stratified SUB1 eCLIP targets into confidence intervals, based on the number, strength and confidence of the reported binding sites. Indeed, eCLIP targets of SUB1 (from HepG2 cells profiled by ENCODE) were significantly downregulated following *EPB41L4A-AS1* KD in MCF-7, with more confident targets experiencing stronger downregulation (**Fig. 3C**). To obtain SUB1-associated transcripts in MCF-7 cells; we performed a native RNA immunoprecipitation followed by sequencing of polyA+ RNAs (RIP-seq) (**Fig. 3D** and **S6C**). As expected, *EPB41L4A-AS1* precipitated with SUB1, as did other top hits from the eCLIP data (**Fig. S6C**). A total of 1,470 genes were enriched in the IP fraction (adjusted P<0.05 and $|\log_2 \text{Fold-change}| > 0.41$) (**Fig. S6D**), and the MCF-7 RIP data were in agreement with that of the eCLIP in HepG2 cells (**Fig. S6E**). Similarly to what was done with the eCLIP data, we divided the RIP-defined targets into confidence intervals based on the significance and overall enrichment in the IP fraction. Strikingly, SUB1-associated transcripts were downregulated upon *EPB41L4A-AS1* KD (**Fig. 3E**), similar to what was observed for the eCLIP-based targets. Since SUB1 is also a transcriptional regulator associated with chromatin, we performed CUT&RUN to uncover the genome-wide chromatin occupancy of SUB1 in MCF-7 cells (**Table S3**). SUB1 binding was enriched around TSSs (**Fig. 3F**) and among its 835 high-confidence genomic peaks, 418 (~50%) overlapped with the promoters of 411 genes. Of these genes, transcripts of only 56 genes were also bound by SUB1 at the RNA level, suggesting largely distinct sets of genes targeted by SUB1 at both the DNA and the RNA levels. Genes with promoter occupancy of SUB1 were enriched for biological processes related to genome maintenance, stress response, and nuclear transport (**Fig. S6F**), which is in line with previous studies[49–52]. These genes were also slightly yet highly significantly downregulated following *EPB41L4A-AS1* KD (**Fig. 3E**). Taken together, *EPB41L4A-AS1* interacts with SUB1 and its KD led to both a decrease in SUB1 mRNA level and the concomitant downregulation of SUB1 targets on chromatin and at the RNA level, positioning this lncRNA as a modulator of SUB1 activity.

Global de-regulation of snoRNAs and histone mRNA expression upon loss of *EPB41L4A-AS1*

Since many classes of noncoding RNAs—including those localizing to the nucleolus—are not polyadenylated, we next sequenced total rRNA-depleted RNAs from *EPB41L4A-AS1* KD and control MCF-7 cells. As expected, using this approach we could quantify various RNA species lacking polyA tails, such as small nuclear RNAs (snRNAs), snoRNAs and histone mRNAs, and could also better quantify intronic expression. Consistent with a *cis*-acting regulation at the transcriptional level, both the exonic and intronic reads of *EPB41L4A* were significantly reduced after *EPB41L4A-AS1* downregulation (**Fig. S7A**). When considering different transcript classes, we observed a marked upregulation of most snRNAs and snoRNAs and a significant downregulation of histone mRNAs (**Fig. 4A**). To quantity SNHG de-regulation at higher resolution, we created a custom annotation in which we divided each SNHG into exonic regions, the snoRNA-encoding region, regions in the snoRNA-containing introns that are found upstream

('pre-') and downstream ('post-') of the snoRNA, and other snoRNA-less introns ('Other') of the SNHG transcripts, while considering the various isoforms annotated for each SNHG (see Methods) (**Fig. 4B**). We then quantified the levels of each region in control and KD conditions. Intriguingly, while SNHG exons and unrelated introns were not affected by *EPB41L4A-AS1* KD, read coverage was increased over both the snoRNAs and the flanking intronic regions (**Fig. 4C**). These changes were shared among all the snoRNA classes and were concordant within individual introns when both (pre- and post-) had enough coverage to be quantified (**Fig. S7B** and **S7C**). Notably, the effect sizes of these changes were small, and they were not evident when plotting metagenes of these regions—that show the median coverage level—which is close to zero even in rRNA-depleted data (**Fig. S7D**), therefore suggesting that snoRNA processing remains an efficient process even in *EPB41L4A-AS1*-depleted cells, and most of the effect we observe is in increased accumulation of the mature and stable snoRNAs. This suggests that *EPB41L4A-AS1* perturbation specifically affects the snoRNA production and/or stability rather than SNHG transcription.

To test if the effects on snoRNA biogenesis are related to the changes in SUB1 expression and/or activity, we sequenced RNA from MCF-7 cells in which SUB1 was depleted by siRNAs (**Fig. S7E** and **S7F**). We identified a total of 243 upregulated and 154 downregulated genes (**Fig. S7G**), of which 41 (~17%) and 74 (~48%) were concordantly affected by *EPB41L4A-AS1* depletion, and the magnitudes of the genome-wide responses were moderately, but significantly correlated (Spearman's $R = 0.138$, $P<10^{-15}$). Strikingly, *SUB1* KD led to a similar increase in snoRNA levels and a decrease in histone mRNAs (**Fig. 4D, 4E** and **S7H**). Interestingly, the changes in snoRNAs expression were highly concordant with those driven by the KD of *EPB41L4A-AS1* (**Fig. 4F**, $R = 0.383$) and intronic regions flanking the snoRNAs were also increased, suggesting that *EPB41L4A-AS1* and SUB1 affect the abundance of the same pool of snoRNAs in MCF-7 cells through likely a shared mechanism.

Lastly, we tested whether the effects of *EPB41L4A-AS1* perturbation could be a generic outcome of a noncoding SNHG (ncSNHG) perturbation. To this end, we used GapmeRs to deplete *GAS5*, another abundant SNHG which encodes 11 snoRNAs within its introns (**Fig. S8A** and **S8B**). As expected from the various functions associated with this ncSNHG[56–64], *GAS5* KD led to substantial changes in gene expression (**Fig. S8C**), including a marked suppression of histone mRNA transcription, in line with its anti-proliferative function (**Fig. S8D**). However, snoRNAs or their host introns were largely unaffected (**Fig. S8E** and **S8G**), suggesting that the effect of *EPB41L4A-AS1* on snoRNA expression is not a common feature of ncSNHGs.

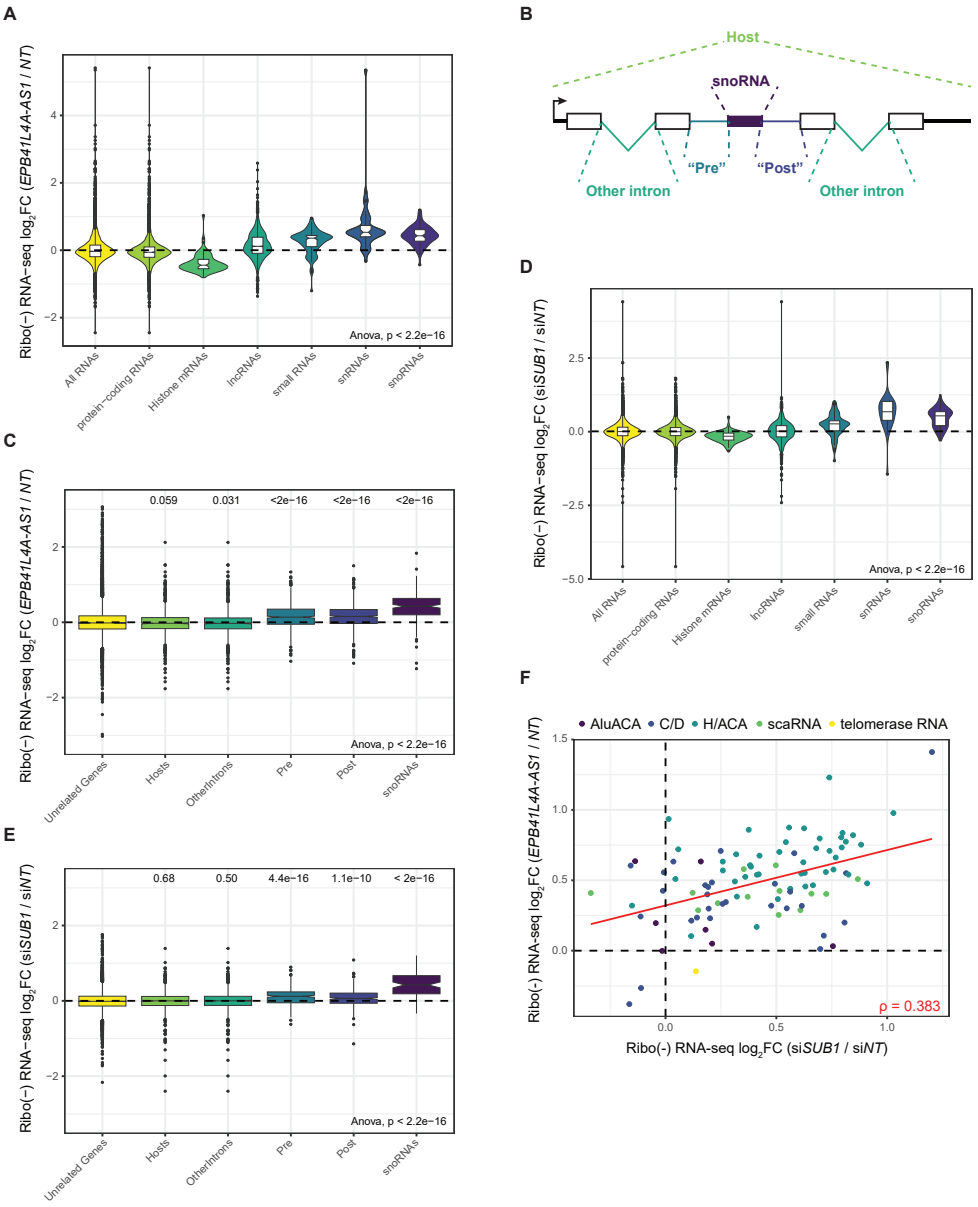


Figure 4. *EPB41L4A-AS1* and SUB1 depletion results in a widespread accumulation of mature snoRNAs. **(A)** Changes in gene expression upon *EPB41L4A-AS1* KD with GapmeRs of the indicated RNA classes. **(B)** Schematics depicting the different regions which were separately quantified in each SNHG. **(C)** Changes in RNA-seq read coverage in different regions upon *EPB41L4A-AS1* KD with GapmeRs. **(D)** As in (B) for *SUB1* KD with siRNAs of the indicated RNA classes. **(E)** as in (C) for *SUB1* KD. **(F)** Correspondence between changes in snoRNAs expression after *EPB41L4A-AS1* and *SUB1* KD. The color indicates the different snoRNA classes, and Spearman's correlation coefficient is shown. All experiments were performed in n = 3 biological replicates. In the boxplots, the thick line, edges of the box, and whiskers represent the median, first and third quartiles, and the upper and lower 1.5 interquartile ranges (IQRs), respectively. Outliers (observations outside the 1.5 IQRs) are drawn as single points, the significance of the different comparisons was computed by a Mann-Whitney test, and a global ANOVA P-value is also reported.

Depletion of *EPB41L4A-AS1* affects SUB1 distribution in the nucleus

The results described so far point to *EPB41L4A-AS1* being important for SUB1 function. To test whether *EPB41L4A-AS1* is important for its proper intracellular distribution, we immunostained for SUB1 and NPM1 in MCF-7 cells transfected with either control or *EPB41L4A-AS1*-targeting GapmeRs. A substantial change in subcellular localization could be observed for both proteins (**Fig. 5A** and **5B**). Specifically, SUB1 normally displays a pan-nuclear localization, sometimes slightly enriching into DAPI-low areas corresponding to the nucleoli (**Fig. 5A** and **5C**). Upon *EPB41L4A-AS1* depletion, SUB1 shifts towards a stronger nucleolar localization while retaining a presence throughout the nucleus. NPM1 instead regularly shows enrichment inside the nucleoli; however, in GapmeR-transfected cells, this is accompanied by the appearance of globular aggregates within the nucleus (**Fig. 5B** and **5D**), reminiscent of the previously reported nucleolar damage/stress response[65,66].

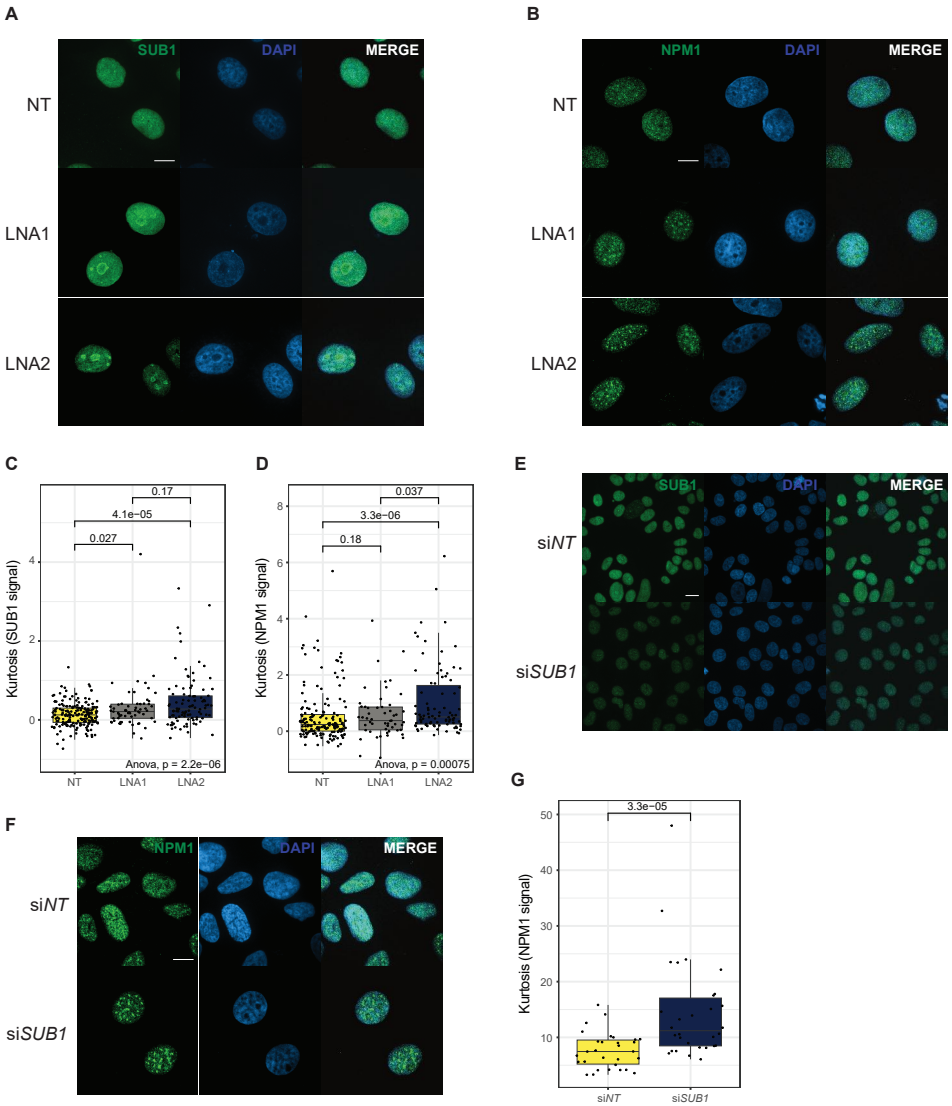


Figure 5. *EPB41L4A-AS1* and SUB1 depletion results in a widespread accumulation of mature snoRNAs. **(A)** Representative immunofluorescence images for SUB1 after *EPB41L4A-AS1* depletion with two distinct GapmeRs (scale bar = 20 µm, 100X magnification). **(B)** Same as in (A), but for NPM1. **(C)** Quantification of the kurtosis of SUB1 nuclear signal in the indicated conditions. **(D)** Same as in (C), but for NPM1. **(E)** Representative immunofluorescence images for SUB1 after *SUB1* KD with siRNAs (scale bar = 20 µm, 60X magnification). **(F)** Same as in (E), but for NPM1. **(G)** Same as in (D), but after *SUB1* depletion. All experiments were performed in n = 3 biological replicates. In the boxplots, the thick line, edges of the box, and whiskers represent the median, first and third quartiles, and the upper and lower 1.5 interquartile ranges (IQRs), respectively. Outliers (observations outside the 1.5 IQRs) are drawn as single points, the significance of the different comparisons was computed by a Mann-Whitney test, and a global ANOVA P-value is also reported. In each boxplot, points represent individual measurements (cell nuclei).

To test if the observed patterns of changes in SUB1 and NPM1 distribution are triggered by nucleolar stress, we exposed MCF-7 cells to CX-5461, an RNA polymerase I inhibitor that is currently being evaluated as a therapeutic agent for solid tumors with DNA repair deficiencies[67–70]. This led to a similar NPM1 pattern as that triggered by *EPB41L4A-AS1* depletion as early as 1h after treatment with the higher dose, and the change was more pronounced in later time points (**Fig. S9A** and **S9B**). SUB1 localization also changed, but only at the last time point (24hrs) (**Fig. S9C** and **S9D**). Lastly, we tested whether SUB1 depletion (**Fig. 5E**) could affect NPM1 distribution, and indeed SUB1 KD via siRNAs made NPM1 assume a similar pattern to that caused by *EPB41L4A-AS1* loss (**Fig. 5F** and **5G**), suggesting that SUB1 acts upstream of NPM1 and this lncRNA might modulate this regulation.

EPB41L4A-AS1 perturbations do not affect steady-state levels of *SNORA13*

A recent study reported an independent function for *SNORA13*, the snoRNA hosted by *EPB41L4A-AS1* in ribosomal assembly and prevention of nucleolar stress in BJ fibroblasts and mice[27]. We therefore wondered if the effects we observe in our experimental system result from changes in expression levels of this snoRNA. *SNORA13* levels did not change in RT-qPCR analysis of perturbed cells (**Fig. 2B**), or in rRNA-depleted RNA-seq data (**Fig. 6A**). We also performed a Northern blot for *SNORA13*, which also did not indicate changes in expression level or changes in RNA size of *SNORA13* upon GapmeR-mediated depletion of *EPB41L4A-AS1* (**Fig. 6B**). To complement these data, we assessed the levels of the 18S:1248$m^1acp^3\Psi$ modification, which includes a pseudouridylation guided by *SNORA13*. We used an aminocarboxyl propyl reverse transcription (aRT)-PCR assay as previously described[26], and found no difference in HinfI cleavage between control cells and those transfected with either *EPB41L4A-AS1*- or *EPB41L4A*-targeting GapmeRs (**Fig. 6C**), suggesting that overall the MACP modification is not affected as well. In addition, variant calling at 18S:1248 exploiting the residual rRNA reads in our polyA+ RNA-seq dataset revealed no distinction in allele frequency between control and transfected cells (**Fig. 6D**), concordant with the aRT-PCR results. In conclusion, we show that in our experimental system—MCF-7 cells transiently transfected with GapmeRs targeting its host—we do not observe changes in steady state *SNORA13* levels nor abundance of the only known RNA modification driven by this snoRNA. Therefore, we postulate that the observed endophenotypes are likely due to changes in the *EPB41L4A-AS1* lncRNA levels and its cooperation with SUB1.

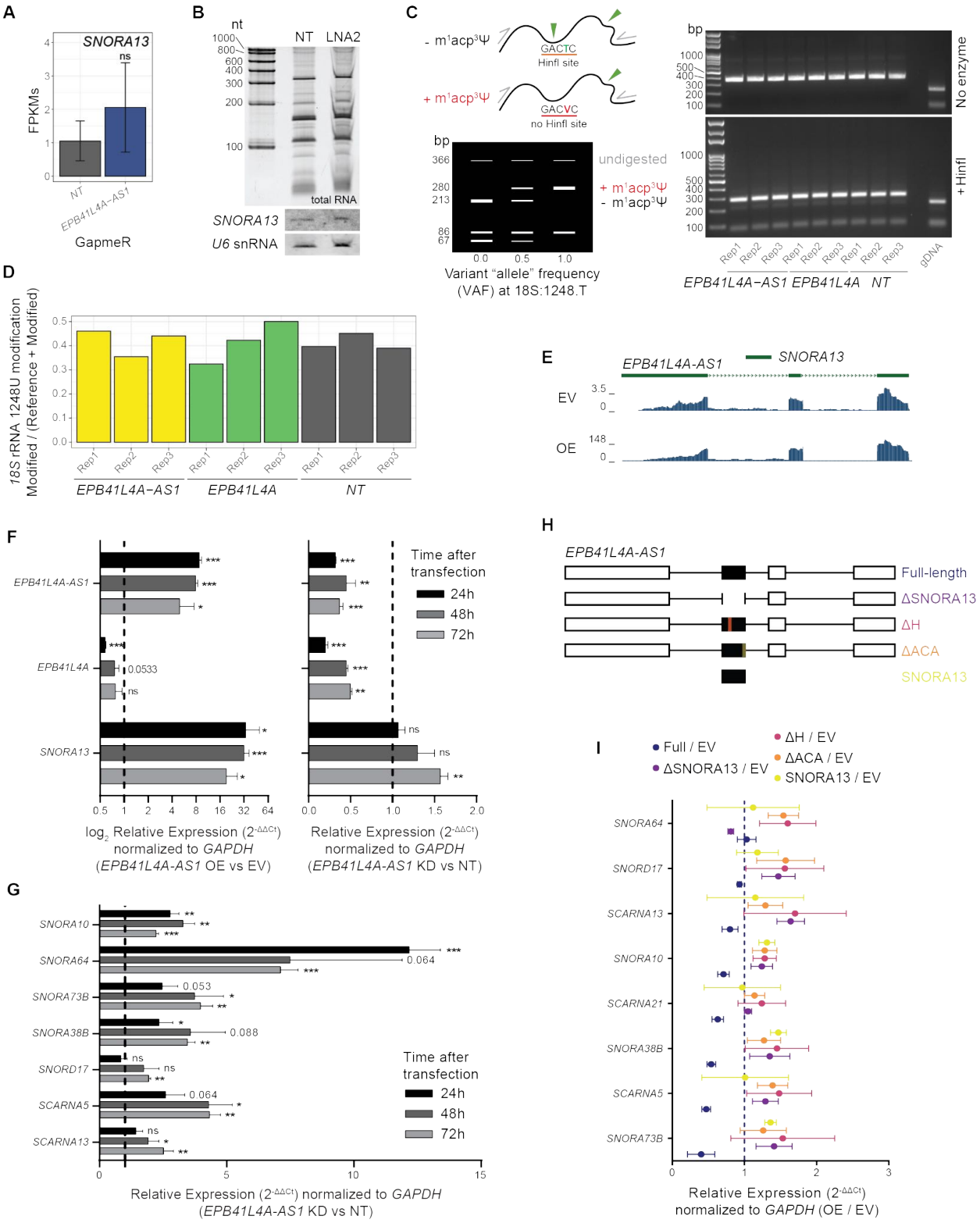


Figure 6. *EPB41L4A-AS1* but not the expression of *SNORA13* is required for proper snoRNAs expression. **(A)** Average expression levels (in FPKM) of *SNORA13* in cells transfected with GapmeRs targeting *EPB41L4A-AS1* (rRNA-depleted RNA-seq), with the error bars representing the standard deviation across the three replicates.

