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]. However, only a limited number of lncRNAs have well-defined mechanisms of action. Long noncoding RNAs exhibit a wide range of expression levels, driven by varying transcription rates and half-lives. Similarly 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. Interestingly, 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].

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 the installation of RNA modifications on other RNAs, mainly ribosomal RNAs (rRNAs). In mammalian cells, snoRNAs are produced from introns of polyadenylated transcripts, some of which code for proteins, and others are instead 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/snoRNA-host gene acting as a regulator in *cis*, 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 for 72 hours 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* with both GapmeRs 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 known to date—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 well 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 playing unspecified roles throughout development[36].

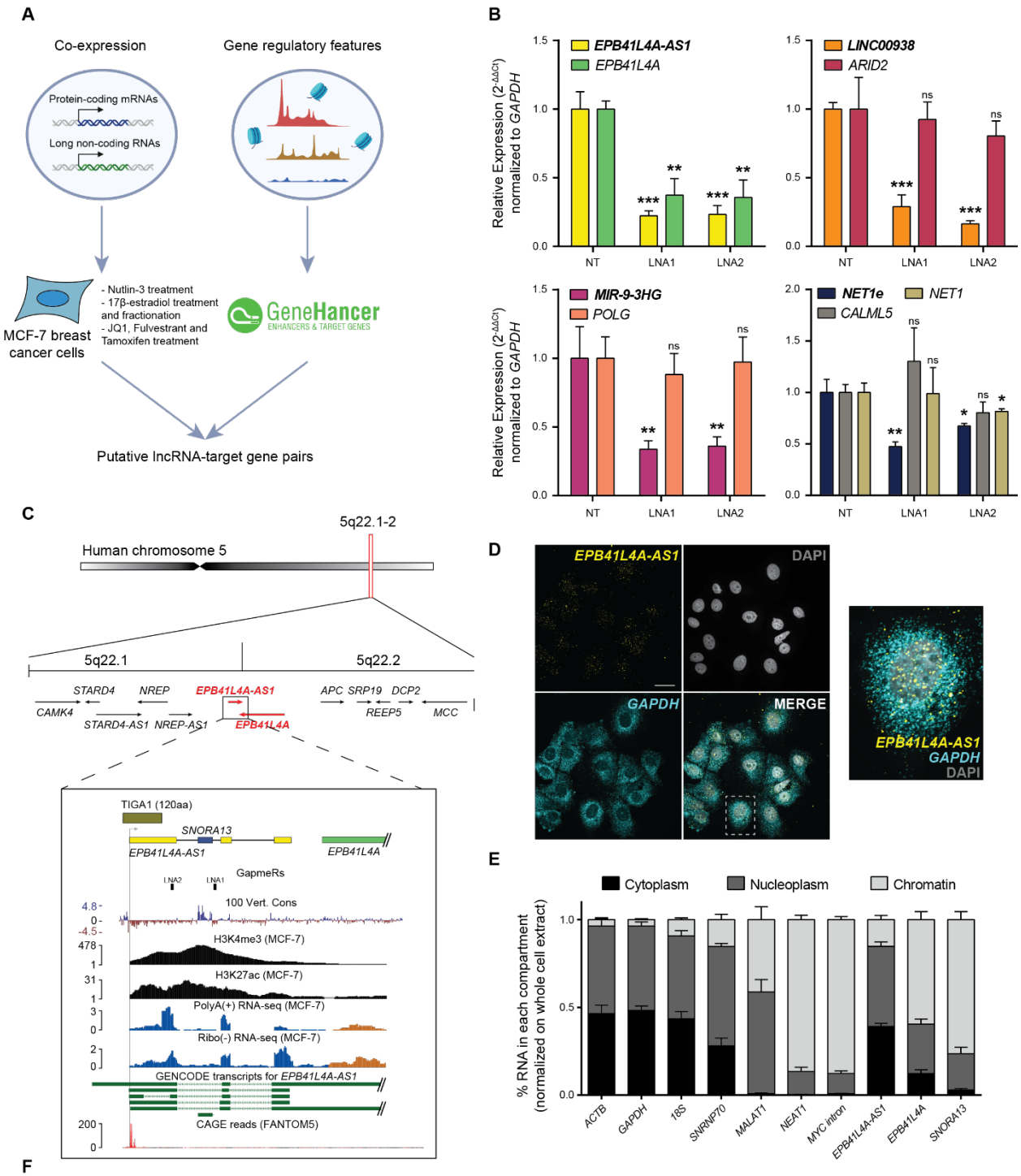

A
Co-expression
Gene regulatory features
Protein-coding mRNAs
Long non-coding RNAs
MCF-7 breast cancer cells
- Nutlin-3 treatment
- 17β-estradiol treatment and fractionation
- JQ1, Fulvestrant and Tamoxifen treatment
GeneHancer
ENHANCERS & TARGET GENES
Putative lncRNA-target gene pairs
B
EPB41L4A-AS1
EPB41L4A
LINC00938
ARID2
MIR-9-3HG
POLG
NET1e
NET1
CALML5
Relative Expression (2^-ΔΔCt) normalized to GAPDH
NT
LNA1
LNA2
ns
C
5q22.1-2
Human chromosome 5
5q22.1
5q22.2
STARD4
NREP
EPB41L4A-AS1
APC
SRP19
DCP2
CAMK4
STARD4-AS1
NREP-AS1
EPB41L4A
REEP5
MCC
TIGA1 (120aa)
SNORA13
GapmeRs
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
EPB41L4A-AS1
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). The location of the GapmeRs used in this study is also reported. **(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* is expressed at an average of 33.37 ± 3.95 molecule per cell, and 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[37–40]. Coding-Potential Assessment (CPAT)[41] 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[42] (**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 therefore 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[42], 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