DESeq2 adjusted P-values compared to control GapmeR are also reported. **(B)** Northern blot for *SNORA13* in cells transfected with GapmeRs targeting *EPB41L4A-AS1*. *U6* was used as a loading control, and total RNA stain is also shown on top. **(C)** Schematics of the expected fragment sizes following HinfI digestion in the aRT-PCR assay (left), and agarose gel following HinfI digestion of the PCR-amplified *18S* rRNA in cells transfected with GapmeRs targeting either *EPB41L4A-AS1* or *EPB41L4A* (right). No enzyme and genomic DNA (gDNA) were used as negative controls. **(D)** Allele frequency at 18S:1248 in the polyA+ RNA-seq dataset, using rRNA reads. **(E)** UCSC Genome Browser view of the *EPB41L4A-AS1* locus, with tracks showing the rRNA-depleted RNA-seq coverage of *EPB41L4A-AS1* OE and control cells. **(F)** RT-qPCR for the indicated genes upon *EPB41L4A-AS1* KD with GapmeRs (right) or overexpression with a full-length unspliced vector (left) over the course of three days post transfection. **(G)** RT-qPCR for the indicated snoRNAs upon *EPB41L4A-AS1* KD with GapmeRs over the course of three days post transfection. **(H)** Schematics of the different *EPB41L4A-AS1* and *SNORA13* overexpressing vectors used in the rescue experiments. **(I)** The ratio between changes in gene expression detected by RT-qPCR in cells transfected with an *EPB41L4A-AS1* or *SNORA13* overexpressing vector vs control cells. All experiments were performed in n = 3 biological replicates, with the error bars in the barplots and forest plots representing the standard deviation. In all cases, ns - P>0.05; * - P<0.05; ** - P<0.01; *** - P<0.001 (two-sided Student's t-test).

A full-length, intron- and *SNORA13*-containing *EPB41L4A-AS1* is required to partially rescue snoRNA expression

We reasoned that if *EPB41L4A-AS1* indeed regulates snoRNA expression, this function likely occurs in *trans*, because none of the genes in the vicinity of the *EPB41L4A-AS1* locus are known to be involved in snoRNA biogenesis or stability. To test this, we performed rescue experiments in which we co-transfected lncRNA-targeting GapmeRs with different GapmeR-resistant *EPB41L4A-AS1* variants expressed from a plasmid driven by a constitutive CMV promoter. We confirmed that the transfected vector is well spliced (**Fig. 6E**), and results in an increase in both *EPB41L4A-AS1* and *SNORA13* expression (**Fig. 6F**, left).

We then assessed both KD dynamics and the abundance of snoRNAs selected based on their sensitivity to *EPB41L4A-AS1* KD in the RNA-seq data at 24, 48 and 72 hours after transfection using qRT-PCR. *EPB41L4A* expression decreased concomitantly with that of *EPB41L4A-AS1*, in line with the transcription of the lncRNA regulating this gene in *cis* (**Fig. 6F**, right). The expression of most *EPB41L4A-AS1*-sensitive snoRNAs increased as early as 24 hours after transfection, and they generally remained upregulated throughout the next two days (**Fig. 6G**). We designed four different *EPB41L4A-AS1* expression vectors to attempt a rescue: 1) full length, including introns; 2) full length, including introns but without *SNORA13* (Δ*SNORA13*); 3) full length, including introns but with either the H or ACA *SNORA13* boxes deleted (ΔH and ΔACA); and 4) *SNORA13* alone (**Fig. 6H**). We then compared gene expression between cells transfected with the overexpressing vectors (OE) or an empty control (EV). We were able to obtain a partial rescue of snoRNA expression only when co-transfecting the full length *EPB41L4A-AS1*, and not any other vector not expressing any or just a mutated *SNORA13* (**Fig. 6I**), indicating that the *SNORA13* sequence is involved in the process. However, the expression of a stand-alone *SNORA13,* using the expression vector used and validated previously[27], was incapable to drive any rescue of the snoRNAs abundance (**Fig. 6I**). This would suggest that it is the whole *EPB41L4A-AS1* transcript and its snoRNA-containing processing that can counter *EPB41L4A-AS1* KD.

EPB41L4A-AS1 affects RNA metabolism of both the genes found in *cis* and of snoRNAs

To distinguish between the effects of *EPB41L4A-AS1* depletion on RNA synthesis and RNA decay, we performed SLAM-seq[71] using ribo-depleted total RNA in *EPB41L4A-AS1* KD cells (**Table S4**). After 48 hours of control or *EPB41L4A-AS1*–targeting GapmeR transfection, we collected RNA at 0, 4, 8 and 24 hours post-addition of 4sU to the media, while replenishing fresh 4sU at each time point (see Methods) (**Fig. 7A**). *EPB41L4A* showed decreased synthesis rate and unchanged half-life (**Fig. 7B**), as expected from a gene regulated at the level of transcription. As a group, genes upregulated after *EPB41L4A-AS1* KD in the regular steady-state RNA-seq data had substantially higher synthesis rates, and those downregulated had lower synthesis rates, when compared to unchanged ones ($P=3.25x10^{-14}$ for the up-regulated genes and P=0.002 for the down-regulated genes, Mann-Whitney test, **Fig. 7C**), whereas only a minor difference was observed in terms of half-lives (P=0.011 for the up-regulated genes and $P=1.82x10^{-8}$ for the down-regulated genes, Mann-Whitney test, **Fig. 7D**). This suggests that the *trans*-regulated genes are also mostly regulated at the level of their transcription. When considering the genes in the two neighboring TADs of *EPB41L4A-AS1*, we also observed a general reduction in both synthesis rates (**Fig. S10A**, left) and half-lives (**Fig. S10B**, left). When comparing the effects across genes in the two neighboring TADs ("Cis"), other TADs on the same chromosome ("Chr5"), or other chromosomes ("Trans"), the genes proximal to *EPB41L4A-AS1* specifically exhibit a reduced synthesis rate compared to the others (**Fig. S10A-B**).

Compared to their hosts and other genes, snoRNAs were, as expected, relatively stable in both KD and control conditions (**Fig. 7E** and **7F**), while the intronic regions exhibited high turnover. When comparing *EPB41L4A-AS1* KD to the control GapmeRs, a slight increase in the synthesis rates of different parts of SNHGs was observed. The snoRNA half-life is difficult to measure accurately due to their relative stability and their short length, which means that the number of incorporated 4sU that can be detected is much lower than that in longer transcripts. With these caveats in mind, there appeared to be a relative increase in snoRNA stability compared to a reduction observed in other genes and other parts of the SNHGs (**Fig. 7G** and **7H**). Altogether, this indicates that the snoRNA level increase driven by *EPB41L4A-AS1* knockdown is primarily due to the increased transcription of their hosts, with a possible additional component of relative stabilization.

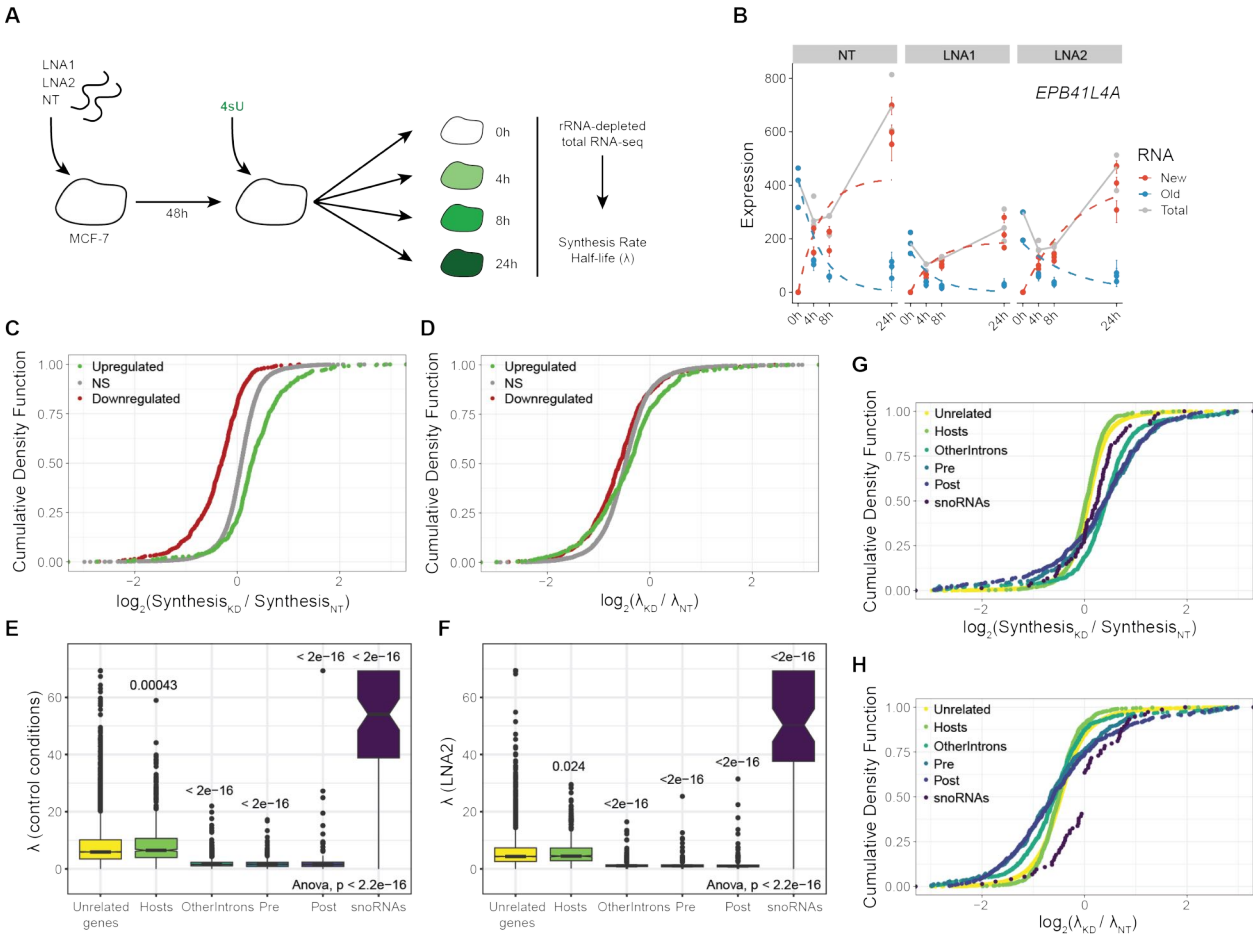


Figure 7. The increased abundance of snoRNAs is primarily due to their hosts' increased transcription and stability. **(A)** Workflow of the SLAM-seq experiment. MCF-7 cells were transfected with the indicated GapmeRs for 48 hours, after which the media was replaced with media containing 4sU and the cells were harvested at different time points. **(B)** Fitted model depicting the synthesis and decay rates of the *EPB41L4A* mRNA upon *EPB41L4A-AS1* KD with GapmeRs. **(C)** Changes in synthesis rate upon *EPB41L4A-AS1* KD with GapmeRs for the indicated group of genes. **(D)** Same as in (C), but for changes in half lives. **(E)** Half lives in control conditions of the indicated SNHG regions as described in Fig. 4B. **(F)** Same as in (E), but after *EPB41L4A-AS1* KD with LNA2. **(G)** Changes in synthesis rate upon *EPB41L4A-AS1* KD with GapmeRs for the indicated SNHG regions. **(H)** Same as in (G), but for changes in half lives. All experiments were performed in n = 3 biological replicates. In the boxplots, the thick line, edges of the box, and whiskers represent the median, first and third quartiles, and the upper and lower 1.5 interquartile ranges (IQRs), respectively. Outliers (observations outside the 1.5 IQRs) are drawn as single points, the significance of the different comparisons was computed by a Mann-Whitney test, and a global ANOVA P-value is also reported. In all cases, ns - P>0.05; * - P<0.05; ** - P<0.01; *** - P<0.001.

EPB41L4A-AS1 loss affects cell proliferation and invasion

Lastly, we aimed to assess the phenotypic consequences of *EPB41L4A-AS1* loss. As mentioned above, the genes affected by the KD were found to be enriched for GO terms related to cell adhesion, movement and division (**Fig. 2G**). Thus, we measured cell migration and proliferation by performing wound healing and cell counting assays in cells with reduced *EPB41L4A-AS1* levels. Cells transfected with *EPB41L4A-AS1* targeting GapmeRs exhibited significantly higher

migration rates, compared to controls (**Fig. 8A** and **8B**), with cells bordering the wounded region displaying a flattened and translucent morphology (**Fig. 8A**). This higher migration is not due to higher cellular proliferation/division, as cells with reduced lncRNA levels proliferate less (**Fig. 8C**), and this deficiency could be rescued by co-transfecting a vector encoding the GapmeR-resistant, full-length *EPB41L4A-AS1* (**Fig. 8C**), suggesting that this modulation takes place in *trans*. Taken together, these findings suggest that proper *EPB41L4A-AS1* expression is required for cellular proliferation, whereas its deficiency results in the onset of more aggressive and migratory behavior, likely linked to the increase of the gene signature of epithelial to mesenchymal transition (EMT) (**Fig. 8D**).

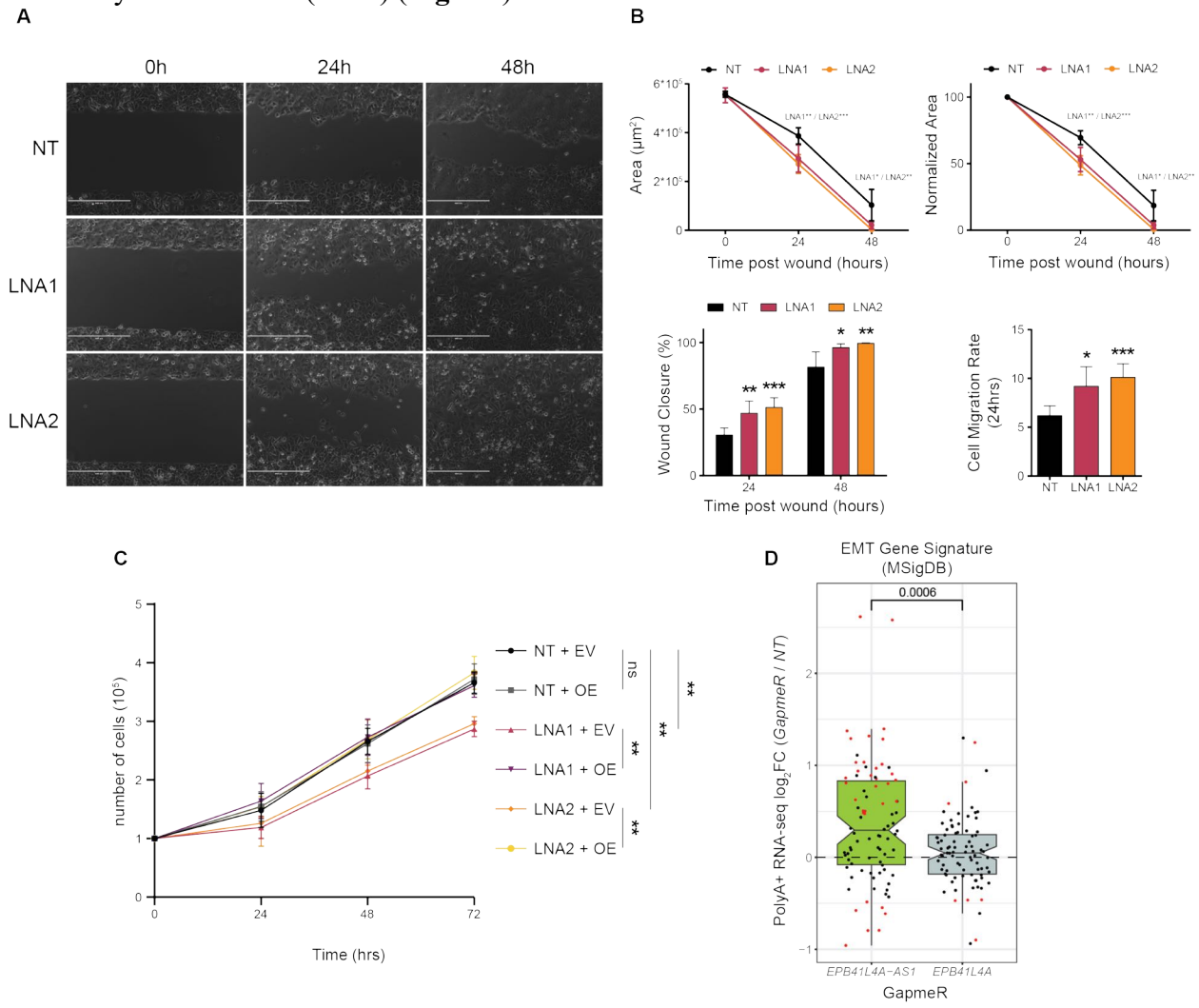


Figure 8. Cells with reduced *EPB41L4A-AS1* expression display reduced proliferation and increased invasion capacity. **(A)** Representative brightfield images of the wound at the indicated time points and conditions. **(B)** Wound area (top-left), normalized wound area (top-right), closure percentage (bottom-left) and migration rate (bottom-right) at the indicated time points and conditions. The significance was calculated by a two-sided Student's t-test. **(C)** Growth curve of the cells in the indicated conditions over the course of three days. The significance was calculated by a two-sided Student's t-test. **(D)** Changes in gene expression of the EMT signature genes (from MSigDB) after KD with GapmeRs targeting either *EPB41L4A-AS1* or *EPB41L4A* (polyA+ RNA-seq data). The significance of the comparison was computed by a Mann-Whitney test. All experiments were performed in n = 3 biological replicates. The error bars in the barplots represent the standard deviation. In the

boxplot, the thick line, edges of the box, and whiskers represent the median, first and third quartiles, and the upper and lower 1.5 interquartile ranges (IQRs), respectively. Outliers (observations outside the 1.5 IQRs) are drawn as single points. The points in the boxplot represent individual genes, and their color whether they were found to be significantly (adjusted P<0.05 and |$\log_2$Fold-change| >0.41) dysregulated (red) or not (black). In all cases, ns - P>0.05 = ns; * - P<0.05 = *; ** - P<0.01 = **; *** - P<0.001 = *** (two-sided Student's t-test).

Discussion

We show here that the *EPB41L4A-AS1* gene locus harbors a pleiotropic gene that, within a single biological system—MCF-7 breast cancer cells—performs diverse functions. At the site of its own transcription, which overlaps a strong TAD boundary, *EPB41L4A-AS1* is required to maintain expression of several adjacent genes, regulated at the level of transcription. Strikingly, the promoter of *EPB41L4A-AS1* ranks in the 99.8th percentile of the strongest TAD boundaries in human H1 embryonic stem cells[43,72]. It features several CTCF binding sites (**Fig. 2A**), and in MCF-7 cells, we demonstrate that it blocks the propagation of the 4C signal between the two flanking TADSs (**Fig. 1F**). Future studies will help elucidate how *EPB41L4A-AS1* transcription and/or the RNA product regulate this boundary. So far, we found that *EPB41L4A-AS1* did not affect CTCF binding to the boundary, and while some peaks in the vicinity of *EPB41L4A-AS1* were significantly affected by its loss, they did not appear to be found near genes that were dysregulated by its KD (**Fig. S10C**). We also found that KD of *EPB41L4A-AS1*—which depletes the RNA product, but may also affect the nascent RNA transcription[73,74]—reduces the spatial contacts between the TAD boundary and the *EPB41L4A* promoter (**Fig. 2E**). Further elucidation of the exact functional entity needed for the *cis*-acting regulation will require detailed genetic perturbations of the locus, that are difficult to carry out in the polypoid MCF-7 cells, without affecting other functional elements of this locus or cell survival as we were unable to generate deletion clones despite several attempts.

Beyond its site of transcription, *EPB41L4A-AS1* associates with SUB1, an abundant protein linked to various functions, and these two players are required for proper distribution of various nuclear proteins. Their dysregulation results in large-scale changes in gene expression, including up-regulation of snoRNA expression, mostly through increased transcription of their hosts, and possibly through a somewhat impaired snoRNA processing and/or stability. Several features of the response to *EPB41L4A-AS1* resemble nucleolar stress, including altered distribution of NPM1[65,66]. SUB1 was shown to be involved in many nuclear processes, including transcription[50], DNA damage response[75,76], telomere maintenance[53], and nucleolar processes including rRNA biogenesis[77]. Our results suggest a complex and multi-faceted relationship between *EPB41L4A-AS1* and SUB1, as SUB1 mRNA levels are reduced by the lncRNA KD (**Fig. 3B**), the distribution of the protein in the nucleus is altered (**Fig. 5A** and **5C**), while the protein is the most prominent binder of the mature *EPB41L4A-AS1* in ENCODE eCLIP data (**Fig. 3A**). The most prominent connection between *EPB41L4A-AS1* and SUB1 is the similar phenotype triggered by their loss (**Fig. 4**). It is difficult to determine which of the connections between these two genes is the most functionally relevant and which may be indirect and/or feedback interactions. For example, it is possible that *EPB41L4A-AS1* primarily acts as a transcriptional regulator of *SUB1* mRNA or that its RNA product is required for proper stability and/or localization of the SUB1 protein or that *EPB41L4A-AS1* acts as a scaffold for the formation of protein-protein interactions of SUB1.