128 comprising five different cell types exposed to a panel of 50 different substances[43] (**Fig. S3D**),
129 as well as exposing MCF-7 cells to LPS (bacterial infection), H_2O_2 (oxidative stress),
130 thapsigargin (ER stress and unfolded protein response) and etoposide (DNA damage) (**Fig.**
131 **S3E**). Strikingly, *EPB41L4A-AS1* expression was induced in nearly all conditions, often together
132 with *EPB41L4A* but not *SNORA13*, suggesting the levels of the snoRNA are typically detached
133 from that of the lncRNA[44], likely because the latter is substantially more stable (see below).

134 ***EPB41L4A-AS1* regulates *EPB41L4A* transcription**

135 We assessed the 3D chromatin architecture around the *EPB41L4A-AS1* locus by performing
136 UMI-4C-seq[45] with baits targeting the lncRNA, *EPB41L4A* and the promoter of *STARD4*—a
137 flanking gene found ~650 Kb upstream of the *EPB41L4A-AS1* promoter. *EPB41L4A-AS1*
138 formed contacts with both genes (**Fig. 1F**, middle), with prominent peaks corresponding to the
139 promoter of *EPB41L4A*. Interestingly, the promoter of *EPB41L4A* displays several contacts with
140 its own gene body (**Fig. 1F**, right panel), consistent with a "stripe" in HiC data where a single
141 anchor interacts with the whole topologically associated domain (TAD) at high frequency[46], with
142 those frequencies reduced to below background levels upstream of the *EPB41L4A-AS1* locus.
143 Conversely, contacts from the *STARD4* promoter also show a peak near *EPB41L4A-AS1* and
144 then reduced below the background (**Fig. 1F**, left panel). These data indicate that *EPB41L4A-*
145 *AS1* is located at a TAD boundary region, forming strong spatial contacts with sites upstream
146 and downstream, including the promoter of *EPB41L4A*. In agreement with this, *EPB41L4A-AS1*
147 is located between the 5q22.1 and 5q22.2 chromosomal bands, delimiting two TADs as evident
148 in both our 4C-seq and Micro-C data from H1 cells (**Fig. 1F** and **2A**). Both *EPB41L4A* and
149 *EPB41L4A-AS1* are characterized by CTCF binding near their promoters and their chromatin is
150 extensively demarcated with H3K27ac (**Fig. 2A**). A recent study identified this boundary as one
151 of the most insulated in human embryonic stem cells (hESCs)[47].

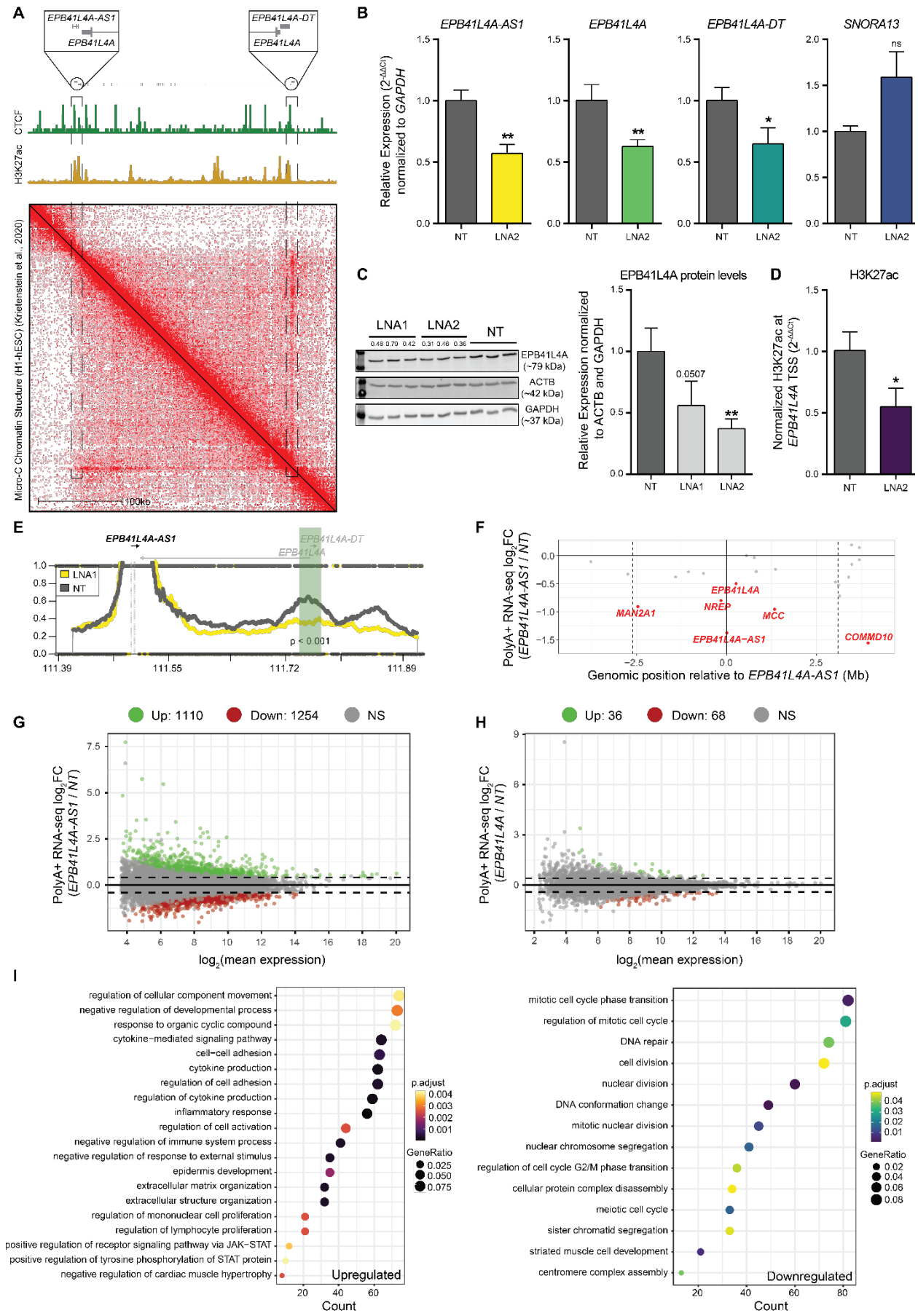

A
EPB41L4A-AS1
EPB41L4A
EPB41L4A-DT
EPB41L4A
CTCF
H3K27ac
Micro-C Chromatin Structure (H1-hESC) (Krietenstein et al., 2020)
100kb
B
EPB41L4A-AS1
EPB41L4A
EPB41L4A-DT
SNORA13
Relative Expression (2^-ΔΔCt) normalized to GAPDH
NT
LNA2
ns
C
LNA1
LNA2
NT
EPB41L4A (~79 kDa)
ACTB (~42 kDa)
GAPDH (~37 kDa)
EPB41L4A protein levels
Relative Expression normalized to ACTB and GAPDH
0.0507
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