In the context of cancer, a systematic study of TCGA data, which is consistent with our analysis (**Fig. S2**) found that *EPB41L4A-AS1* is downregulated in tumors compared to tumor-adjacent normal tissues[29], in accordance with both the downregulation of the TIGA1 protein[29] and the reduction in the *SNORA13*-guided rRNA modification[26]. The association with poor prognosis[29] is consistent with the observed increased invasiveness that we find in MCF-7 cells. Effects of *EPB41L4A-AS1* on migration were also reported in colorectal cancer cells[32].

Several recent studies have described the functions of *EPB41L4A-AS1* in various biological systems, highlighting its dysregulation in cancer[29,32,78], diabetes[31], aging[79,80], and Alzheimer's disease (AD)[80]. Both p53 and PGC-1α were shown to act upstream of *EPB41L4A-AS1* and regulate its expression[29]. In these different systems, different pathways were described as regulated by this lncRNA, including glycolysis, glutamine metabolism[29], and autophagy[80]. Mechanistically, it was reported that this lncRNA regulates gene expression via its association with and upregulation of KAT2A (GCN5/GNC5L2)[31,80]. These prior studies generally relied on candidate-gene or co-expression approaches to identify the genes regulated by *EPB41L4A-AS1*, and did not systematically examine its *cis*- and *trans*-acting effects, with the exception of a *cis*-acting regulation, that was briefly hinted in a preprint posted four years ago[81], but not published so far to the best of our knowledge. These studies also did not systematically investigate the *trans*-regulated targets of *EPB41L4A-AS1*, did not separately deplete the nascent and the mature forms of the lncRNA, and did not attempt to decouple its functions from those of *SNORA13*.

A recent study found that *SNORA13* negatively regulates ribosome biogenesis in immortalized human fibroblasts, by decreasing the incorporation of RPL23 into the maturing 60S ribosomal subunits, eventually triggering p53-mediated cellular senescence[27]. In their experimental setting, *SNORA13* was constitutively reduced by either CRISPRi of the *EPB41L4A-AS1* locus for 7 days or its KO leaving the lncRNA sequence expression unaltered, whereas our approach relies solely on transient knockdown for 48–72hrs that does not appreciably affect *SNORA13* (**Fig. 1B, 2B, 6A** and **6B**), or the levels of the RNA modification that it guides (**Fig. 6C** and **6D**). Therefore, while both studies implicate *EPB41L4A-AS1* in nucleolar biology and its perturbation in nucleolar stress, there is an apparent discrepancy between our results on the role that *SNORA13* plays. In our hands, the differential expression of snoRNAs and cell proliferation can be rescued by expression of *EPB41L4A-AS1* with the snoRNA, and not by separate expression of *SNORA13*, suggesting the snoRNA is not playing a central role in the observed phenotypes. Also, the study by Cheng et al. showed that activity of the p53 pathway was less active in *SNORA13* knockout cells, with or without induction of oncogenic stress, whereas we see an opposite trend of some up-regulation of the p53 pathway upon transient *EPB41L4A-AS1* KD (**Fig. S10D**). We cannot however completely rule out potential contributions of this snoRNA to the phenotypes observed in this study. For instance, the phenotype of snoRNAs expression is partially rescued by *EPB41L4A-AS1* overexpression in *trans* only when a full-length, unspliced and *SNORA13*-containing vector is used, and not when the snoRNA is deleted, mutated or expressed alone (**Fig. 6I**). This would suggest that the processing of *EPB41L4A-AS1* is involved in its mode of action, although further experiments will be needed to address whether *SNORA13* or its DNA sequence are indeed required to regulate the expression of the snoRNAs. We note that compensatory effects are possible in cells that grow for extended periods of time without *SNORA13*.

Methods

<u>Cell culture, transfection and CX-5461 treatment</u>

MCF-7 cells were cultured in DMEM supplemented with 4.5 g/L of glucose, 4 mM L-Glutamine, 10% FBS, 1% PenStrep solution, grown in incubators maintained at 37°C with 5% CO_2, and were detached for passaging using a Trypsin (0.05%) – EDTA (0.02%) solution. DNA plasmid, siRNAs (final concentration 25 µM) and GapmeRs (50 µM) transfections were performed by using Lipofectamine 3000 (ThermoFisher Scientific), according to the manufacturer's protocol protocol. In all cases, cells were plated the day before the transfection in an antibiotic-free medium which was subsequently changed prior to adding the transfection mix, and again after 24 hr to maximize cell viability. The complete list with the sequences of both siRNAs and LNA/GapmeRs is reported in **Table S5**.

CX-5461 (Sigma) was diluted in sterile-filtered 50 mM NaH_2PO_4 at a stock concentration of 10mM and added to the cell culture media at a final concentration of 0.5, 1 or 5 µM. As a control, cells were supplemented with an equal volume of 50 mM NaH_2PO_4.

Plasmids construction

The sequence corresponding to the cDNA of human *EPB41L4A-AS1* (ENST00000688370.2_3) was synthetized and cloned by Hylabs inside a pcDNA3.1(+) vector. The vector to express the *EPB41L4A* ORF (NM_022140.5) was purchased from Genecopoeia (EX-A8889-Lv201). The vector expressing a full length and unspliced *EPB41L4A-AS1* (pLV[ncRNA]-Puro-CMV) was purchased from Vectorbuilder, and the target site of LNA2 was mutated by introducing four mismatches (AACTTAAAAGCAGCGT → ACCTTAGAAGTAGAGT) via restriction-free cloning. To generate the Δ*SNORA13*, ΔH, and ΔACA vectors, we ordered *EPB41L4A-AS1* gene fragments with the desired mutations from TWIST. These fragments corresponded to the sequence between the unique SalI and MluI restriction sites and were inserted via restriction digestion. The pLKO.1 *SNORA13* plasmid[27] was a gift from Prof. Mendell (UT Southwestern). All gRNAs were ordered as single-stranded oligos (Sigma) and cloned inside pKLV-Puro vectors as previously described[82]. The complete list of gRNAs and their sequences is available in **Table S5.** All plasmids were sent for Sanger and whole plasmid (Plasmidsaurus) sequencing prior to use.

Wound healing assay and analysis

Wound healing experiments were carried out as previously described[83]. We employed the µ-Plate 24 Well system (Ibidi). We transfected the targeting and control GapmeRs via reverse transfection, diluting the transfection mixes directly into the cell mixtures while seeding the cells. After reaching confluence overnight, the culture inserts were gently removed with sterile forceps and the wells were washed with 1X PBS. Brightfield images were taken using an EVOS Cell Imaging System (Thermo Fisher Scientific) at different time points. Data analysis was then performed using the Wound_healing_size_tool[84] plugin for ImageJ, which calculates the total (μm^2) and percentage (%) of the wounded area. From these data, we calculated the rate of cell migration (R_M) as:

$$R_M = \frac{W_i - W_t}{t}$$

Where W_i represents the initial average wound width, and W_t is the average wound width at a given time t. Additionally, we derived the percentage of wound closure defined as:

$$Wound\ Closure\ (\%) = \frac{A_{t=0} - A_{t=\Delta t}}{A_{t=0}}$$

With $A_{t=0}$ and $A_{t=\Delta t}$ being the wounded area (in μm^2) at the beginning and after time t, respectively.

Gene expression analysis

RNA was extracted by using TRI-Reagent as previously described[83], according to the manufacturer's instructions. The isolated RNA was then resuspended in DNase/RNase-free H_2O and DNase-treated using the Baseline-ZERO™ DNase (Biosearch Technologies), following the manufacturer's recommendations and incubating the RNAs for 1 hour at 37°C. RNA was then retrotranscribed using the qScript™ Flex cDNA Synthesis Kit (Quanta Bio), following standard conditions as suggested by the manufacturer. The resulting cDNA was finally diluted 1:5-1:10 with DNase/RNase-free H_2O and used as input for Quantitative Real-Time PCR using the Fast SYBR™ Green Master Mix (Applied Biosystems). Each reaction was amplified using the following PCR program: 20'' at 95°C, 40 cycles by denaturing for 1" at 95°C followed by annealing/extension for 20'' at 60°C. Results were analyzed using the ΔΔCt method. The complete list of primers and their sequences is available in **Table S5**.

Aminocarboxyl propyl reverse transcription (aRT)-PCR assay to estimate 18S:1248m^{1}acp^{3}Ψ

The aminocarboxylpropyl reverse transcription (aRT)-PCR assay to detect and quantify the 1248.m1acp3Ψ modification was performed as previously described[26], with minor modifications. Briefly, total RNA was extracted as described above, and 1 µg per sample was DNase-treated using TURBO DNase (Invitrogen), according to the manufacturer's protocol. Retrotrascription was then carried out using SuperScript III (Invitrogen), following the manual's recommendations and by using only random hexamers. The cDNA was then diluted 1:5 with DNase/RNase-free H_2O and used for end-point PCR using the Q5 High-Fidelity DNA Polymerase (NEB), according to the standard protocol and using the primers listed in the original protocol[26] and available in **Table S5**. From each PCR reaction, 5 µL were digested in duplicate with HinFI (NEB) in rCutSmart buffer (NEB) and an additional 5 µL were incubated without the enzyme as a negative control. Lastly, the resulting fragments were then run on a 2% agarose gel.

Northern Blot

Northern Blot to detect *SNORA13* was performed as previously described[27], with minor modifications. A total of 20 µg of RNA per sample was loaded onto an 8% TBE-UREA polyacrylamide gel, which had been pre-equilibrated by running at low voltage in 1X TBE. After electrophoresis, the gel was stained for 15 minutes in 0.5X TBE with a few drops of ethidium bromide while rocking, to assess the quality of the run. The RNA was then transferred to an HyBond-NX membrane (GE Healthcare) in 0.5X TBE at room temperature using a constant current of 0.25 A for 2 hours. Following the transfer, the membrane was rinsed in 2X SSC,

crosslinked at 1200 J/cm² in a UV oven, and pre-hybridized for 30 minutes at 42°C in 10 mL of pre-warmed ULTRAhyb™ Ultrasensitive Hybridization Buffer while rotating. IR fluorescent probes targeting *U6* and *SNORA13* were then added to a final concentration of 10 nM, and the membrane was incubated O/N at 42°C while rotating. The next day, the membrane was washed twice in 2X SSC, 0.1% SDS at 42°C for 5 minutes while rotating, followed by two additional washes in 0.1X SSC, 0.1% SDS, and then finally imaged. The complete list of probes is available in **Table S5**.

Protein extraction and Western Blot

Proteins were extracted using RIPA as previously described[83]. Briefly, cells were detached using a 0.05% Trypsin – 0.02% EDTA solution, pelleted by centrifugation, resuspended in ice-cold RIPA buffer, and incubated on ice for 20 minutes with intermittent vortexing. The lysates were then centrifuged at 15,000 rpm for 15 minutes at 4°C, after which the supernatants were collected and quantified. For Western blots, equal amounts of protein were loaded onto a polyacrylamide gel and run in an SDS running buffer. Proteins were then transferred onto a pre-activated PVDF membrane in a cold Tris-Glycine buffer supplemented with 20% methanol, with a constant current of 0.30 A for 2 hours. Membranes were then blocked in 5% milk in 1X PBS-0.1% Tween 20 for 1 hour, followed by overnight incubation at 4°C with primary antibodies diluted in the same blocking solution. The next day, membranes were washed three times with 1X PBS-0.1% Tween 20, incubated for 2 hours at room temperature with the appropriate secondary antibodies, and washed three more times before image acquisition. The complete list of primary and secondary antibodies and relative dilutions is available in **Table S5**.

Immunostaining

Immunofluorescence was performed as previously described[83]. Cells were plated and transfected on either sterile glass coverslips or 8-well chambers (Ibidi). They were first washed once with 1X PBS, then fixed with cold 4% PFA in 1X PBS for 15 minutes at room temperature. After two 5-minute washes with 1X PBS, cells were permeabilized by incubating them in a permeabilization/blocking solution containing 5% horse serum, 1 mg/mL BSA, and 0.1% Triton X-100 in 1X PBS for 30 minutes at room temperature. Following three washes with 1X PBS, cells were incubated overnight at 4°C with primary antibodies diluted in blocking solution (5% horse serum, 1 mg/mL BSA in 1X PBS). The next day, cells were washed three times with 1X PBS and incubated for 2 hours at room temperature with secondary antibodies in blocking solution, protected from light. After three additional washes with 1X PBS, cells were stained with DAPI and mounted using ProLong™ Glass Antifade Mountant (Invitrogen). The slides were left to cure overnight in the dark at room temperature before imaging with a Zeiss Spinning Disk confocal microscope. The complete list of primary and secondary antibodies and relative dilutions is available in **Table S5**.

Subcellular fractionation

To fractionate cells and isolate the cytoplasmic, nucleoplasmic and chromatin RNA fractions we

employed a published protocol[85] with minor modifications. Confluent 10 cm plates were harvested with a Trypsin (0.05%) – EDTA (0.02%) solution, pelleted by centrifugation, washed once with ice-cold 1X PBS and resuspended in 380 µL of ice-cold HLB buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.3% NP-40, 10% Glycerol) freshly supplemented with RNase Inhibitors (EURX). The samples were then incubated on ice for 20' with occasional vortexing, spun down for 3' at 4°C / 1000 g, the supernatant was then transferred to a new tube, centrifuged again for 10' at 4°C / 1000 g, after which the supernatant (cytoplasmic fraction) was moved to a new tube, supplemented with 1 mL of RNA precipitation solution (RPS, 150mM Sodium Acetate in 100% ethanol) and stored at -20°C for at least 1 h. The semi-pure nuclei were then washed three times with ice-cold HLB and centrifugations for 2' at 4°C / 500 g, after which the pellets were resuspended in 380 µL of ice-cold MWS (10 mM Tris-HCl, pH 7.0, 4 mM EDTA, pH 8.0, 0.3 M NaCl, 1 M Urea, 1% NP-40) freshly supplemented with RNase Inhibitors, quickly vortexed and incubated on ice for 5 minutes, vortexed again and incubated for an additional 10 minutes. The tubes were then centrifuged for 3' at 4°C / 1000 g, the supernatant (nucleoplasmic / nuclear soluble fraction) was saved to a new tube, supplemented with 1 mL of RPS and stored at -20°C for at least 1 h. The chromatin pellets were finally washed three times with ice-cold MWS and centrifugations for 2' at 4°C / 500 g, after which 1 mL of TRI-Reagent was added and samples were stored at -20°C. After 1+ hours at -20°C, the tubes with the cytoplasmic and nucleoplasmic fractions in RPS were briefly vortexed, centrifuged for 15' at 4°C / 15000 g, the pellets were washed with 70% ethanol by vortexing and centrifuged again for 15' at 4°C / 15000 g, after which they were air-dried and resuspended in 1 mL of TRI-Reagent. To each sample in 1 mL of TRI-Reagent, 10 µL of 0.5 M EDTA, pH 8.0 were added followed by an incubation with vortexing for ~10 minutes at 65°C, or until the RNA pellets fully dissolved. The RNA was then extracted following the standard TRI-Reagent protocol.

Native RNA Immunoprecipitation (RIP)

To perform Native RNA Immunoprecipitation, we have employed a slightly modified version of the ENCODE IP protocol[86]. Briefly, $5x10^6$ MCF-7 cells were harvested by trypsinization, washed once with 1X PBS, and flash frozen as pellets in 50 µL of 1X PBS. The pellets were then thawed on ice, resuspended in 450 µL of ice-cold lysis buffer (50 mM Tris-HCl, pH 7.4; 100 mM NaCl; 1% NP-40; 0.1% SDS; 0.5% sodium deoxycholate) supplemented with EDTA-free Protease Inhibitor Cocktail (APExBIO)) and RNase Inhibitor (EURX), and incubated on ice for 20' with occasional vortexing to allow complete lysis. The lysates were then centrifuged at max speed for 15' at 4°C and the supernatants were collected into a new tube. A total of 20 µL/sample of Protein A/G were then washed twice with ice-cold lysis buffer and resuspended in 250 µL of lysis buffer, of which 50 µL were used to pre-clear the lysate by incubation at RT for 30' with gentle rotation, and the remaining 200 µL were instead incubated for 1 h/RT with 3 µg of antibody. After 30' of preclearing, the clear lysate was moved to a new tube in ice, and 25 µL (5%) of it was saved as INPUT sample. After 1 h, the antibody/beads were washed three times with ice-cold lysis buffer and incubated with the pre-cleared lysate O/N at 4°C while gently rotating. The following day, 25 µL (5%) of the supernatant was saved as UNBOUND sample, and the beads (BOUND sample) were gently washed twice with ice-cold wash buffer (20 mM Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20) freshly supplemented with RNase Inhibitors (EUR_X). Finally, 500 µL of TRI-Reagent was added to all three fractions and RNA isolation was carried out as previously described. The complete list of antibodies and relative dilutions is available in **Table S5**.

Single-molecule RNA FISH with HCR technology (smFISH HCR)

To visualize the subcellular localization of EPB41L4A-AS1 in vivo, we performed single-molecule fluorescence in situ hybridization (smFISH) using HCR™ amplifiers. Probe sets (n = 30 unique probes) targeting *EPB41L4A-AS1* and *GAPDH* (positive control) were designed and ordered from Molecular Instruments. We followed the Multiplexed HCR v3.0 protocol with minor modifications. MCF-7 cells were plated in 8-well chambers (Ibidi) and cultured O/N as described above. The next day, cells were fixed with cold 4% paraformaldehyde (PFA) in 1X PBS for 10 minutes at RT and then permeabilized O/N in 70% ethanol at -20°C. Following permeabilization, cells were washed twice with 2X SSC and incubated at 37°C for 30 minutes in hybridization buffer (HB). The HB was then replaced with a probe solution containing 1.2 pmol of *EPB41L4A-AS1* probes and 0.6 pmol of *GAPDH* probes in HB. The slides were incubated O/N at 37°C. To remove excess probes, the slides were washed four times with probe wash buffer at 37°C for 5 minutes each, followed by two washes with 5X SSCT at RT for 5 minutes. The samples were then pre-amplified in amplification buffer for 30 minutes at RT and subsequently incubated O/N in the dark at RT in amplification buffer supplemented with 18 pmol of the appropriate hairpins. Finally, excess hairpins were removed by washing the slides five times in 5X SSCT at RT. The slides were mounted with ProLong™ Glass Antifade Mountant (Invitrogen), cured O/N in the dark at RT, and imaged using a Nikon CSU-W1 spinning disk confocal microscope.

Cleavage Under Targets and Release Using Nuclease (CUT&RUN)

CUT&RUN reactions were performed following the V3 of the protocol[87], with minor modifications which were described previously[83]. An equal number of cells was harvested and washed once with 1X PBS at room temperature. After three washes with Wash Buffer (20mM HEPES-NaOH pH 7.5, 150mM NaCl, 0.5mM Spermidine supplemented with EDTA-free Protease Inhibitor Cocktail (APExBIO)) at RT, cells were resuspended in 1 mL Wash Buffer and incubated for 10 minutes while gently rotating with 20 µL of pre-activated Concanavalin A-coated magnetic beads (EpiCypher). The buffer was then removed, and the beads were resuspended in 150 µL of Antibody Buffer (Wash Buffer with 0.1% Digitonin (Sigma) and 2mM EDTA, pH 8) containing the antibody of interest, and left gently rotating overnight with a 180° angle at 4°C. On the next day, cells were washed two times with ice-cold Dig-Wash Buffer (Wash Buffer with 0.1% Digitonin), resuspended in 150 µL of Dig-Wash Buffer supplemented with 1 µL of custom-made pA/G-MNase every mL and left rotating for 1h at 4°C. After that, cells were washed again two times with Dig-Wash Buffer, resuspended in 100 µL of the same buffer, and chilled into a thermoblock sitting on ice. To initiate the cleavage reaction, 2 µL of 100mM $CaCl_2$ was added to each tube, and the tubes were left at 0°C for 30 minutes. To halt the reaction, 100 µL of a 2X STOP Buffer (340mM NaCl, 20mM EDTA pH 8, 4mM EGTA pH 8, 0.05% Digitonin, 100 µg/mL RNAse A, 50 µg/mL Glycogen) were added to each tube, and cleaved DNA fragments were released by incubating the samples for 30' at 37°C. The tubes were then centrifuged for 5 minutes at 4°C / 16,000g and the supernatant was collected. Finally, the DNA was purified by standard Phenol/Chloroform extraction using the 5PRIME Phase Lock Gel Heavy tubes (QuantaBio), and the success of the CUT&RUN reaction was assessed by running the positive control (H3K27me3 or H3K4me3) on a Tapestation (Agilent Technologies) using a High Sensitivity D1000 ScreenTape (Agilent Technologies). The complete list of primary and secondary antibodies and relative dilutions is available in **Table S5**.

RNA-seq and CUT&RUN Library Preparation

RNA-seq libraries were prepared using the CORALL mRNA-Seq Kit V2 (Lexogen), following the manufacturer's instructions. Prior to library generation, 1 µg of total RNA was either PolyA-enriched using the Poly(A) RNA Selection Kit V1.5 (Lexogen), or depleted of rRNAs using the RiboCop rRNA depletion kit for HMR V2 (Lexogen). To prepare the CUT&RUN libraries, we followed the original protocol[87], with slight modifications which were described previously[83]. The libraries were then quality-checked by both dsDNA Qubit (Thermo Fisher Scientific) and Tapestation (Agilent Technologies). All libraries were sequenced on an Illumina NovaSeq 6000 or Novaseq X instrument, aiming for either 10M (CUT&RUN) or 15M (RNA-seq) reads per sample.

RNA-seq data analysis

Analysis of the RNA-seq libraries was performed as previously described[83]. Raw FASTQ files were processed by a customized Lexogen CORALL analysis pipeline script (https://github.com/Lexogen-Tools/corall_analysis). Adaptors were trimmed by using Cutadapt[88], and the resulting trimmed FASTQ files were used as input for mapping to the hg19 human genome with STAR aligner[89], using the `--quantMode GeneCounts` option in order to count the number of reads mapping to each feature provided by an appropriate GTF file. To call for differential expression, we used DESeq2[90], considering a gene to be differentially expressed when adjusted $P<0.05$ and absolute $\log_2 FC > 0.41$ (~33% change in expression). Prior to any subsequent analysis, we filtered out poorly expressed genes, pseudogenes, and genes with an exonic length shorter than 200nt. To perform GO enrichment analysis and GSEA, we used the Clusterprofiler R package[91] by considering a term to be enriched when the adjusted $P<0.05$. Whenever several redundant terms were found, we employed the *simplify* function with a cutoff=0.7 to reduce the redundancy.

CUT&RUN data analysis

Raw FASTQ files were aligned to the hg19 reference genome with Bowtie2[92], using the -p 8, -X 2000 and --no-unal options. Peaks were called using the MACS2 callpeak function[93]. Peak quantification and differential expression were performed using the HOMER tool suite[94], using the *annotatePeaks.pl* and *getDiffExpression.pl* functions, respectively. Lastly, to plot the metagene profiles and coverage heatmaps, we employed deepTools[95].

Computational screen to identify putative *cis*-acting lncRNAs

We reasoned that lncRNAs that are likely to act in *cis* will be occasionally co-regulated with their target genes and also found in spatial proximity to the target promoters. We therefore sought such lncRNA-target pairs using the GeneHancer database[22] that connects between regulatory elements and their targets by combining tissue co-expression data, ChIP-seq of enhancer-enriched histone modifications and transcription factor binding, Hi-C data and eQTLs within putative enhancer regions. We complemented this by analyzing several RNA-seq datasets of MCF-7 breast cancer cells exposed to different conditions[23–25], to infer genome-wide transcriptional alterations across a variety of different experimental settings. We then focused on

those cases in which a lncRNA 1) overlaps a high-confidence GeneHancer (GH) element, 2) is differentially expressed in at least one of the conditions analyzed, and 3) the targets of the aforementioned GH element are also concomitantly differentially expressed. Because this approach identified over 600 candidate lncRNA-target pairs (**Table S1**), we first discarded lncRNAs which have been already reported to act in *cis*, and prioritized those cases in which the lncRNA is well expressed and/or the target genes are biologically interesting. This led us to focus on four lncRNA-target pairs connected in GeneHancer and in which both genes were significantly either upregulated or downregulated in the treatment condition, compared to controls (**Fig. S1A-D**).

Chromosome Conformation Capture (3C)

Between 5 and $10x10^6$ cells for each condition/replicate were collected and washed once with 1X PBS at RT. Cells were then resuspended in 10 mL of ice-cold 1X PBS with 10% FBS, supplemented with another 10 mL of 1X PBS with 10% FBS and 4% Formaldehyde, and left gently rotating for 5' at RT. Quenching was achieved by adding 1020 µL 2.5 M Glycine to each tube and placing them on ice to cool down, after that the cells were centrifuged for 8' at 4°C / 300 g and washed once with ice-cold 1X PBS. The pellet was then resuspended in 10 mL of Nuclear Permeabilization Buffer (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 0.5% NP-40) supplemented with EDTA-free Protease Inhibitor Cocktail (APExBIO)), gently rotated for 1 h in a cold room and centrifuged for 5' at 4°C / 600 g. After discarding the supernatant, the nuclei were resuspended in ~1 mL of ice-cold 1X PBS, moved to a 1.5 mL centrifuge tube, centrifuged and washed again with 1mL ice-cold 1X PBS, and eventually frozen as a pellet at -80 °C.

After 1+ days at -80 °C, each pellet was thawed on ice and gently resuspended in 175.5 µL of nuclease-free H_2O, after which 24.5 µL of 10X DpnII Reaction Buffer was added and the tubes were moved to a thermomixer set at 37°C. Each reaction was then supplemented with 3 µL of pre-warmed 20% SDS, incubated for 1h at 37°C / 900 rpm, supplemented again with 20 µL of pre-warmed 20% Triton-X100 and incubated for an additional 1h at 37°C / 900 rpm, after which 15 µL of DpnII (NEB, 50.000 U/mL) were added and the reaction was finally incubated O/N at 37°C / 900 rpm. The following day, the restriction reaction was removed by centrifugation for 5' at 4°C / 600 g, and the pellets were washed once with ice-cold 1X PBS, eventually leaving a small volume (~30 µL) on top of the pellet. The tubes were then incubated for 20' at 65°C, to inactivate DpnII activity, followed by thorough resuspension in 445 µL of nuclease-free H_2O and were transferred to ice. After chilling, 50 µL of 10X Ligation Buffer (NEB), 0.5 µg BSA and 5 µL of Quick Ligase (NEB) were added to each tube, and left incubating O/N at 16°C. After the O/N incubation, the ligation reaction was removed by centrifugation for 5' at 4°C / 600 g, the pellets were washed once with ice-cold 1X PBS and eventually resuspended in 250 µL 10 mM Tris-HCl, pH 7.5. A total of 30 µg of RNase A were then added to each sample, which were subsequentially incubated for 45' at 37°C / 750 rpm, after which 15 µL of PCR-grade Proteinase K were added and incubated O/N at 65°C / 750 rpm. The following day, the de-crosslinked reactions (3C template) were cleaned with 0.5X Sera-Mag Size Selection Beads (Cytiva Life Sciences), according to manufacturer's instructions, eluted in 200 µL Elution Buffer (Qiagen) and quantified by dsDNA Qubit.

UMI Circular Chromosome Conformation Capture (UMI-4C)

Quantitative UMI-4C was performed as previously described[41], with minor modifications. An equal amount of 3C template (5-10 µg) for each sample/replicate was diluted to a final volume of 200 µL with Elution Buffer and sonicated to an extent the fragments size ranges between 250 and 1000bp. To the sonicated fragments, 20 µL of 10X End-Repair Buffer (NEB) and 10 µL of End Repair Mix (NEB) were added and the samples were incubated for 30' at 20°C, after which they were cleaned with 2.2X Sera-Mag Selection Beads and eluted in 78 µL 10 mM Tris-HCl, pH 8.0. To 76 µL of the cleaned DNA, 10 µL 10X NEBuffer 2, 10 µL 10 mM dATP, and 4 µL Klenow (3'-5' exo, NEB) were added, and the reaction was incubated for 30' at 37°C. After this, the enzymes were inactivated by incubating the samples for 20' at 75°C, then 2 µL of FastCIP (NEB) were added and the reactions were further incubated for 1 h at 50°C, followed by purification with 2X Sera-Mag Size Selection Beads and elution in 67 µL of 10mM Tris-HCl, pH 8.0. Illumina-compatible adapters were ligated by adding 80 µL Quick Ligase Buffer (NEB), 10 µL Quick Ligase (NEB) and 4 µL of 15µM Y-Shaped Adaptors (IDT) to 66 µL of the eluted DNA from the previous step, followed by an incubation at 25°C for 15'. After that, the samples were denatured for 2' at 95°C, placed on ice, brought to volume by adding 160 µL of nuclease-free H_2O, cleaned with 1X Sera-Mag Size Selection Beads and eluted in 40 µL 10mM Tris-HCl, pH 8.0. To quantify the resulting 4C template, 1 µL was diluted in 4 µL of 10mM Tris-HCl, pH 8.0, denatured for 5' at 95°C, placed on ice and analyzed by ssDNA Qubit.

Libraries were then prepared using a nested PCR approach, preparing five reactions for each sample and employing three bait primers for each target (EPB41L4A-AS1 TSS, EPB41L4A TSS, STARD4 TSS, see next for design criteria) to increase complexity of the 4C profiles. For the first PCR (PCR1), upstream bait primers were designed close (<80-85 bp) to a DpnII restriction site (hg19 genome build); for the second PCR (PCR2), the downstream bait primers were designed to follow – but not overlap – the upstream ones, ending ideally <20bp from the DpnII site ("pad" sequence), and containing a sequence consisting of the Universal Illumina P5 on their 5' end. Both primers were designed to have a standard annealing temperature of ~60°C, and their sequences are available in **Table S5**. PCR1 consisted of 200 ng of 4C template, 10 µL of 10X Phusion Buffer (NEB), 1 µL of 10 mM dNTPs, 2.5 µL of pooled upstream bait primers with a final concentration of 10 µM, 2.5 µL of 10 µM Universal Illumina P7 primer, 0.5 µL Phusion Polymerase (NEB), and nuclease-free H_2O up to a final volume of 50 µL. The reactions were incubated for 30'' at 98°C, amplified with 20 cycles consisting of 10'' at 98°C, 30'' at 56°C and 60'' at 72°C, with a final extension of 10' at 72°C, individually purified with 0.8X Sera-Mag Size Selection Beads and eluted in 20 µL of 10 mM Tris-Hcl, pH 8.0. PCR2 consisted of 15 µL of purified template from PCR1, 10 µL of 10X Phusion Buffer (NEB), 1 µL of 10 mM dNTPs, 2.5 µL of pooled downstream bait primers with a final concentration of 10 µM, 2.5 µL of 10µM Universal Illumina P7 primer, 0.5 µL Phusion Polymerase (NEB), and nuclease-free H_2O up to a final volume of 50 µL. The reactions were incubated with the same program of PCR1 but with 18 cycles of amplification instead of 20, all five replicates of each sample were then pooled together, purified with 0.8X Sera-Mag Size Selection Beads and eluted in 40 µL of Elution Buffer. The libraries were then quality-checked by both dsDNA Qubit and Tapestation, and eventually pooled according to the relative number of reads needed. All libraries were sequenced on an Illumina NovaSeq 6000 instrument, aiming at ~5M reads per sample.

UMI-4C Data Analysis

To analyze the UMI-4C data we have employed the umi4cpackage R package (https://github.com/tanaylab/umi4cpackage). After setting up the configuration file as described, they were loaded with the *p4cLoadConfFiles* function, and the raw fastq files were analyzed using the *p4cCreate4CseqTrack* function. The 4C profiles were then built and plotted using the *p4cNewProfile* and plot functions, respectively, specifying the *cis* window and overlaying two profiles when needed. The p-values and $\log_2$FC between conditions and across a genomic region of interest were finally calculated using the *p4cIntervalsMean* function.

Metabolic labeling, SLAM-seq and data analysis

For half-life and synthesis rate estimation upon the KD of *EPB41L4A-AS1*, cells were transfected with control and targeting GapmeRs as described above. Metabolic labeling and SLAM-seq were then performed as previously described[96], with minor modifications. All steps were performed in the dark until the chemical conversion with iodoacetamide (IAA). After 48h hours, the growth medium was replaced with fresh media supplemented with 500 µM 4-Thiouridine (4sU, Sigma), and replaced again with fresh 4sU after every time point. In total, we collected RNA from cells at time 0 (unlabeled), and after 4, 8 and 24h hours post-4sU addition. RNA extraction was carried out as described above with the addition of DTT (0.1 mM final concentration) to keep the samples under reducing conditions, and IAA conversion was performed as described previously[71,96]. Ribosomal RNAs were then depleted using the NEBNext rRNA Depletion Kit V2 (NEB) and libraries were generated with the NEBNext Ultra II Directional Library Kit (NEB), both according the manufacturer's instructions. Estimation of half-lives and synthesis rates was performed using GRAND-SLAM/grandR[97], and are listed in **Table S4**. Notably, the estimation assumes steady state expression in control samples, but not in the KD ones.

Data availability

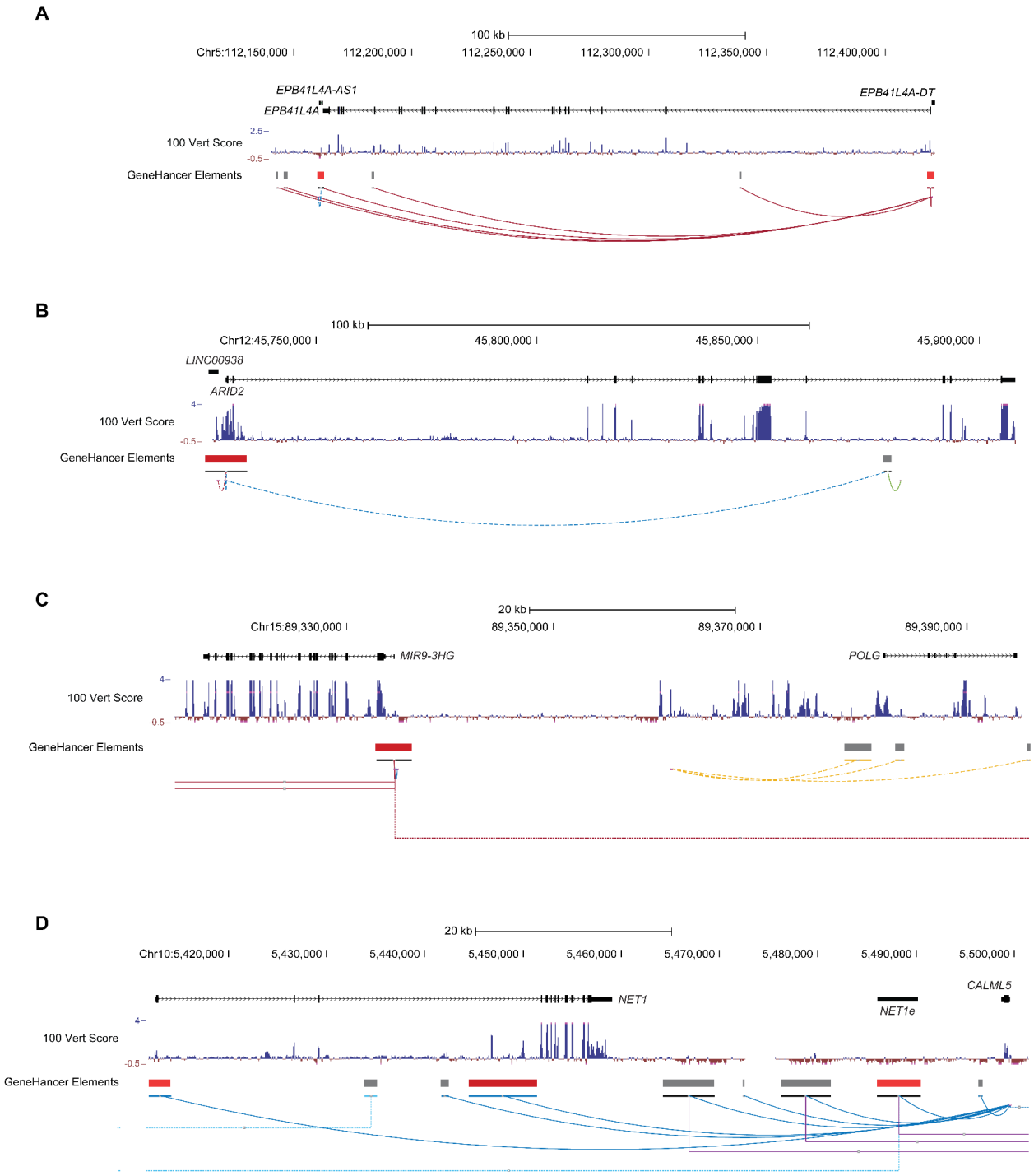
All the RNA-seq, RIP-seq, UMI-4C, and Cut&Run datasets were deposited to the GEO database under the accession GSE292272.

Author contributions

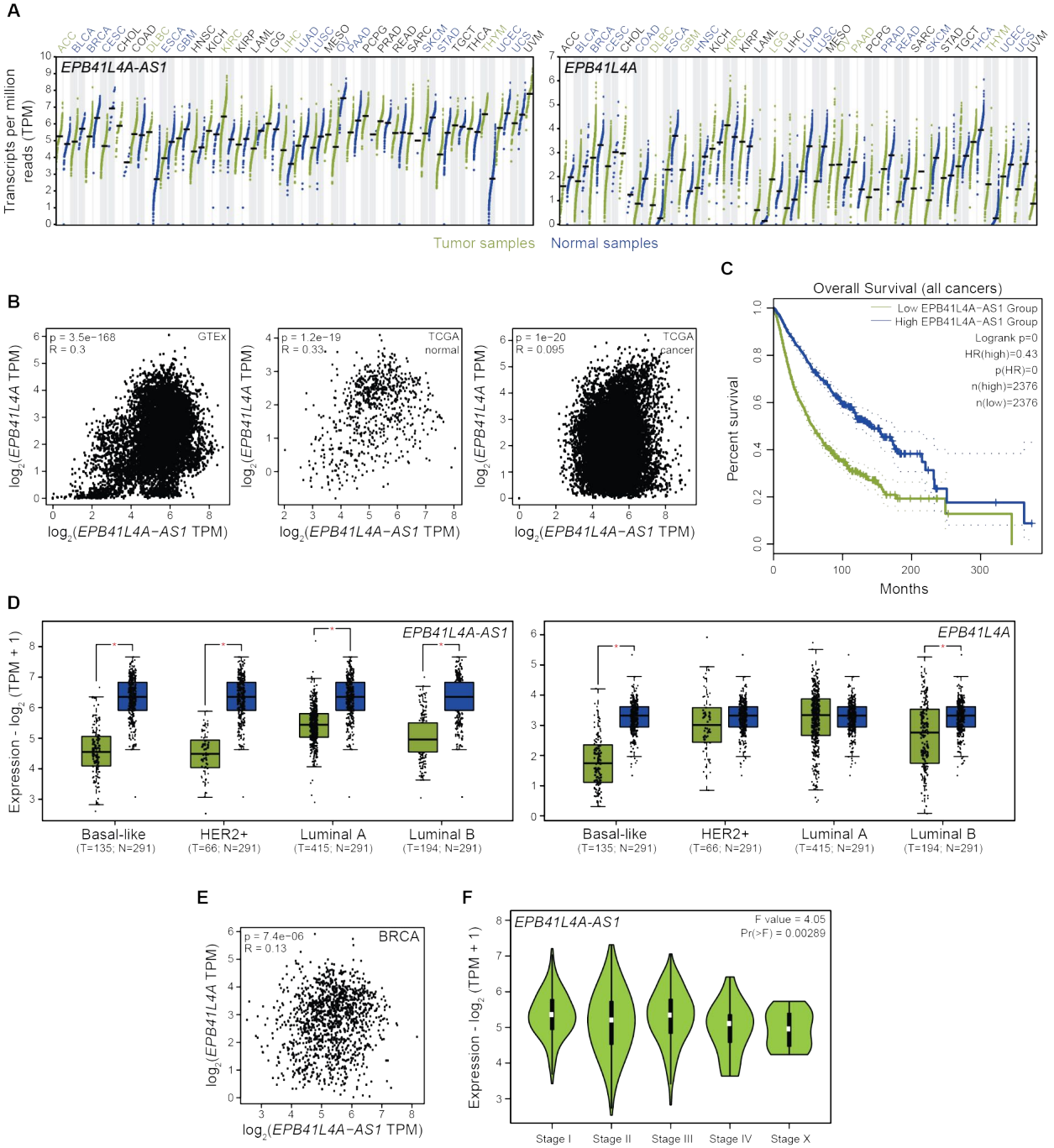
A.M. and I.U. conceived the study. A.M. carried out most experiments and analyzed the data. J.P.U. helped with experimental designs, UMI-4C experiments and provided insightful input. T.C. assisted with stress induction experiments. A.M. and I.U. wrote the manuscript with input from the other authors.

Acknowledgements

We thank members of the Ulitsky lab for insightful discussions and comments on the manuscript. We thank Dana Hirsch and Liat Alyagor for assistance with smFISH experiments, Ehud Sivan for imaging data analysis, Noa Gil for discussions about UMI-4C, Florian Erhard for assistance with SLAM-seq analysis and Josh Mendell's lab for the SNORA13 over-expression plasmid. This work was supported by the ERC Consolidator Grant lncIMPACT to IU.

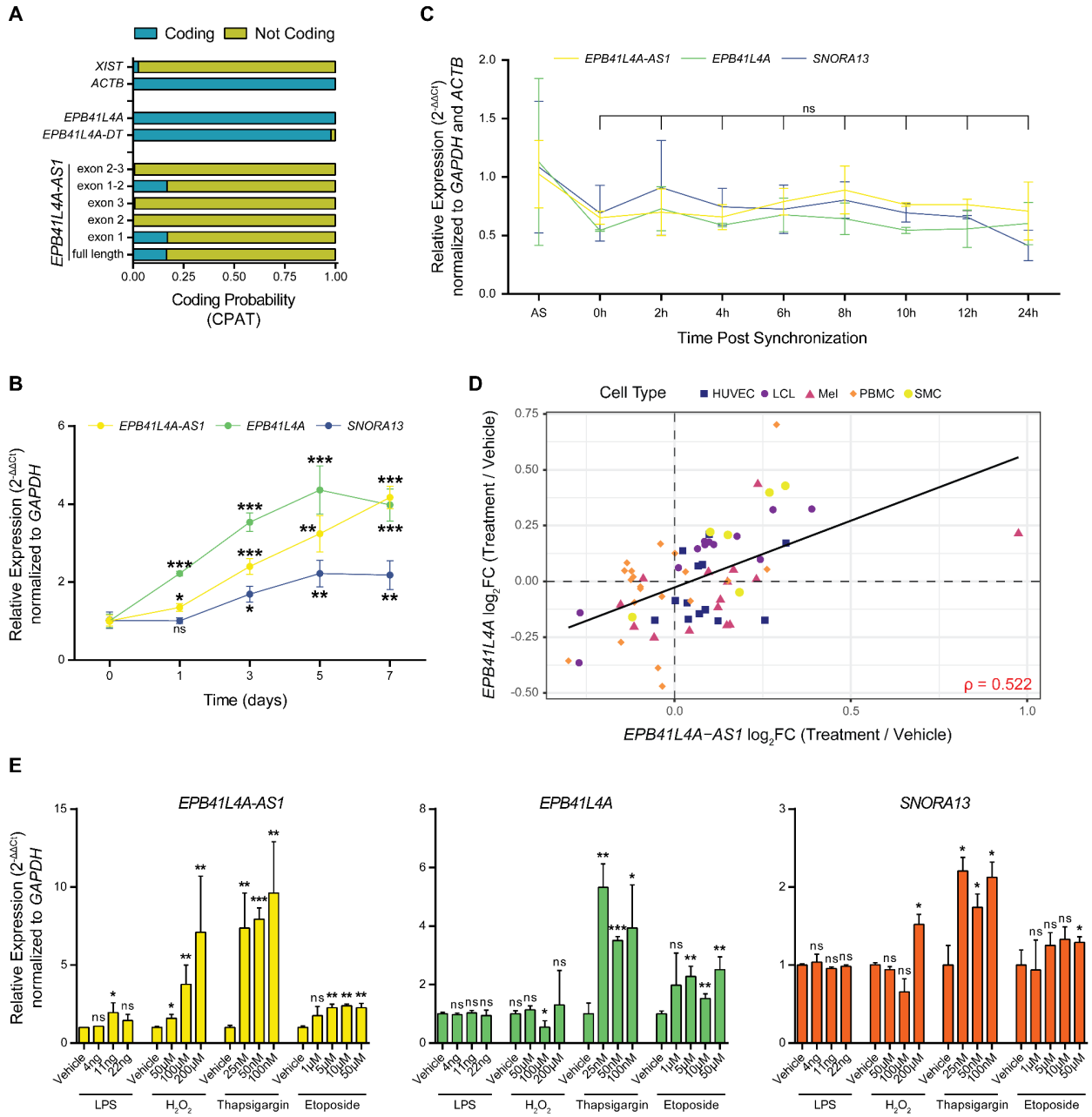


Supplementary Figure 1. **UCSC Genome Browser view of the loci containing the selected lncRNAs and target genes for validation**. View of the *EPB41L4A-AS1* - *EPB41L4A* **(A)**, *LINC00938* - *ARID2* **(B)**, *MIR9-3HG* - *POLG* **(C)** and *NET1e* - *NET1* - *CALML5* **(D)** loci. In each view, the PhyloP conservation score and GeneHancer (GH) elements are also reported.