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[48] 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 flanking TADs of the lncRNA in cells transfected with GapmeRs targeting *EPB41L4A-AS1*. The vertical dotted lines represent the TAD boundaries (as assessed by TADmap[49]), 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[50,51], 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**). The former suggests that endogenous levels of *EPB41L4A-AS1*—that are far greater than those of *EPB41L4A*—are sufficient to sustain the maximal expression of this target gene in MCF-7 cells. 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 a *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

176 the expression of *EPB41L4A*, possibly by ensuring spatial proximity between the TAD boundary
177 and the *EPB41L4A* promoter.

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

179 To evaluate the global transcriptional response to *EPB41L4A-AS1* levels, we sequenced
180 polyadenylated RNA from cells transfected with *EPB41L4A-AS1*-, *EPB41L4A*-targeting, or
181 control GapmeRs by RNA-seq (**Table S2**). Because LNA1 induced the expression of the stress
182 responsive gene *CDKN1A* (p21)—a known offtarget effect of some GapmeR[52–54] (**Fig. S5A-C**)
183 —we have decided to initially sequence only the cells transfected with LNA2. Consistently with
184 the RT-qPCR data, KD of *EPB41L4A-AS1* reduced *EPB41L4A* expression, and also reduced
185 expression of several, but not all other genes in the TADs flanking the lncRNA (**Fig. 2F**). Based
186 on these data, *EPB41L4A-AS1* is a significant *cis*-acting activator according to TransCistor[55]
187 (P=0.005 using the digital mode). The *cis*-regulated genes reduced by *EPB41L4A-AS1* KD
188 included *NREP*, a gene important for brain development, whose homolog was downregulated
189 by genetic manipulations of regions homologous to the lncRNA locus in mice[47]. Depletion of
190 *EPB41L4A-AS1* thus affects several genes in its vicinity.

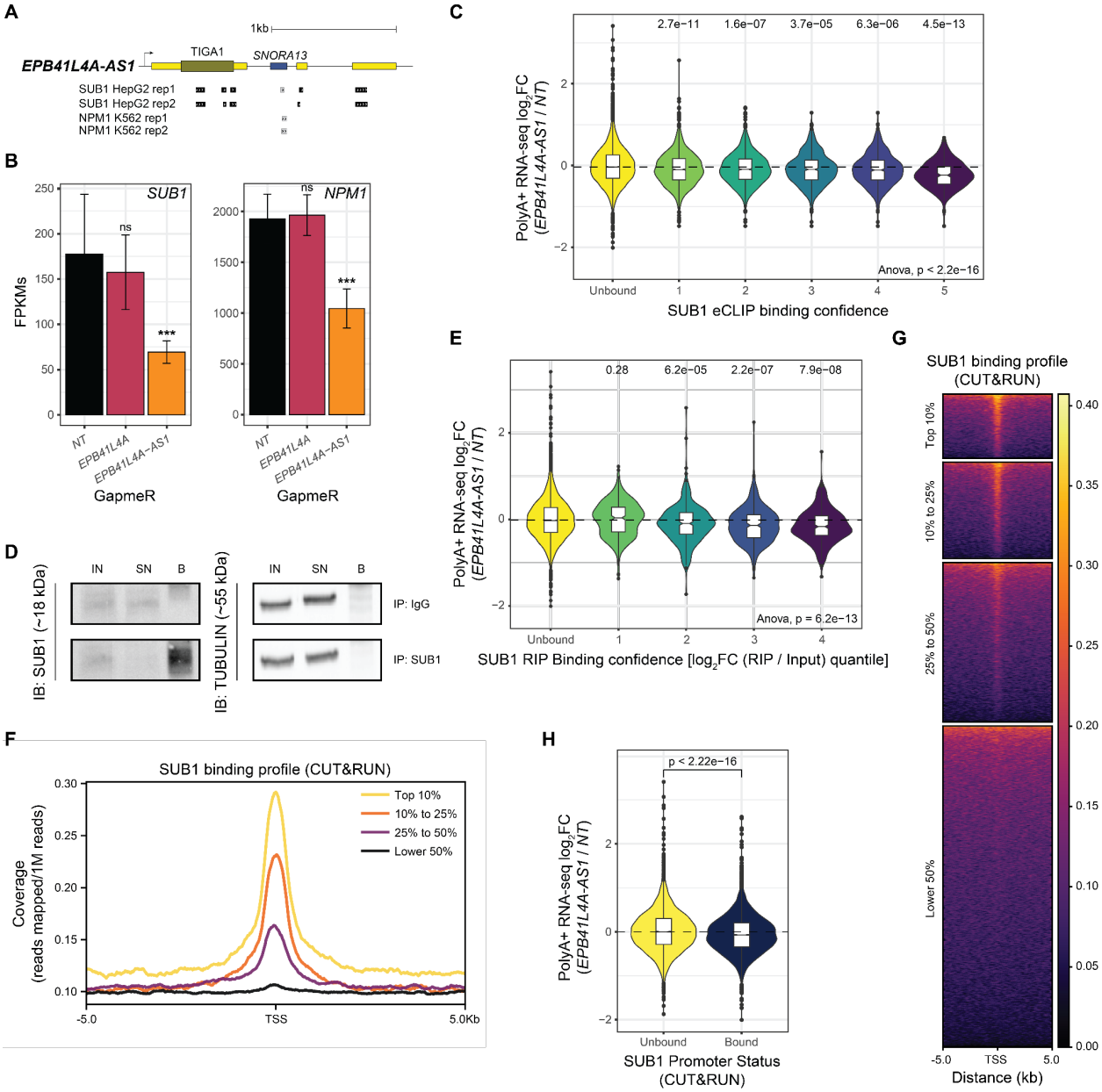
191 Downregulation of *EPB41L4A-AS1* led to broad changes in gene expression beyond the
192 proximal TADs (**Fig. S5D**). A total of 2,364 genes were differentially expressed (1,110 up, 1,254
193 down, adjusted $P<0.05$ and $|\log_2\text{Fold-change}| >0.41$) (**Fig. 2G**), in contrast to only 104 (36 up,
194 68 down) when *EPB41L4A* was targeted (**Fig. 2H**). A total of 25 (69%) of the upregulated, and
195 22 (32%) of the downregulated genes after *EPB41L4A* KD were also regulated in the same
196 direction by the KD of *EPB41L4A-AS1* (**Fig. S5E,** $P<10^{-15}$ for both). In agreement with studies
197 on *EPB41L4A* function[34,35], these genes were enriched with those functioning in cell-cell
198 adhesion and cell surface receptor signaling, although with a borderline significance (**Fig. S5F**).
199 Most of the consequences of the depletion of *EPB41L4A-AS1* are thus not directly explained by
200 changes in *EPB41L4A* levels. An additional *trans*-acting function for *EPB41L4A-AS1* would
201 therefore be consistent with its high expression levels compared to most lncRNAs detected in
202 MCF-7 (**Fig. S5G**). To strengthen these findings, we have transfected MCF-7 cells with LNA1
203 and a second control GapmeR (NT2), as well as the previous one (NT1) and LNA2, and
204 sequenced the polyadenylated RNA fraction as before. Notably, the expression levels (in
205 FPKMs) of the replicates of both control samples are highly correlated with each other (**Fig.
206 S6A**), and the global transcriptomic changes triggered by the two *EPB41L4A-AS1*-targeting
207 LNAs are largely concordant (**Fig. S6B** and **S6C**). Because of this concordance and the cleaner
208 (i.e., no CDKN1A upregulation) readout in LNA2-transfected cells, we focused mainly on these
209 cells for subsequent analyses.