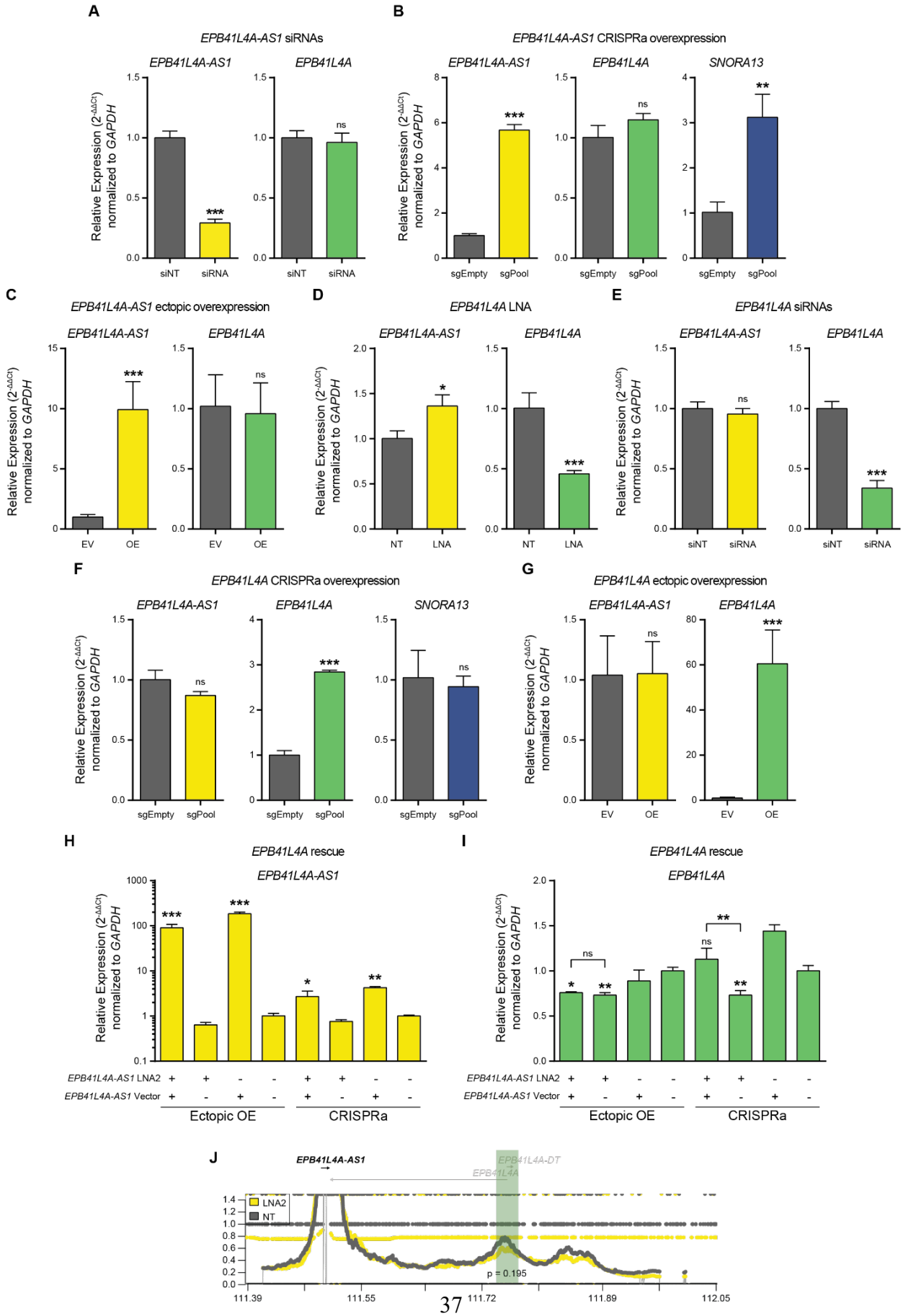
Supplementary Figure 2. ***EPB41L4A-AS1* is dysregulated in several cancer types and correlates with survival**. **(A)** Expression level of *EPB41L4A-AS1* (left) and *EPB41L4A* (right) across the TCGA cohort. The thick lines and each point represent the median and an individual sample, respectively, and the color on top reflects whether the indicated genes are significantly more expressed in control or tumor conditions. **(B)** Correspondence between the expression of *EPB41L4A-AS1* and *EPB41L4A* in GTEx v8 (left), TCGA control (center) and tumor (right) samples. The Spearman's correlation coefficient and p-value is also reported. **(C)** Kaplan-Meier survival curve stratified into individuals with the lower and higher 50% expression level of *EPB41L4A-AS1*. **(D)** Expression level of *EPB41L4A-AS1* (left) and *EPB41L4A* (right) in patients with the different BRCA subtypes and matched controls. The thick lines, edges of the box, whiskers and each point represent the median, first and third quartiles, the upper and lower 1.5 interquartile ranges (IQRs), and an individual sample, respectively. **(E)**

Correspondence between the expression of *EPB41L4A-AS1* and *EPB41L4A* in BRCA samples. **(F)** Expression level of *EPB41L4A-AS1* in patients at different BRCA stages. The white dot and edges of the box represent the median and the first and third quartiles, respectively. The one-way ANOVA p-value for differential expression is also reported. In all cases, P<0.05 = *. Data was accessed using the GEPIA2 portal[98].

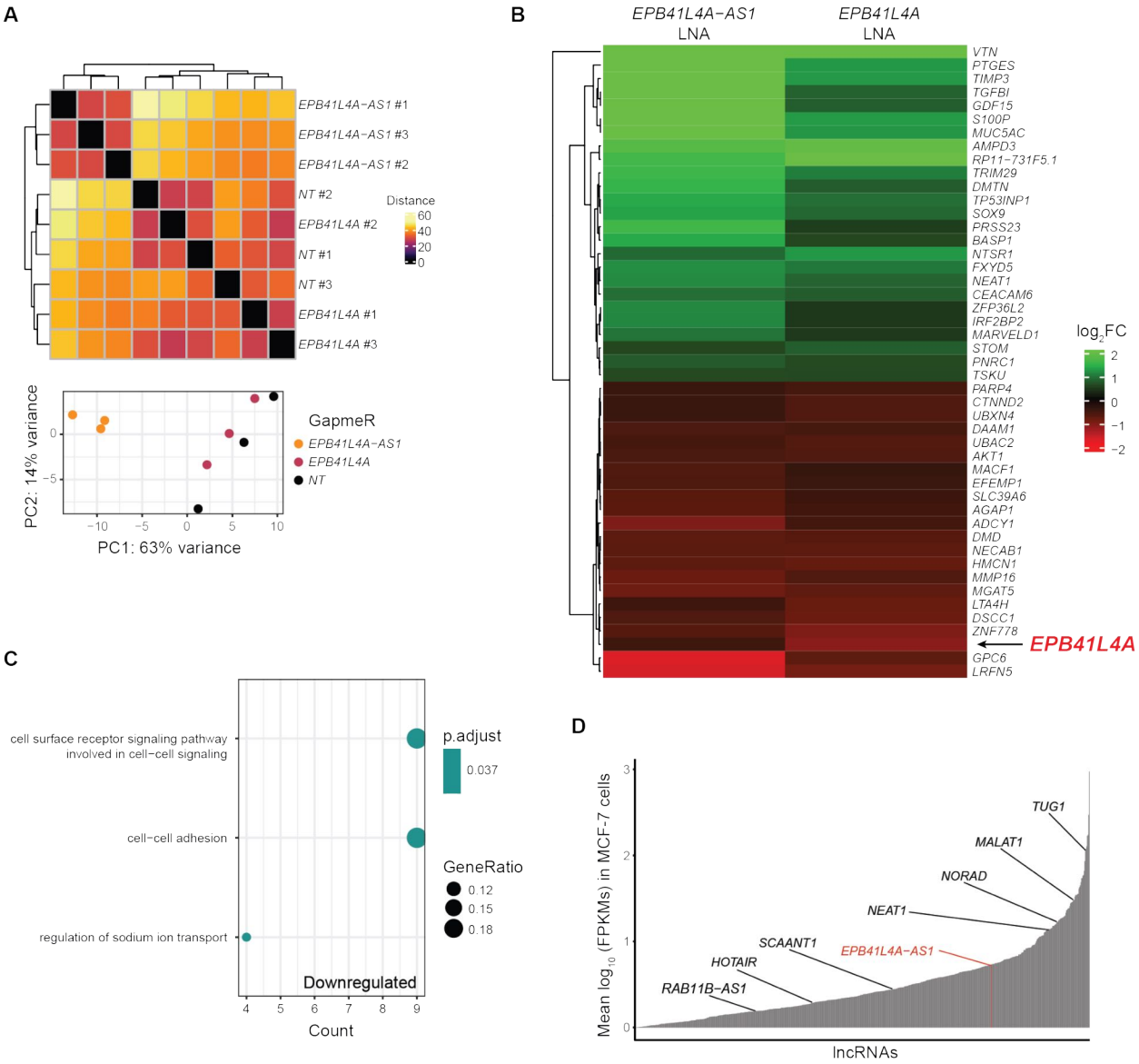


Supplementary Figure 3. ***EPB41L4A-AS1* expression is altered upon multiple stimuli**. **(A)** CPAT[37] analysis of the indicated transcripts. As controls for a non-coding and a coding transcript, *XIST* and *ACTB* are reported. Changes in gene expression for the indicated genes **(B)** during a time course of 7 days in serum starvation conditions, **(C)** following the release from a double thymidine cell cycle block, **(D)** in a panel of 250 unique cell

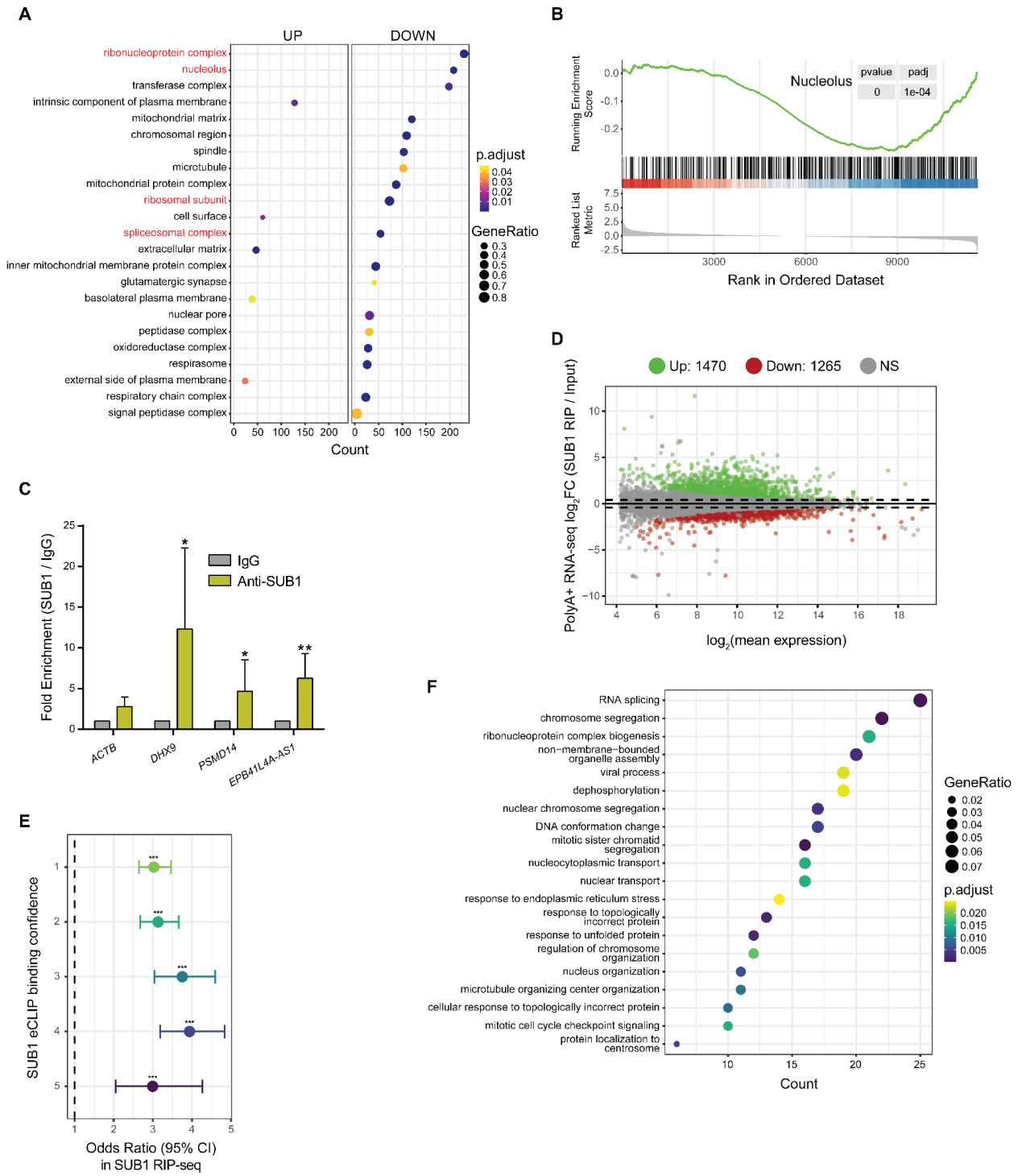
type-treatment combination (the Spearman's correlation coefficient and trendline are also reported), **(E)** and after exposing MCF-7 cells to LPS, H_2O_2, thapsigargin and etoposide. When applicable, all experiments were performed in n = 3 biological replicates, with the error bars in the dotplots and boxplots representing the standard deviation. In all cases, ns - P>0.05 = ns; * - P<0.05 = *; ** - P<0.01 = **; *** - P<0.001 = *** (two-sided Student's t-test).


A
EPB41L4A-AS1 siRNAs
EPB41L4A-AS1
EPB41L4A
Relative Expression (2^-ΔΔCt) normalized to GAPDH
siNT
siRNA
***
ns
B
EPB41L4A-AS1 CRISPRa overexpression
SNORA13
sgEmpty
sgPool
C
EPB41L4A-AS1 ectopic overexpression
EV
OE
D
EPB41L4A LNA
NT
LNA
*
E
EPB41L4A siRNAs
F
EPB41L4A CRISPRa overexpression
G
EPB41L4A ectopic overexpression
H
EPB41L4A rescue
EPB41L4A-AS1
EPB41L4A-AS1 LNA2
EPB41L4A-AS1 Vector
Ectopic OE
CRISPRa
I
EPB41L4A rescue
EPB41L4A
J
EPB41L4A-AS1
EPB41L4A-DT
EPB41L4A
LNA2
NT
p = 0.195
111.39
111.55
111.72
111.89
112.05

Supplementary Figure 4. ***EPB41L4A-AS1* unidirectionally facilitates *EPB41L4A* expression in *cis***. RT-qPCR to assess the expression of the reported genes after **(A)** transfection with siRNAs against *EPB41L4A-AS1*, **(B)** CRISPRa with guides targeting the *EPB41L4A-AS1* promoter, **(C)** transfection with plasmid encoding the *EPB41L4A-AS1* cDNA, transfection with GapmeRs **(D)** and siRNAs **(E)** targeting *EPB41L4A*, **(F)** CRISPRa with guides targeting the *EPB41L4A* promoter and **(G)** transfection with plasmid encoding the *EPB41L4A* cDNA. **(H)** Changes in *EPB41L4A-AS1* expression after rescuing *EPB41L4A-AS1* with an ectopic plasmid or CRISPRa following its KD with GapmeRs. Asterisks indicate significance relative to the –/– control. **(I)** Same as in (H), but for changes in *EPB41L4A* expression. **(J)** UMI-4C contact profiles in control and LNA2-transfected cells using baits targeting the TSS of *EPB41L4A-AS1*. The green area represents the quantified genomic interval, and the p-value was calculated using a Chi-squared test. All experiments were performed in n = 3 biological replicates, except UMI-4C with n=2, with the error bars in the boxplots representing the standard deviation. In all cases, ns - P>0.05 = ns; * - P<0.05 = *; ** - P<0.01 = **; *** - P<0.001 = *** (two-sided Student's t-test).

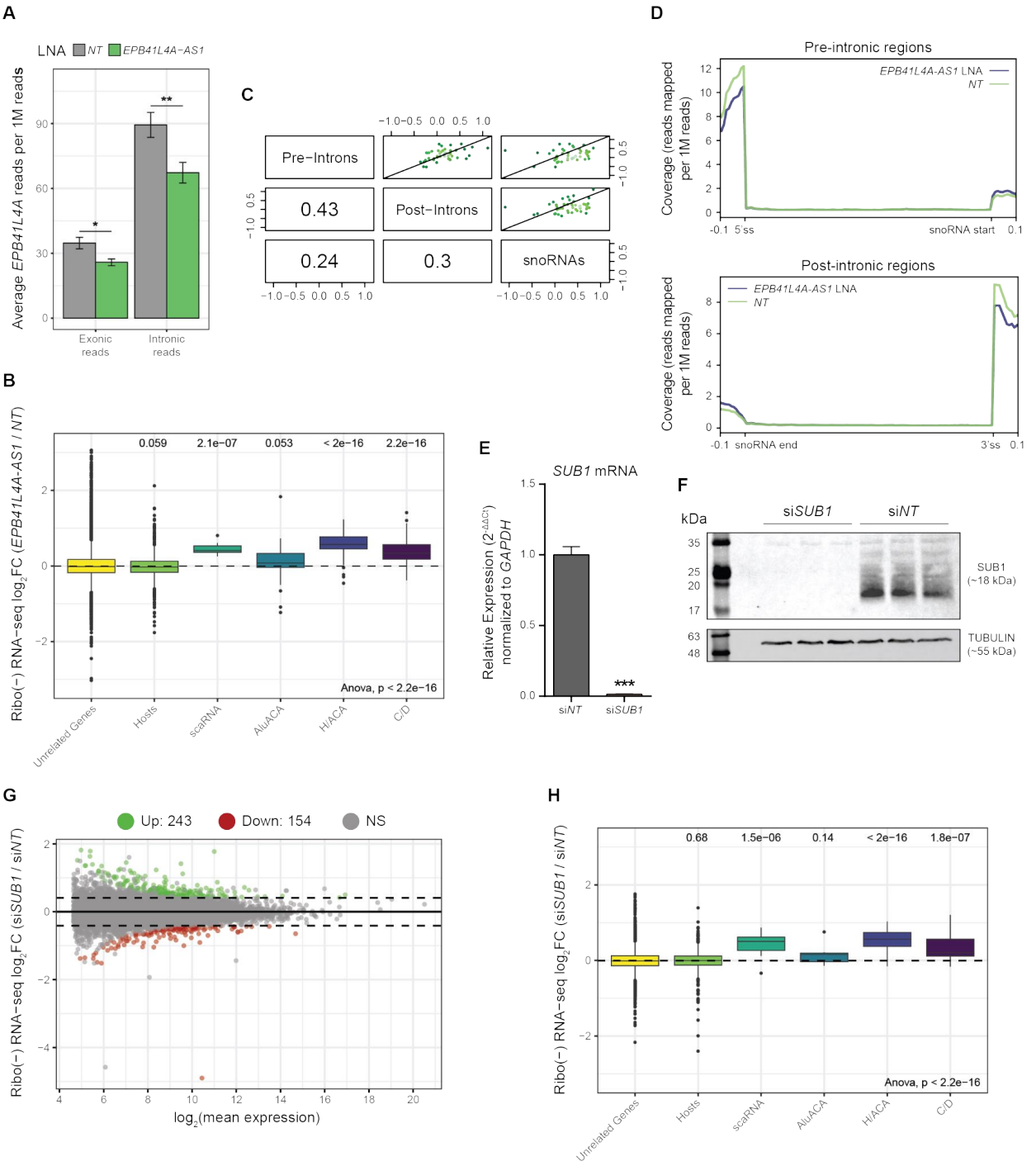


Supplementary Figure 5. **Most of the transcriptomic changes following *EPB41L4A-AS1* downregulation are not explained by *EPB41L4A*. (A)** Hierarchical clustering (top) and Principal Component Analysis (PCA, bottom) of the polyA+ RNA-seq libraries. The color in the heatmap represents the Euclidean distance between the libraries. **(B)** Heatmap of the differentially expressed genes shared between *EPB41L4A-AS1* and *EPB41L4A* depletions. Color intensity reflects the changes in gene expression (log_2FC). **(C)** GO enrichment analysis of the genes downregulated following *EPB41L4A* depletion. **(D)** Barplot of the lncRNAs (annotated in GENCODE) detected in MCF-7 cells, ordered by their expression. *EPB41L4A-AS1* and other representative *cis*- or *trans*-acting lncRNAs are highlighted. All experiments were performed in n = 3 biological replicates, and a gene was considered to be differentially expressed if both adjusted $P<0.05$ and $|log_2Fold\text{-}change| >0.41$ (corresponding to a change of 33%)



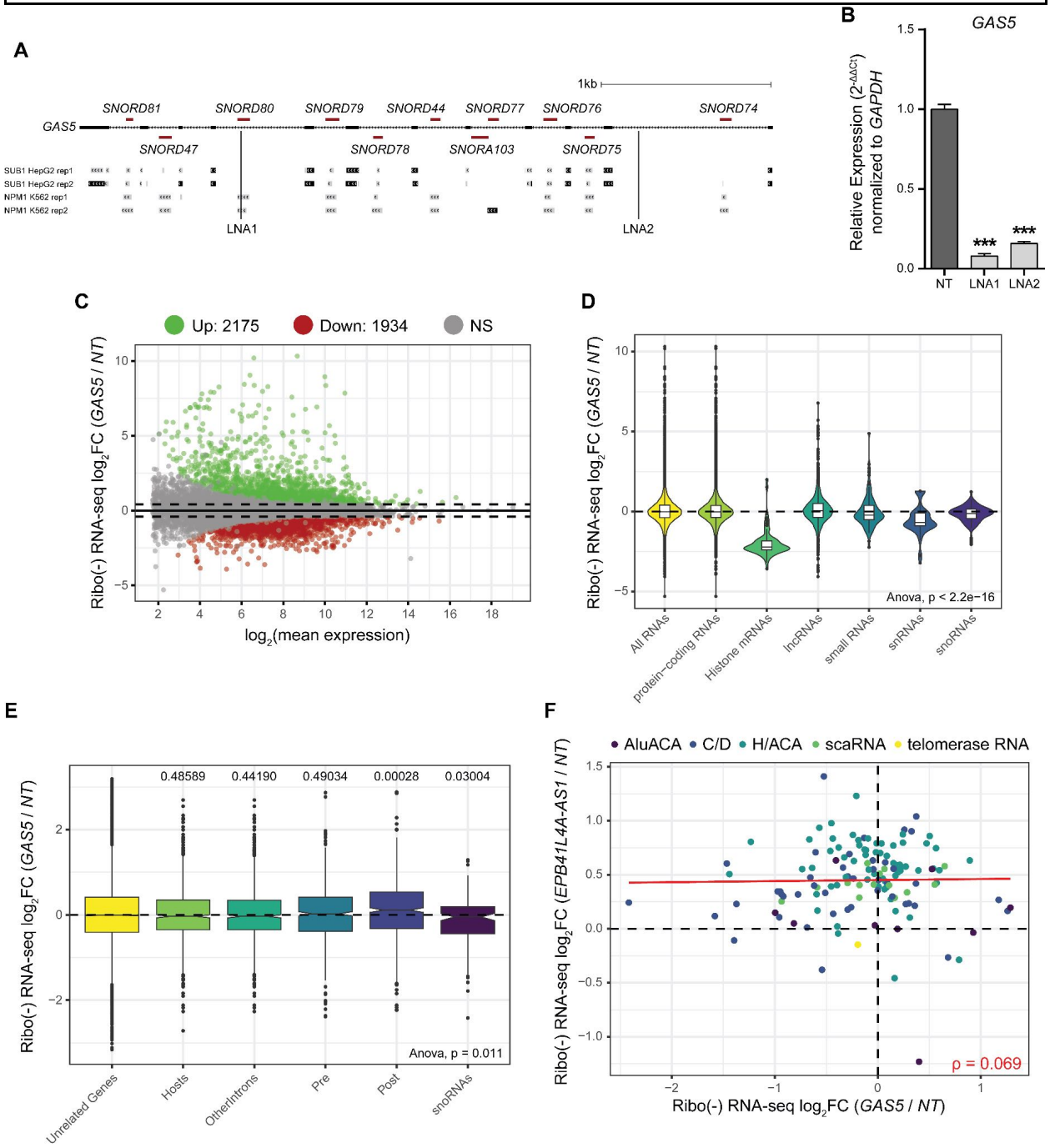
Supplementary Figure 6. **SUB1 is both a chromatin-associated and an RNA-binding protein. (A)** Gene set enrichment analysis (GSEA) using *cellular component* as ontology. **(B)** Running score and preranked list of the nucleolus-associated genes in the GSEA analysis. **(C)** RT-qPCR for the indicated genes after RIP with either an anti-SUB1 or control IgG antibody. **(D)** MA plot showing the relative enrichment in RIP-seq after using either an anti-SUB1 or control IgG antibody. **(E)** Forest plot for the enrichment of the genes in the ENCODE SUB1 eCLIP in our RIP-seq data. The genes in the eCLIP dataset were ranked according to their enrichment score. **(F)** GO enrichment analysis of the genes with a SUB1 peak as determined from our CUT&RUN data. All experiments

were performed in n = 3 biological replicates, with the error bars representing the standard deviation (C) or the 95% confidence interval (E). A gene was considered to be differentially expressed if both adjusted $P<0.05$ and $|\log_2 \text{Fold-change}| > 0.41$ (corresponding to a change of 33%). In all cases, ns - $P>0.05$ = ns; * - $P<0.05$ = *; ** - $P<0.01$ = **; *** - $P<0.001$ = *** (two-sided Student's t-test in (C), Fisher's exact test in (E)).

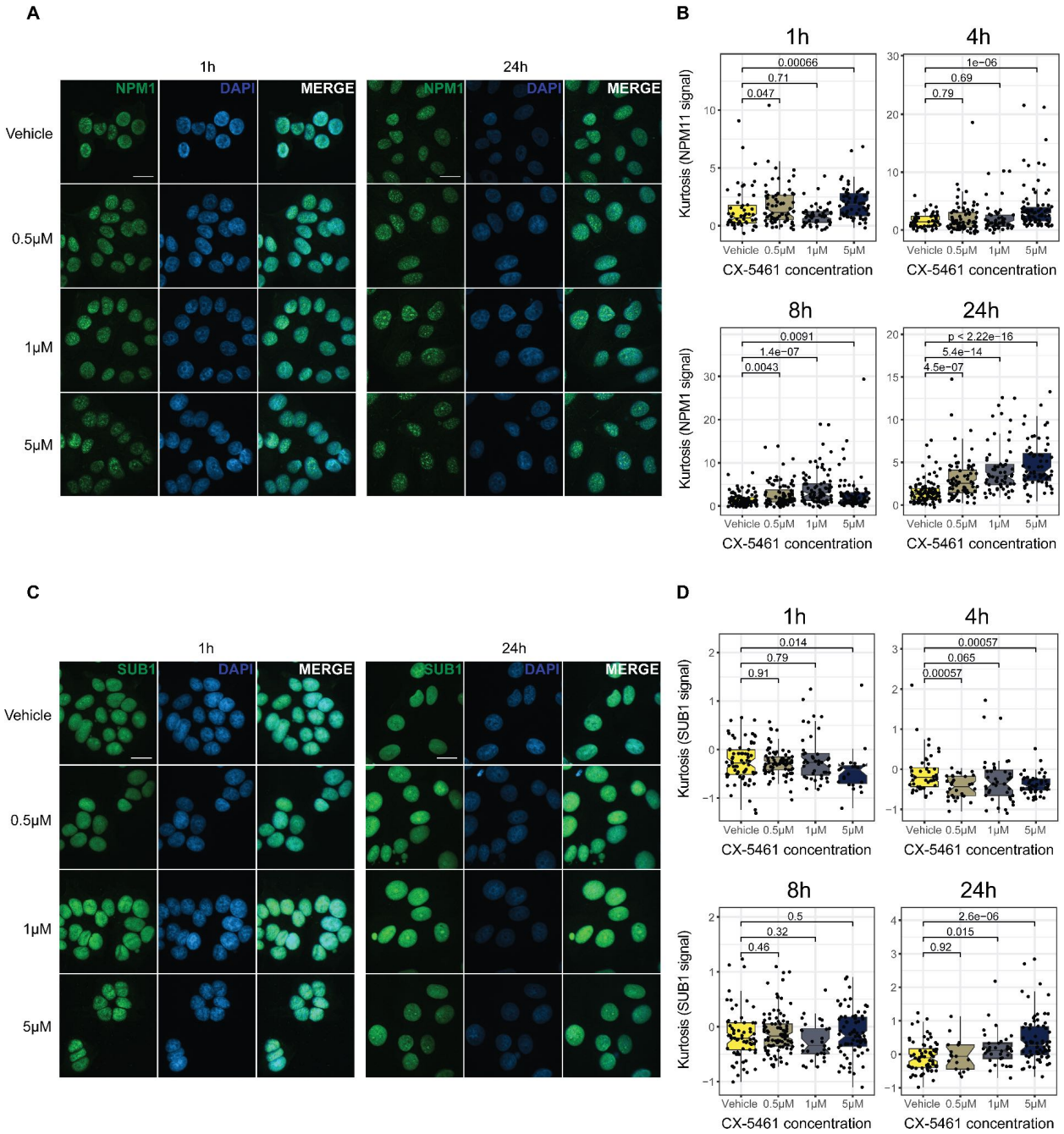


Supplementary Figure 7. ***EPB41L4A-AS1* and SUB1 depletion affect the expression of different classes of snoRNAs. (A)** Barplot depicting the normalized read density in the Ribo(-) RNA-seq data in *EPB41L4A* exons and introns. **(B)** Boxplot of gene expression changes ($\log_2$FC) after *EPB41L4A-AS1* depletion for the indicated RNA categories. **(C)** Correspondence of the changes in gene expression after *EPB41L4A-AS1* KD between the pre-intronic, post-intronic and snoRNA regions. Spearman's correlation coefficients are also reported. **(D)** Normalized Ribo(-) RNA-seq signal over the intronic regions before (top) and after (bottom) the expressed snoRNAs. RT-qPCR **(E)** and Western blot **(F)** upon *SUB1* depletion with siRNAs to assess KD efficiency. **(G)** MA plot showing the genome-wide gene expression changes after *SUB1* KD with siRNAs. **(H)** Same as in (B),

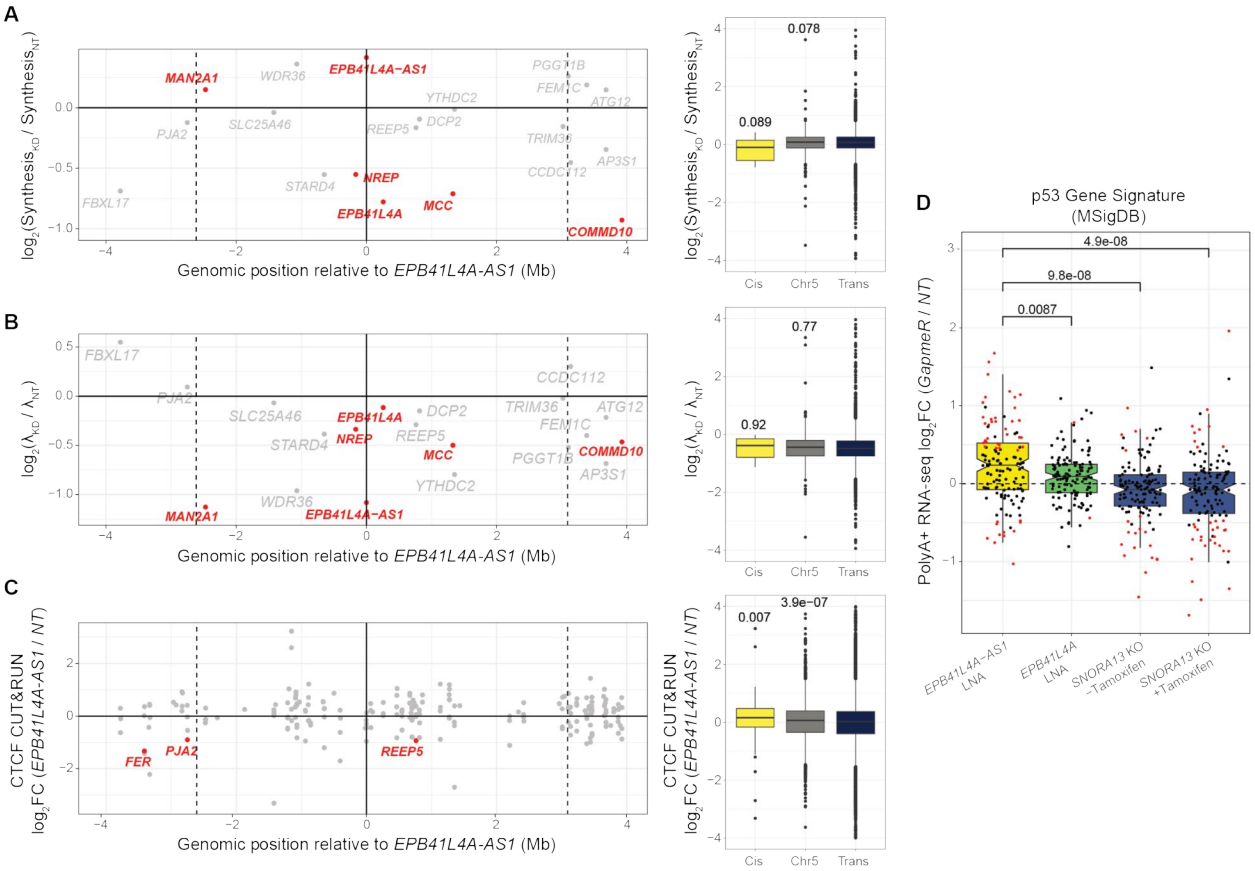
but after *SUB1* KD with siRNAs. All experiments were performed in n = 3 biological replicates, with the error bars in the barplots representing the standard deviation. In the boxplots, the thick line, edges of the box, and whiskers represent the median, first and third quartiles, and the upper and lower 1.5 interquartile ranges (IQRs), respectively. Outliers (observations outside the 1.5 IQRs) are drawn as single points, the significance of the different comparisons was computed by a Mann-Whitney test, and a global ANOVA p-value is also reported. A gene was considered to be differentially expressed if both adjusted P<0.05 and |$\log_2$Fold-change| >0.41 (corresponding to a change of 33%). In all cases, ns - P>0.05 = ns; * - P<0.05 = *; ** - P<0.01 = **; *** - P<0.001 = *** (two-sided Student's t-test).



Supplementary Figure 8. ***GAS5* depletion does not affect snoRNAs expression**. **(A)** Schematics of the *GAS5* locus with tracks depicting the eCLIP peaks for both SUB1 and NPM1 (source: ENCODE), as well as the location of the two GapmeRs that were used. **(B)** RT-qPCR upon *GAS5* depletion with GapmeRs to assess KD efficiency. **(C)** MA plot showing the genome-wide gene expression changes after *GAS5* KD with GapmeRs. **(D)** Violin/boxplots of gene expression changes (log_2FC) after *GAS5* depletion for the indicated RNA classes. **(E)** Boxplot of gene expression changes (log_2FC) after *GAS5* depletion for the indicated RNA categories. **(F)** Correspondence between the changes in gene expression of the color-coded snoRNA categories after *EPB41L4A-AS1* and *GAS5* depletion. The trendline and Spearman correlation coefficient are also reported. All experiments were performed in n = 3 biological replicates, with the error bars in the barplots representing the standard deviation. In the boxplots, the thick line, edges of the box, and whiskers represent the median, first and third quartiles, and the upper and lower 1.5 interquartile ranges (IQRs), respectively. Outliers (observations outside the 1.5 IQRs) are drawn as single points, the significance of the different comparisons was computed by a Mann-Whitney test, and a global ANOVA p-value is also reported. A gene was considered to be differentially expressed if both adjusted $P<0.05$ and $|log_2Fold\text{-}change| >0.41$ (corresponding to a change of 33%). In all cases, ns - $P>0.05$ = ns; * - $P<0.05$ = *; ** - $P<0.01$ = **; *** - $P<0.001$ = *** (two-sided Student's t-test).



Supplementary Figure 9. **Nucleolar stress induced by CX-5461 treatment affects SUB1 and NPM1 nuclear patterns. (A)** Representative immunofluorescence images for NPM1 after treating cells with the indicated CX-5461 concentrations and time (scale bar = 20 μm). **(B)** Boxplot depicting the changes of the Kurtosis of the nuclear NPM1 signal after treating MCF-7 cells with the indicated CX-5461 concentrations and time. **(C)** Same as in (A), but for SUB1. **(D)** Same as in (B), but for SUB1. All experiments were performed in n = 3 biological replicates. In the boxplots, the thick line, edges of the box, and whiskers represent the median, first and third quartiles, and the upper and lower 1.5 interquartile ranges (IQRs), respectively. Outliers (observations outside the 1.5 IQRs) are drawn as single points, the significance of the different comparisons was computed by a Mann-Whitney test.



Supplementary Figure 10. ***EPB41L4A-AS1* affects local RNA metabolism. (A)** Changes in synthesis rate in KD cells of the genes in the two flaking TADs of *EPB41L4A-AS1*. The vertical dotted lines represent the TAD boundaries (as assessed by TADmap), the continuous vertical line the lncRNA locus and inter-TAD boundary, and the horizontal continuous line a $\log_2$Fold-change equal to 0. The dots represent individual genes, with the significant (adjusted P<0.05 and |$\log_2$Fold-change| >0.41) ones highlighted in red. **(B)** Same as in (A), but for RNA half lives. **(C)** Same as in (A) and (B), but for CTCF binding as assessed by CUT&RUN. Each dot and name represent a single CTCF peak and the closest gene, respectively. **(D)** Changes in gene expression of the p53 signature genes (from MSigDB) after KD with GapmeRs targeting either *EPB41L4A-AS1* or *EPB41L4A* (polyA+ RNA-seq data), and *SNORA13* KO cells[27]. The points in the boxplot represent individual genes, and their color whether they were found to be significantly dysregulated (red) or not (black). All experiments were performed in n = 3 biological replicates. In the boxplots, the thick line, edges of the box, and whiskers represent the median, first and third quartiles, and the upper and lower 1.5 interquartile ranges (IQRs), respectively. Outliers (observations outside the 1.5 IQRs) are drawn as single points, the significance of the different comparisons was computed by a Mann-Whitney test.

References

1. Statello, L., Guo, C.-J., Chen, L.-L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. Nat. Rev. Mol. Cell Biol. *22*, 96–118.

2. Wu, H., Yang, L., and Chen, L.-L. (2017). The Diversity of Long Noncoding RNAs and Their Generation. Trends Genet. *33*, 540–552.

3. Ransohoff, J.D., Wei, Y., and Khavari, P.A. (2018). The functions and unique features of long intergenic non-coding RNA. Nat. Rev. Mol. Cell Biol. *19*, 143–157.

4. Mattick, J.S., Amaral, P.P., Carninci, P., Carpenter, S., Chang, H.Y., Chen, L.-L., Chen, R., Dean, C., Dinger, M.E., Fitzgerald, K.A., et al. (2023). Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat. Rev. Mol. Cell Biol. *24*, 430–447.

5. Hangauer, M.J., Vaughn, I.W., and McManus, M.T. (2013). Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. PLoS Genet. *9*, e1003569.

6. Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., et al. (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559–1563.

7. ENCODE Project Consortium, Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

8. Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermüller, J., Hofacker, I.L., et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science *316*, 1484–1488.

9. Mudge, J.M., Carbonell-Sala, S., Diekhans, M., Martinez, J.G., Hunt, T., Jungreis, I., Loveland, J.E., Arnan, C., Barnes, I., Bennett, R., et al. (2024). GENCODE 2025: reference gene annotation for human and mouse. Nucleic Acids Res. https://doi.org/10.1093/nar/gkae1078.

10. Ruiz-Orera, J., Messeguer, X., Subirana, J.A., and Alba, M.M. (2014). Long non-coding RNAs as a source of new peptides. Elife *3*, e03523.

11. Chen, J., Brunner, A.-D., Cogan, J.Z., Nuñez, J.K., Fields, A.P., Adamson, B., Itzhak, D.N., Li, J.Y., Mann, M., Leonetti, M.D., et al. (2020). Pervasive functional translation of noncanonical human open reading frames. Science *367*, 1140–1146.

12. Barczak, W., Carr, S.M., Liu, G., Munro, S., Nicastri, A., Lee, L.N., Hutchings, C., Ternette, N., Klenerman, P., Kanapin, A., et al. (2023). Long non-coding RNA-derived peptides are immunogenic and drive a potent anti-tumour response. Nat. Commun. *14*, 1078.

13. Kesner, J.S., Chen, Z., Shi, P., Aparicio, A.O., Murphy, M.R., Guo, Y., Trehan, A., Lipponen, J.E., Recinos, Y., Myeku, N., et al. (2023). Noncoding translation mitigation. Nature *617*, 395–402.

14. Chothani, S.P., Adami, E., Widjaja, A.A., Langley, S.R., Viswanathan, S., Pua, C.J., Zhihao, N.T., Harmston, N., D'Agostino, G., Whiffin, N., et al. (2022). A high-resolution map of human RNA

translation. Mol. Cell *82*, 2885–2899.e8.

15. Housman, G., and Ulitsky, I. (2016). Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs. Biochim. Biophys. Acta *1859*, 31–40.

16. Paralkar, V.R., Taborda, C.C., Huang, P., Yao, Y., Kossenkov, A.V., Prasad, R., Luan, J., Davies, J.O.J., Hughes, J.R., Hardison, R.C., et al. (2016). Unlinking an lncRNA from Its Associated cis Element. Mol. Cell *62*, 104–110.

17. Ali, T., and Grote, P. (2020). Beyond the RNA-dependent function of LncRNA genes. Elife *9*. https://doi.org/10.7554/eLife.60583.

18. Beucher, A., Miguel-Escalada, I., Balboa, D., De Vas, M.G., Maestro, M.A., Garcia-Hurtado, J., Bernal, A., Gonzalez-Franco, R., Vargiu, P., Heyn, H., et al. (2022). The HASTER lncRNA promoter is a cis-acting transcriptional stabilizer of HNF1A. Nat. Cell Biol. *24*, 1528–1540.

19. De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.-L., and Natoli, G. (2010). A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol. *8*, e1000384.

20. Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. Nat. Rev. Genet. *21*, 102–117.

21. Monziani, A., and Ulitsky, I. (2023). Noncoding snoRNA host genes are a distinct subclass of long noncoding RNAs. Trends Genet. https://doi.org/10.1016/j.tig.2023.09.001.

22. Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., et al. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database *2017*. https://doi.org/10.1093/database/bax028.

23. Sun, M., Gadad, S.S., Kim, D.-S., and Kraus, W.L. (2015). Discovery, Annotation, and Functional Analysis of Long Noncoding RNAs Controlling Cell-Cycle Gene Expression and Proliferation in Breast Cancer Cells. Mol. Cell *59*, 698–711.

24. Nagarajan, S., Rao, S.V., Sutton, J., Cheeseman, D., Dunn, S., Papachristou, E.K., Prada, J.-E.G., Couturier, D.-L., Kumar, S., Kishore, K., et al. (2020). ARID1A influences HDAC1/BRD4 activity, intrinsic proliferative capacity and breast cancer treatment response. Nat. Genet. *52*, 187–197.

25. Janky, R. 's, Verfaillie, A., Imrichová, H., Van de Sande, B., Standaert, L., Christiaens, V., Hulselmans, G., Herten, K., Naval Sanchez, M., Potier, D., et al. (2014). iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput. Biol. *10*, e1003731.

26. Babaian, A., Rothe, K., Girodat, D., Minia, I., Djondovic, S., Milek, M., Spencer Miko, S.E., Wieden, H.-J., Landthaler, M., Morin, G.B., et al. (2020). Loss of m1acp3Ψ Ribosomal RNA Modification Is a Major Feature of Cancer. Cell Rep. *31*, 107611.

27. Cheng, Y., Wang, S., Zhang, H., Lee, J.-S., Ni, C., Guo, J., Chen, E., Wang, S., Acharya, A., Chang, T.-C., et al. (2024). A non-canonical role for a small nucleolar RNA in ribosome biogenesis and senescence. Cell *187*, 4770–4789.e23.

28. Yabuta, N., Onda, H., Watanabe, M., Yoshioka, N., Nagamori, I., Funatsu, T., Toji, S., Tamai, K.,

and Nojima, H. (2006). Isolation and characterization of the TIGA genes, whose transcripts are induced by growth arrest. Nucleic Acids Res. *34*, 4878–4892.

29. Hegre, S.A., Samdal, H., Klima, A., Stovner, E.B., Nørsett, K.G., Liabakk, N.B., Olsen, L.C., Chawla, K., Aas, P.A., and Sætrom, P. (2021). Joint changes in RNA, RNA polymerase II, and promoter activity through the cell cycle identify non-coding RNAs involved in proliferation. Sci. Rep. *11*, 18952.

30. Liao, M., Liao, W., Xu, N., Li, B., Liu, F., Zhang, S., Wang, Y., Wang, S., Zhu, Y., Chen, D., et al. (2019). LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2. EBioMedicine *41*, 200–213.

31. Zhu, Y., Liu, Q., Liao, M., Diao, L., Wu, T., Liao, W., Wang, Z., Li, B., Zhang, S., Wang, S., et al. (2019). Overexpression of lncRNA EPB41L4A-AS1 Induces Metabolic Reprogramming in Trophoblast Cells and Placenta Tissue of Miscarriage. Mol. Ther. Nucleic Acids *18*, 518–532.

32. Liao, W., Xu, N., Zhang, H., Liao, W., Wang, Y., Wang, S., Zhang, S., Jiang, Y., Xie, W., and Zhang, Y. (2022). Persistent high glucose induced EPB41L4A-AS1 inhibits glucose uptake via GCN5 mediating crotonylation and acetylation of histones and non-histones. Clin. Transl. Med. *12*, e699.

33. Bin, J., Nie, S., Tang, Z., Kang, A., Fu, Z., Hu, Y., Liao, Q., Xiong, W., Zhou, Y., Tang, Y., et al. (2021). Long noncoding RNA EPB41L4A-AS1 functions as an oncogene by regulating the Rho/ROCK pathway in colorectal cancer. J. Cell. Physiol. *236*, 523–535.

34. Wang, Z., Liao, W., Liu, F., Yang, T., Xie, W., Liao, M., Gu, D., and Zhang, Y. (2021). Downregulation of lncRNA EPB41L4A-AS1 Mediates Activation of MYD88-Dependent NF-κB Pathway in Diabetes-Related Inflammation. Diabetes Metab. Syndr. Obes. *14*, 265–277.

35. Ishiguro, H., Furukawa, Y., Daigo, Y., Miyoshi, Y., Nagasawa, Y., Nishiwaki, T., Kawasoe, T., Fujita, M., Satoh, S., Miwa, N., et al. (2000). Isolation and characterization of human NBL4, a gene involved in the beta-catenin/tcf signaling pathway. Jpn. J. Cancer Res. *91*, 597–603.

36. Comiskey, D.F., Jr, He, H., Liyanarachchi, S., Sheikh, M.S., Hendrickson, I.V., Yu, L., Brock, P.L., and de la Chapelle, A. (2020). Characterizing the function of EPB41L4A in the predisposition to papillary thyroid carcinoma. Sci. Rep. *10*, 19984.

37. Guo, Y., Christine, K.S., Conlon, F., Gessert, S., and Kühl, M. (2011). Expression analysis of epb41l4a during Xenopus laevis embryogenesis. Dev. Genes Evol. *221*, 113–119.

38. Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.-P., and Li, W. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res *41*, e74.

39. Moyerbrailean, G.A., Richards, A.L., Kurtz, D., Kalita, C.A., Davis, G.O., Harvey, C.T., Alazizi, A., Watza, D., Sorokin, Y., Hauff, N., et al. (2016). High-throughput allele-specific expression across 250 environmental conditions. Genome Res. *26*, 1627–1638.

40. Fafard-Couture, É., Bergeron, D., Couture, S., Abou-Elela, S., and Scott, M.S. (2021). Annotation of snoRNA abundance across human tissues reveals complex snoRNA-host gene relationships. Genome Biol *22*, 172.

41. Schwartzman, O., Mukamel, Z., Oded-Elkayam, N., Olivares-Chauvet, P., Lubling, Y., Landan, G., Izraeli, S., and Tanay, A. (2016). UMI-4C for quantitative and targeted chromosomal contact profiling. Nat. Methods *13*, 685–691.

42. Vian, L., Pękowska, A., Rao, S.S.P., Kieffer-Kwon, K.-R., Jung, S., Baranello, L., Huang, S.-C., El Khattabi, L., Dose, M., Pruett, N., et al. (2018). The energetics and physiological impact of cohesin extrusion. Cell *175*, 292–294.

43. Salnikov, P., Korablev, A., Serova, I., Belokopytova, P., Yan, A., Stepanchuk, Y., Tikhomirov, S., and Fishman, V. (2024). Structural variants in the Epb41l4a locus: TAD disruption and Nrep gene misregulation as hypothetical drivers of neurodevelopmental outcomes. Sci Rep *14*, 5288.

44. Krietenstein, N., Abraham, S., Venev, S.V., Abdennur, N., Gibcus, J., Hsieh, T.-H.S., Parsi, K.M., Yang, L., Maehr, R., Mirny, L.A., et al. (2020). Ultrastructural Details of Mammalian Chromosome Architecture. Mol. Cell *78*, 554–565.e7.

45. Liang, X.-H., Vickers, T.A., Guo, S., and Crooke, S.T. (2011). Efficient and specific knockdown of small non-coding RNAs in mammalian cells and in mice. Nucleic Acids Res. *39*, e13.

46. Hagedorn, P.H., Persson, R., Funder, E.D., Albæk, N., Diemer, S.L., Hansen, D.J., Møller, M.R., Papargyri, N., Christiansen, H., Hansen, B.R., et al. (2018). Locked nucleic acid: modality, diversity, and drug discovery. Drug Discov. Today *23*, 101–114.

47. Dhaka, B., Zimmerli, M., Hanhart, D., Moser, M.B., Guillen-Ramirez, H., Mishra, S., Esposito, R., Polidori, T., Widmer, M., García-Pérez, R., et al. (2024). Functional identification of cis-regulatory long noncoding RNAs at controlled false discovery rates. Nucleic Acids Res. *52*, 2821–2835.

48. Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.-Y., Cody, N.A.L., Dominguez, D., et al. (2020). A large-scale binding and functional map of human RNA-binding proteins. Nature *583*, 711–719.

49. Garavís, M., and Calvo, O. (2017). Sub1/PC4, a multifaceted factor: from transcription to genome stability. Curr. Genet. *63*, 1023–1035.

50. Conesa, C., and Acker, J. (2010). Sub1/PC4 a chromatin associated protein with multiple functions in transcription. RNA Biol. *7*, 287–290.

51. Das, C., Hizume, K., Batta, K., Kumar, B.R.P., Gadad, S.S., Ganguly, S., Lorain, S., Verreault, A., Sadhale, P.P., Takeyasu, K., et al. (2006). Transcriptional coactivator PC4, a chromatin-associated protein, induces chromatin condensation. Mol. Cell. Biol. *26*, 8303–8315.

52. Hou, Y., Gan, T., Fang, T., Zhao, Y., Luo, Q., Liu, X., Qi, L., Zhang, Y., Jia, F., Han, J., et al. (2022). G-quadruplex inducer/stabilizer pyridostatin targets SUB1 to promote cytotoxicity of a transplatinum complex. Nucleic Acids Res. *50*, 3070–3082.

53. Dubois, J.-C., Bonnell, E., Filion, A., Frion, J., Zimmer, S., Riaz Khan, M., Teplitz, G.M., Casimir, L., Méthot, É., Marois, I., et al. (2025). The single-stranded DNA-binding factor SUB1/PC4 alleviates replication stress at telomeres and is a vulnerability of ALT cancer cells. Proc Natl Acad Sci U S A *122*, e2419712122.

54. Salgado, S., Abreu, P.L., Moleirinho, B., Guedes, D.S., Larcombe, L., and Azzalin, C.M. (2024). Human PC4 supports telomere stability and viability in cells utilizing the alternative lengthening of

telomeres mechanism. EMBO Rep *25*, 5294–5315.

55. Nachmani, D., Bothmer, A.H., Grisendi, S., Mele, A., Bothmer, D., Lee, J.D., Monteleone, E., Cheng, K., Zhang, Y., Bester, A.C., et al. (2019). Germline NPM1 mutations lead to altered rRNA 2'-O-methylation and cause dyskeratosis congenita. Nat. Genet. *51*, 1518–1529.

56. Mourtada-Maarabouni, M., Hedge, V.L., Kirkham, L., Farzaneh, F., and Williams, G.T. (2008). Growth arrest in human T-cells is controlled by the non-coding RNA growth-arrest-specific transcript 5 (GAS5). J. Cell Sci. *121*, 939–946.

57. Schneider, C., King, R.M., and Philipson, L. (1988). Genes specifically expressed at growth arrest of mammalian cells. Cell *54*, 787–793.

58. Mourtada-Maarabouni, M., Pickard, M.R., Hedge, V.L., Farzaneh, F., and Williams, G.T. (2009). GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene *28*, 195–208.

59. Yacqub-Usman, K., Pickard, M.R., and Williams, G.T. (2015). Reciprocal regulation of GAS5 lncRNA levels and mTOR inhibitor action in prostate cancer cells. Prostate *75*, 693–705.

60. Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci. Signal. *3*, ra8.

61. Zhang, Z., Zhu, Z., Watabe, K., Zhang, X., Bai, C., Xu, M., Wu, F., and Mo, Y.-Y. (2013). Negative regulation of lncRNA GAS5 by miR-21. Cell Death Differ. *20*, 1558–1568.

62. Hu, G., Lou, Z., and Gupta, M. (2014). The long non-coding RNA GAS5 cooperates with the eukaryotic translation initiation factor 4E to regulate c-Myc translation. PLoS One *9*, e107016.

63. Sun, D., Yu, Z., Fang, X., Liu, M., Pu, Y., Shao, Q., Wang, D., Zhao, X., Huang, A., Xiang, Z., et al. (2017). LncRNA GAS5 inhibits microglial M2 polarization and exacerbates demyelination. EMBO Rep. *18*, 1801–1816.

64. He, X., Chen, X., Zhang, X., Duan, X., Pan, T., Hu, Q., Zhang, Y., Zhong, F., Liu, J., Zhang, H., et al. (2015). An Lnc RNA (GAS5)/SnoRNA-derived piRNA induces activation of TRAIL gene by site-specifically recruiting MLL/COMPASS-like complexes. Nucleic Acids Res. *43*, 3712–3725.

65. Potapova, T.A., Unruh, J.R., Conkright-Fincham, J., Banks, C.A.S., Florens, L., Schneider, D.A., and Gerton, J.L. (2023). Distinct states of nucleolar stress induced by anticancer drugs. https://doi.org/10.7554/eLife.88799.

66. Yang, K., Wang, M., Zhao, Y., Sun, X., Yang, Y., Li, X., Zhou, A., Chu, H., Zhou, H., Xu, J., et al. (2016). A redox mechanism underlying nucleolar stress sensing by nucleophosmin. Nat. Commun. *7*, 13599.

67. Haddach, M., Schwaebe, M.K., Michaux, J., Nagasawa, J., O'Brien, S.E., Whitten, J.P., Pierre, F., Kerdoncuff, P., Darjania, L., Stansfield, R., et al. (2012). Discovery of CX-5461, the first direct and selective inhibitor of RNA polymerase I, for cancer therapeutics. ACS Med. Chem. Lett. *3*, 602–606.

68. Xu, H., Di Antonio, M., McKinney, S., Mathew, V., Ho, B., O'Neil, N.J., Santos, N.D., Silvester, J., Wei, V., Garcia, J., et al. (2017). CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. Nat. Commun. *8*, 14432.

69. Mars, J.-C., Tremblay, M.G., Valere, M., Sibai, D.S., Sabourin-Felix, M., Lessard, F., and Moss, T. (2020). The chemotherapeutic agent CX-5461 irreversibly blocks RNA polymerase I initiation and promoter release to cause nucleolar disruption, DNA damage and cell inviability. NAR Cancer *2*, zcaa032.

70. Hilton, J., Gelmon, K., Bedard, P.L., Tu, D., Xu, H., Tinker, A.V., Goodwin, R., Laurie, S.A., Jonker, D., Hansen, A.R., et al. (2022). Results of the phase I CCTG IND.231 trial of CX-5461 in patients with advanced solid tumors enriched for DNA-repair deficiencies. Nat. Commun. *13*, 3607.

71. Herzog, V.A., Reichholf, B., Neumann, T., Rescheneder, P., Bhat, P., Burkard, T.R., Wlotzka, W., von Haeseler, A., Zuber, J., and Ameres, S.L. (2017). Thiol-linked alkylation of RNA to assess expression dynamics. Nat Methods *14*, 1198–1204.

72. Open2C, Abdennur, N., Abraham, S., Fudenberg, G., Flyamer, I.M., Galitsyna, A.A., Goloborodko, A., Imakaev, M., Oksuz, B.A., Venev, S.V., et al. (2024). Cooltools: Enabling high-resolution Hi-C analysis in Python. PLoS Comput. Biol. *20*, e1012067.

73. Lee, J.-S., and Mendell, J.T. (2020). Antisense-Mediated Transcript Knockdown Triggers Premature Transcription Termination. Mol. Cell *77*, 1044–1054.e3.

74. Lai, F., Damle, S.S., Ling, K.K., and Rigo, F. (2020). Directed RNase H Cleavage of Nascent Transcripts Causes Transcription Termination. Mol. Cell *77*, 1032–1043.e4.

75. Yu, L., Ma, H., Ji, X., and Volkert, M.R. (2016). The Sub1 nuclear protein protects DNA from oxidative damage. Mol. Cell. Biochem. *412*, 165–171.

76. Mortusewicz, O., Roth, W., Li, N., Cardoso, M.C., Meisterernst, M., and Leonhardt, H. (2008). Recruitment of RNA polymerase II cofactor PC4 to DNA damage sites. J. Cell Biol. *183*, 769–776.

77. Tafforeau, L., Zorbas, C., Langhendries, J.-L., Mullineux, S.-T., Stamatopoulou, V., Mullier, R., Wacheul, L., and Lafontaine, D.L.J. (2013). The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. Mol. Cell *51*, 539–551.

78. Wang, M., Zheng, S., Li, X., Ding, Y., Zhang, M., Lin, L., Xu, H., Cheng, Y., Zhang, X., Xu, H., et al. (2020). Integrated analysis of lncRNA-miRNA-mRNA ceRNA network identified lncRNA EPB41L4A-AS1 as a potential biomarker in non-small cell lung cancer. Front. Genet. *11*, 511676.

79. Yang, T., Wang, Y., Liao, W., Zhang, S., Wang, S., Xu, N., Xie, W., Luo, C., Wang, Y., Wang, Z., et al. (2021). Down-regulation of EPB41L4A-AS1 mediated the brain aging and neurodegenerative diseases via damaging synthesis of NAD+ and ATP. Cell Biosci. *11*, 192.

80. Wang, Z., Wang, R., Niu, L., Zhou, X., Han, J., and Li, K. (2024). EPB41L4A-AS1 is required to maintain basal autophagy to modulates Aβ clearance. NPJ Aging *10*, 24.

81. Samdal, H., Hegre, S.A., Chawla, K., Liabakk, N.-B., Aas, P.A., Sporsheim, B., and Sætrom, P. (2021). The lncRNA EPB41L4A-AS1 regulates gene expression in the nucleus and exerts cell type-dependent effects on cell cycle progression. bioRxiv, 2021.02.10.430566. https://doi.org/10.1101/2021.02.10.430566.

82. Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583–588.

83. Suarez-Arnedo, A., Torres Figueroa, F., Clavijo, C., Arbeláez, P., Cruz, J.C., and Muñoz-Camargo, C. (2020). An image J plugin for the high throughput image analysis of in vitro scratch wound healing assays. PLoS One *15*, e0232565.

84. Gagnon, K.T., Li, L., Janowski, B.A., and Corey, D.R. (2014). Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. Nat. Protoc. *9*, 2045–2060.

85. Sundararaman, B., Zhan, L., Blue, S.M., Stanton, R., Elkins, K., Olson, S., Wei, X., Van Nostrand, E.L., Pratt, G.A., Huelga, S.C., et al. (2016). Resources for the Comprehensive Discovery of Functional RNA Elements. Mol. Cell *61*, 903–913.

86. Meers, M.P., Bryson, T.D., Henikoff, J.G., and Henikoff, S. (2019). Improved CUT&RUN chromatin profiling tools. Elife *8*. https://doi.org/10.7554/eLife.46314.

87. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 10.

88. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

89. Love, M., Anders, S., and Huber, W. (2014). Differential analysis of count data--the DESeq2 package. Genome Biol. *15*, 550.

90. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb) *2*, 100141.

91. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

92. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

93. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

94. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160–W165.

95. Zuckerman, B., Ron, M., Mikl, M., Segal, E., and Ulitsky, I. (2020). Gene Architecture and Sequence Composition Underpin Selective Dependency of Nuclear Export of Long RNAs on NXF1 and the TREX Complex. Mol Cell *79*, 251–267.e6.

96. Jürges, C., Dölken, L., and Erhard, F. (2018). Dissecting newly transcribed and old RNA using GRAND-SLAM. Bioinformatics *34*, i218–i226.

97. Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res. *47*. https://doi.org/10.1093/nar/gkz430.

1. Statello, L., Guo, C.-J., Chen, L.-L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. Nat. Rev. Mol. Cell Biol. *22*, 96–118.

2. Wu, H., Yang, L., and Chen, L.-L. (2017). The Diversity of Long Noncoding RNAs and Their Generation. Trends Genet. *33*, 540–552.

3. Ransohoff, J.D., Wei, Y., and Khavari, P.A. (2018). The functions and unique features of long intergenic non-coding RNA. Nat. Rev. Mol. Cell Biol. *19*, 143–157.

4. Mattick, J.S., Amaral, P.P., Carninci, P., Carpenter, S., Chang, H.Y., Chen, L.-L., Chen, R., Dean, C., Dinger, M.E., Fitzgerald, K.A., et al. (2023). Long non-coding RNAs: definitions, functions, challenges and recommendations. Nat. Rev. Mol. Cell Biol. *24*, 430–447.

5. Hangauer, M.J., Vaughn, I.W., and McManus, M.T. (2013). Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs. PLoS Genet. *9*, e1003569.

6. Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M.C., Maeda, N., Oyama, R., Ravasi, T., Lenhard, B., Wells, C., et al. (2005). The transcriptional landscape of the mammalian genome. Science *309*, 1559–1563.

7. ENCODE Project Consortium, Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

8. Kapranov, P., Cheng, J., Dike, S., Nix, D.A., Duttagupta, R., Willingham, A.T., Stadler, P.F., Hertel, J., Hackermüller, J., Hofacker, I.L., et al. (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science *316*, 1484–1488.

9. Mudge, J.M., Carbonell-Sala, S., Diekhans, M., Martinez, J.G., Hunt, T., Jungreis, I., Loveland, J.E., Arnan, C., Barnes, I., Bennett, R., et al. (2024). GENCODE 2025: reference gene annotation for human and mouse. Nucleic Acids Res. https://doi.org/10.1093/nar/gkae1078.

10. Ruiz-Orera, J., Messeguer, X., Subirana, J.A., and Alba, M.M. (2014). Long non-coding RNAs as a source of new peptides. Elife *3*, e03523.

11. Chen, J., Brunner, A.-D., Cogan, J.Z., Nuñez, J.K., Fields, A.P., Adamson, B., Itzhak, D.N., Li, J.Y., Mann, M., Leonetti, M.D., et al. (2020). Pervasive functional translation of noncanonical human open reading frames. Science *367*, 1140–1146.

12. Barczak, W., Carr, S.M., Liu, G., Munro, S., Nicastri, A., Lee, L.N., Hutchings, C., Ternette, N., Klenerman, P., Kanapin, A., et al. (2023). Long non-coding RNA-derived peptides are immunogenic and drive a potent anti-tumour response. Nat. Commun. *14*, 1078.

13. Kesner, J.S., Chen, Z., Shi, P., Aparicio, A.O., Murphy, M.R., Guo, Y., Trehan, A., Lipponen, J.E., Recinos, Y., Myeku, N., et al. (2023). Noncoding translation mitigation. Nature *617*, 395–402.

14. Chothani, S.P., Adami, E., Widjaja, A.A., Langley, S.R., Viswanathan, S., Pua, C.J., Zhihao, N.T., Harmston, N., D'Agostino, G., Whiffin, N., et al. (2022). A high-resolution map of human RNA translation. Mol. Cell *82*, 2885-2899.e8.

15. Housman, G., and Ulitsky, I. (2016). Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs. Biochim. Biophys. Acta *1859*, 31–40.

16. Paralkar, V.R., Taborda, C.C., Huang, P., Yao, Y., Kossenkov, A.V., Prasad, R., Luan, J., Davies, J.O.J., Hughes, J.R., Hardison, R.C., et al. (2016). Unlinking an lncRNA from Its Associated cis Element. Mol. Cell *62*, 104–110.

17. Ali, T., and Grote, P. (2020). Beyond the RNA-dependent function of LncRNA genes. Elife *9*. https://doi.org/10.7554/eLife.60583.

18. Beucher, A., Miguel-Escalada, I., Balboa, D., De Vas, M.G., Maestro, M.A., Garcia-Hurtado, J., Bernal, A., Gonzalez-Franco, R., Vargiu, P., Heyn, H., et al. (2022). The HASTER lncRNA promoter is a cis-acting transcriptional stabilizer of HNF1A. Nat. Cell Biol. *24*, 1528–1540.

19. De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B.K., Muller, H., Ragoussis, J., Wei, C.-L., and Natoli, G. (2010). A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol. *8*, e1000384.

20. Gil, N., and Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. Nat. Rev. Genet. *21*, 102–117.

21. Monziani, A., and Ulitsky, I. (2023). Noncoding snoRNA host genes are a distinct subclass of long noncoding RNAs. Trends Genet. https://doi.org/10.1016/j.tig.2023.09.001.

22. Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., et al. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database *2017*. https://doi.org/10.1093/database/bax028.

23. Sun, M., Gadad, S.S., Kim, D.-S., and Kraus, W.L. (2015). Discovery, Annotation, and Functional Analysis of Long Noncoding RNAs Controlling Cell-Cycle Gene Expression and Proliferation in Breast Cancer Cells. Mol. Cell *59*, 698–711.

24. Nagarajan, S., Rao, S.V., Sutton, J., Cheeseman, D., Dunn, S., Papachristou, E.K., Prada, J.-E.G., Couturier, D.-L., Kumar, S., Kishore, K., et al. (2020). ARID1A influences HDAC1/BRD4 activity, intrinsic proliferative capacity and breast cancer treatment response. Nat. Genet. *52*, 187–197.

25. Janky, R., Verfaillie, A., Imrichová, H., Van de Sande, B., Standaert, L., Christiaens, V.,

Hulselmans, G., Herten, K., Naval Sanchez, M., Potier, D., et al. (2014). iRegulon: from a gene list to a gene regulatory network using large motif and track collections. PLoS Comput. Biol. *10*, e1003731.

26. Babaian, A., Rothe, K., Girodat, D., Minia, I., Djondovic, S., Milek, M., Spencer Miko, S.E., Wieden, H.-J., Landthaler, M., Morin, G.B., et al. (2020). Loss of m1acp3Ψ Ribosomal RNA Modification Is a Major Feature of Cancer. Cell Rep. *31*, 107611.

27. Cheng, Y., Wang, S., Zhang, H., Lee, J.-S., Ni, C., Guo, J., Chen, E., Wang, S., Acharya, A., Chang, T.-C., et al. (2024). A non-canonical role for a small nucleolar RNA in ribosome biogenesis and senescence. Cell *187*, 4770-4789.e23.

28. Hegre, S.A., Samdal, H., Klima, A., Stovner, E.B., Nørsett, K.G., Liabakk, N.B., Olsen, L.C., Chawla, K., Aas, P.A., and Sætrom, P. (2021). Joint changes in RNA, RNA polymerase II, and promoter activity through the cell cycle identify non-coding RNAs involved in proliferation. Sci. Rep. *11*, 18952.

29. Liao, M., Liao, W., Xu, N., Li, B., Liu, F., Zhang, S., Wang, Y., Wang, S., Zhu, Y., Chen, D., et al. (2019). LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by mediating nucleolar translocation of HDAC2. EBioMedicine *41*, 200–213.

30. Zhu, Y., Liu, Q., Liao, M., Diao, L., Wu, T., Liao, W., Wang, Z., Li, B., Zhang, S., Wang, S., et al. (2019). Overexpression of lncRNA EPB41L4A-AS1 Induces Metabolic Reprogramming in Trophoblast Cells and Placenta Tissue of Miscarriage. Mol. Ther. Nucleic Acids *18*, 518–532.

31. Liao, W., Xu, N., Zhang, H., Liao, W., Wang, Y., Wang, S., Zhang, S., Jiang, Y., Xie, W., and Zhang, Y. (2022). Persistent high glucose induced EPB41L4A-AS1 inhibits glucose uptake via GCN5 mediating crotonylation and acetylation of histones and non-histones. Clin. Transl. Med. *12*, e699.

32. Bin, J., Nie, S., Tang, Z., Kang, A., Fu, Z., Hu, Y., Liao, Q., Xiong, W., Zhou, Y., Tang, Y., et al. (2021). Long noncoding RNA EPB41L4A-AS1 functions as an oncogene by regulating the Rho/ROCK pathway in colorectal cancer. J. Cell. Physiol. *236*, 523–535.

33. Wang, Z., Liao, W., Liu, F., Yang, T., Xie, W., Liao, M., Gu, D., and Zhang, Y. (2021). Downregulation of lncRNA EPB41L4A-AS1 Mediates Activation of MYD88-Dependent NF-κB Pathway in Diabetes-Related Inflammation. Diabetes Metab. Syndr. Obes. *14*, 265–277.

34. Ishiguro, H., Furukawa, Y., Daigo, Y., Miyoshi, Y., Nagasawa, Y., Nishiwaki, T., Kawasoe, T., Fujita, M., Satoh, S., Miwa, N., et al. (2000). Isolation and characterization of human NBL4, a gene involved in the beta-catenin/tcf signaling pathway. Jpn. J. Cancer Res. *91*, 597–603.

35. Comiskey, D.F., Jr, He, H., Liyanarachchi, S., Sheikh, M.S., Hendrickson, I.V., Yu, L., Brock, P.L., and de la Chapelle, A. (2020). Characterizing the function of EPB41L4A in the predisposition to papillary thyroid carcinoma. Sci. Rep. *10*, 19984.

36. Guo, Y., Christine, K.S., Conlon, F., Gessert, S., and Kühl, M. (2011). Expression analysis of epb41l4a during Xenopus laevis embryogenesis. Dev. Genes Evol. *221*, 113–119.

37. Wang, L., Park, H.J., Dasari, S., Wang, S., Kocher, J.-P., and Li, W. (2013). CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res *41*, e74.

38. Yabuta, N., Onda, H., Watanabe, M., Yoshioka, N., Nagamori, I., Funatsu, T., Toji, S., Tamai, K., and Nojima, H. (2006). Isolation and characterization of the TIGA genes, whose transcripts are induced by growth arrest. Nucleic Acids Res. *34*, 4878–4892.

39. Moyerbrailean, G.A., Richards, A.L., Kurtz, D., Kalita, C.A., Davis, G.O., Harvey, C.T., Alazizi, A., Watza, D., Sorokin, Y., Hauff, N., et al. (2016). High-throughput allele-specific expression across 250 environmental conditions. Genome Res. *26*, 1627–1638.

40. Fafard-Couture, É., Bergeron, D., Couture, S., Abou-Elela, S., and Scott, M.S. (2021). Annotation of snoRNA abundance across human tissues reveals complex snoRNA-host gene relationships. Genome Biol *22*, 172.

41. Schwartzman, O., Mukamel, Z., Oded-Elkayam, N., Olivares-Chauvet, P., Lubling, Y., Landan, G., Izraeli, S., and Tanay, A. (2016). UMI-4C for quantitative and targeted chromosomal contact profiling. Nat. Methods *13*, 685–691.

42. Vian, L., Pękowska, A., Rao, S.S.P., Kieffer-Kwon, K.-R., Jung, S., Baranello, L., Huang, S.-C., El Khattabi, L., Dose, M., Pruett, N., et al. (2018). The energetics and physiological impact of cohesin extrusion. Cell *175*, 292–294.

43. Salnikov, P., Korablev, A., Serova, I., Belokopytova, P., Yan, A., Stepanchuk, Y., Tikhomirov, S., and Fishman, V. (2024). Structural variants in the Epb41l4a locus: TAD disruption and Nrep gene misregulation as hypothetical drivers of neurodevelopmental outcomes. Sci Rep *14*, 5288.

44. Krietenstein, N., Abraham, S., Venev, S.V., Abdennur, N., Gibcus, J., Hsieh, T.-H.S., Parsi, K.M., Yang, L., Maehr, R., Mirny, L.A., et al. (2020). Ultrastructural Details of Mammalian Chromosome Architecture. Mol. Cell *78*, 554-565.e7.

45. Liang, X.-H., Vickers, T.A., Guo, S., and Crooke, S.T. (2011). Efficient and specific knockdown of small non-coding RNAs in mammalian cells and in mice. Nucleic Acids Res. *39*, e13.

46. Hagedorn, P.H., Persson, R., Funder, E.D., Albæk, N., Diemer, S.L., Hansen, D.J., Møller, M.R., Papargyri, N., Christiansen, H., Hansen, B.R., et al. (2018). Locked nucleic acid: modality, diversity, and drug discovery. Drug Discov. Today *23*, 101–114.

47. Dhaka, B., Zimmerli, M., Hanhart, D., Moser, M.B., Guillen-Ramirez, H., Mishra, S., Esposito, R., Polidori, T., Widmer, M., García-Pérez, R., et al. (2024). Functional identification of cis-regulatory long noncoding RNAs at controlled false discovery rates. Nucleic Acids Res. *52*, 2821–2835.

48. Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue, S.M., Chen, J.-Y., Cody, N.A.L., Dominguez, D., et al. (2020). A large-scale binding and functional map of human RNA-binding proteins. Nature *583*, 711–719.

49. Garavís, M., and Calvo, O. (2017). Sub1/PC4, a multifaceted factor: from transcription to genome stability. Curr. Genet. *63*, 1023–1035.

50. Conesa, C., and Acker, J. (2010). Sub1/PC4 a chromatin associated protein with multiple functions in transcription. RNA Biol. *7*, 287–290.

51. Das, C., Hizume, K., Batta, K., Kumar, B.R.P., Gadad, S.S., Ganguly, S., Lorain, S., Verreault, A., Sadhale, P.P., Takeyasu, K., et al. (2006). Transcriptional coactivator PC4, a chromatin-associated protein, induces chromatin condensation. Mol. Cell. Biol. *26*, 8303–8315.

52. Hou, Y., Gan, T., Fang, T., Zhao, Y., Luo, Q., Liu, X., Qi, L., Zhang, Y., Jia, F., Han, J., et al. (2022). G-quadruplex inducer/stabilizer pyridostatin targets SUB1 to promote cytotoxicity of a transplatinum complex. Nucleic Acids Res. *50*, 3070–3082.

53. Dubois, J.-C., Bonnell, E., Filion, A., Frion, J., Zimmer, S., Riaz Khan, M., Teplitz, G.M., Casimir, L., Méthot, É., Marois, I., et al. (2025). The single-stranded DNA-binding factor SUB1/PC4 alleviates replication stress at telomeres and is a vulnerability of ALT cancer cells. Proc Natl Acad Sci U S A *122*, e2419712122.

54. Salgado, S., Abreu, P.L., Moleirinho, B., Guedes, D.S., Larcombe, L., and Azzalin, C.M. (2024). Human PC4 supports telomere stability and viability in cells utilizing the alternative lengthening of telomeres mechanism. EMBO Rep *25*, 5294–5315.

55. Nachmani, D., Bothmer, A.H., Grisendi, S., Mele, A., Bothmer, D., Lee, J.D., Monteleone, E., Cheng, K., Zhang, Y., Bester, A.C., et al. (2019). Germline NPM1 mutations lead to altered rRNA 2'-O-methylation and cause dyskeratosis congenita. Nat. Genet. *51*, 1518–1529.

56. Mourtada-Maarabouni, M., Hedge, V.L., Kirkham, L., Farzaneh, F., and Williams, G.T. (2008). Growth arrest in human T-cells is controlled by the non-coding RNA growth-arrest-specific transcript 5 (GAS5). J. Cell Sci. *121*, 939–946.

57. Schneider, C., King, R.M., and Philipson, L. (1988). Genes specifically expressed at growth arrest of mammalian cells. Cell *54*, 787–793.

58. Mourtada-Maarabouni, M., Pickard, M.R., Hedge, V.L., Farzaneh, F., and Williams, G.T. (2009). GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene *28*, 195–208.

59. Yacqub-Usman, K., Pickard, M.R., and Williams, G.T. (2015). Reciprocal regulation of GAS5 lncRNA levels and mTOR inhibitor action in prostate cancer cells. Prostate *75*, 693–705.

60. Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Noncoding RNA

gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci. Signal. *3*, ra8.

61. Zhang, Z., Zhu, Z., Watabe, K., Zhang, X., Bai, C., Xu, M., Wu, F., and Mo, Y.-Y. (2013). Negative regulation of lncRNA GAS5 by miR-21. Cell Death Differ. *20*, 1558–1568.

62. Hu, G., Lou, Z., and Gupta, M. (2014). The long non-coding RNA GAS5 cooperates with the eukaryotic translation initiation factor 4E to regulate c-Myc translation. PLoS One *9*, e107016.

63. Sun, D., Yu, Z., Fang, X., Liu, M., Pu, Y., Shao, Q., Wang, D., Zhao, X., Huang, A., Xiang, Z., et al. (2017). LncRNA GAS5 inhibits microglial M2 polarization and exacerbates demyelination. EMBO Rep. *18*, 1801–1816.

64. He, X., Chen, X., Zhang, X., Duan, X., Pan, T., Hu, Q., Zhang, Y., Zhong, F., Liu, J., Zhang, H., et al. (2015). An Lnc RNA (GAS5)/SnoRNA-derived piRNA induces activation of TRAIL gene by site-specifically recruiting MLL/COMPASS-like complexes. Nucleic Acids Res. *43*, 3712–3725.

65. Potapova, T.A., Unruh, J.R., Conkright-Fincham, J., Banks, C.A.S., Florens, L., Schneider, D.A., and Gerton, J.L. (2023). Distinct states of nucleolar stress induced by anticancer drugs. https://doi.org/10.7554/eLife.88799.

66. Yang, K., Wang, M., Zhao, Y., Sun, X., Yang, Y., Li, X., Zhou, A., Chu, H., Zhou, H., Xu, J., et al. (2016). A redox mechanism underlying nucleolar stress sensing by nucleophosmin. Nat. Commun. *7*, 13599.

67. Haddach, M., Schwaebe, M.K., Michaux, J., Nagasawa, J., O'Brien, S.E., Whitten, J.P., Pierre, F., Kerdoncuff, P., Darjania, L., Stansfield, R., et al. (2012). Discovery of CX-5461, the first direct and selective inhibitor of RNA polymerase I, for cancer therapeutics. ACS Med. Chem. Lett. *3*, 602–606.

68. Xu, H., Di Antonio, M., McKinney, S., Mathew, V., Ho, B., O'Neil, N.J., Santos, N.D., Silvester, J., Wei, V., Garcia, J., et al. (2017). CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2 deficient tumours. Nat. Commun. *8*, 14432.

69. Mars, J.-C., Tremblay, M.G., Valere, M., Sibai, D.S., Sabourin-Felix, M., Lessard, F., and Moss, T. (2020). The chemotherapeutic agent CX-5461 irreversibly blocks RNA polymerase I initiation and promoter release to cause nucleolar disruption, DNA damage and cell inviability. NAR Cancer *2*, zcaa032.

70. Hilton, J., Gelmon, K., Bedard, P.L., Tu, D., Xu, H., Tinker, A.V., Goodwin, R., Laurie, S.A., Jonker, D., Hansen, A.R., et al. (2022). Results of the phase I CCTG IND.231 trial of CX-5461 in patients with advanced solid tumors enriched for DNA-repair deficiencies. Nat. Commun. *13*, 3607.

71. Herzog, V.A., Reichholf, B., Neumann, T., Rescheneder, P., Bhat, P., Burkard, T.R., Wlotzka, W., von Haeseler, A., Zuber, J., and Ameres, S.L. (2017). Thiol-linked alkylation of RNA to assess expression dynamics. Nat Methods *14*, 1198–1204.

72. Open2C, Abdennur, N., Abraham, S., Fudenberg, G., Flyamer, I.M., Galitsyna, A.A., Goloborodko, A., Imakaev, M., Oksuz, B.A., Venev, S.V., et al. (2024). Cooltools: Enabling high-resolution Hi-C analysis in Python. PLoS Comput. Biol. *20*, e1012067.

73. Lee, J.-S., and Mendell, J.T. (2020). Antisense-Mediated Transcript Knockdown Triggers Premature Transcription Termination. Mol. Cell *77*, 1044-1054.e3.

74. Lai, F., Damle, S.S., Ling, K.K., and Rigo, F. (2020). Directed RNase H Cleavage of Nascent Transcripts Causes Transcription Termination. Mol. Cell *77*, 1032-1043.e4.

75. Yu, L., Ma, H., Ji, X., and Volkert, M.R. (2016). The Sub1 nuclear protein protects DNA from oxidative damage. Mol. Cell. Biochem. *412*, 165–171.

76. Mortusewicz, O., Roth, W., Li, N., Cardoso, M.C., Meisterernst, M., and Leonhardt, H. (2008). Recruitment of RNA polymerase II cofactor PC4 to DNA damage sites. J. Cell Biol. *183*, 769–776.

77. Tafforeau, L., Zorbas, C., Langhendries, J.-L., Mullineux, S.-T., Stamatopoulou, V., Mullier, R., Wacheul, L., and Lafontaine, D.L.J. (2013). The complexity of human ribosome biogenesis revealed by systematic nucleolar screening of Pre-rRNA processing factors. Mol. Cell *51*, 539–551.

78. Wang, M., Zheng, S., Li, X., Ding, Y., Zhang, M., Lin, L., Xu, H., Cheng, Y., Zhang, X., Xu, H., et al. (2020). Integrated analysis of lncRNA-miRNA-mRNA ceRNA network identified lncRNA EPB41L4A-AS1 as a potential biomarker in non-small cell lung cancer. Front. Genet. *11*, 511676.

79. Yang, T., Wang, Y., Liao, W., Zhang, S., Wang, S., Xu, N., Xie, W., Luo, C., Wang, Y., Wang, Z., et al. (2021). Down-regulation of EPB41L4A-AS1 mediated the brain aging and neurodegenerative diseases via damaging synthesis of NAD+ and ATP. Cell Biosci. *11*, 192.

80. Wang, Z., Wang, R., Niu, L., Zhou, X., Han, J., and Li, K. (2024). EPB41L4A-AS1 is required to maintain basal autophagy to modulates Aβ clearance. NPJ Aging *10*, 24.

81. Samdal, H., Hegre, S.A., Chawla, K., Liabakk, N.-B., Aas, P.A., Sporsheim, B., and Sætrom, P. (2021). The lncRNA EPB41L4A-AS1 regulates gene expression in the nucleus and exerts cell type-dependent effects on cell cycle progression. bioRxiv, 2021.02.10.430566. https://doi.org/10.1101/2021.02.10.430566.

82. Konermann, S., Brigham, M.D., Trevino, A.E., Joung, J., Abudayyeh, O.O., Barcena, C., Hsu, P.D., Habib, N., Gootenberg, J.S., Nishimasu, H., et al. (2015). Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. Nature *517*, 583–588.

83. Monziani, A., Ben-Tov Perry, R., Hezroni, H., and Ulitsky, I. (2025). BAHCC1 promotes gene expression in neuronal cells by antagonizing SIN3A-HDAC1. bioRxiv. https://doi.org/10.1101/2025.01.24.634719.

84. Suarez-Arnedo, A., Torres Figueroa, F., Clavijo, C., Arbeláez, P., Cruz, J.C., and Muñoz-Camargo, C. (2020). An image J plugin for the high throughput image analysis of in vitro

scratch wound healing assays. PLoS One *15*, e0232565.

85. Gagnon, K.T., Li, L., Janowski, B.A., and Corey, D.R. (2014). Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. Nat. Protoc. *9*, 2045–2060.

86. Sundararaman, B., Zhan, L., Blue, S.M., Stanton, R., Elkins, K., Olson, S., Wei, X., Van Nostrand, E.L., Pratt, G.A., Huelga, S.C., et al. (2016). Resources for the Comprehensive Discovery of Functional RNA Elements. Mol. Cell *61*, 903–913.

87. Meers, M.P., Bryson, T.D., Henikoff, J.G., and Henikoff, S. (2019). Improved CUT&RUN chromatin profiling tools. Elife *8*. https://doi.org/10.7554/eLife.46314.

88. Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. *17*, 10.

89. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

90. Love, M., Anders, S., and Huber, W. (2014). Differential analysis of count data--the DESeq2 package. Genome Biol. *15*, 550.

91. Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS *16*, 284–287.

92. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

93. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., et al. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

94. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

95. Ramírez, F., Ryan, D.P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dündar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. *44*, W160-5.

96. Zuckerman, B., Ron, M., Mikl, M., Segal, E., and Ulitsky, I. (2020). Gene Architecture and Sequence Composition Underpin Selective Dependency of Nuclear Export of Long RNAs on NXF1 and the TREX Complex. Mol Cell *79*, 251-267.e6.

97. Jürges, C., Dölken, L., and Erhard, F. (2018). Dissecting newly transcribed and old RNA using GRAND-SLAM. Bioinformatics *34*, i218–i226.

98. Tang, Z., Kang, B., Li, C., Chen, T., and Zhang, Z. (2019). GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res. *47*. https://doi.org/10.1093/nar/gkz430.