210 Several GO terms were found to be enriched among both up- and down-regulated genes
211 following *EPB41L4A-AS1* KD. Upregulated genes were linked to cell-cell adhesion,
212 inflammatory response and JAK/STAT/ERK pathway, and the downregulated ones to cell cycle,
213 division and DNA repair, among others (**Fig. 2I**). When considering cellular components, the
214 downregulated genes were associated with ribonucleoprotein and spliceosomal complexes, as
215 well as to ribosome subunit and, interestingly, the nucleolus (**Fig. S7A** and **S7B**). Thus, since
216 *EPB41L4A-AS1* is a snoRNA-host gene (SNHG) and snoRNAs typically localize to and act in

217 the nucleolus, we aimed to investigate whether this lncRNA might interact with some factors that
218 can explain the observed enrichment in nucleolar-associated genes.

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

220 In order to identify potential factors that might be associated with *EPB41L4A-AS1*, we inspected
221 protein-RNA binding data from the ENCODE eCLIP dataset[56]. The exons of the *EPB41L4A-AS1*
222 lncRNA were densely and strongly bound by SUB1 (also known as PC4) in both HepG2 and
223 K562 cells (**Fig. 3A**). SUB1 interacts with all three RNA polymerases and was reported to be
224 involved in transcription initiation and elongation, response to DNA damage, chromatin
225 condensation[57–60], telomere maintenance[61,62] and rDNA transcription[63]. SUB1 normally localizes
226 throughout the nucleus in various cell lines, yet staining experiments show a moderate
227 enrichment for the nucleolus (source: Human Protein Atlas;
228 https://www.proteinatlas.org/ENSG00000113387-SUB1/subcellular)[63]. Another nucleolus-
229 related protein, NPM1, which binds C/D-box snoRNAs and directs 2'-O-methylation of rRNAs[64],
230 had a single moderately confident binding site within *EPB41L4A-AS1* intron overlapping
231 *SNORA13* (**Fig. 3A**). Interestingly, both *SUB1* and *NPM1* mRNAs were significantly
232 downregulated following *EPB41L4A-AS1* KD (**Fig. 3B**).



233

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 FPKMs) 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_2FC RIP / Input). **(F)** Metagene profile around TSS of the normalized SUB1 CUT&RUN signal, stratified by gene expression levels in MCF-7 cells. **(G)** Heatmap of SUB1 reads coverage around the TSSs of all human genes, stratified by gene expression levels in MCF-7 cells. **(H)** 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.

234 We hypothesized that loss of *EPB41L4A-AS1* might affect SUB1, either via the reduction in its
235 expression or by affecting its functions. We stratified SUB1 eCLIP targets into confidence
236 intervals, based on the number, strength and confidence of the reported binding sites. Indeed,
237 eCLIP targets of SUB1 (from HepG2 cells profiled by ENCODE) were significantly
238 downregulated following *EPB41L4A-AS1* KD in MCF-7, with more confident targets
239 experiencing stronger downregulation (**Fig. 3C**). Importantly, this still holds true when controlling
240 for gene expression levels (**Fig. S7C**), suggesting that this negative trend is not due to
241 differences in their baseline expression. To obtain SUB1-associated transcripts in MCF-7 cells,
242 we performed a native RNA immunoprecipitation followed by sequencing of polyA+ RNAs (RIP-
243 seq) (**Fig. 3D**, **S7D** and **S7E**). As expected, *EPB41L4A-AS1* precipitated with SUB1, as did
244 other top hits from the eCLIP data (**Fig. S7D**). A total of 1,470 genes were enriched in the IP
245 fraction (adjusted P<0.05 and $|\log_2$Fold-change| >0.41) (**Fig. S7E**), and our MCF-7 RIP data
246 was in agreement with that of the eCLIP in HepG2 cells (**Fig. S7F**). Similarly to what was done
247 with the eCLIP data, we divided the RIP-defined targets into confidence intervals based on the
248 significance and overall enrichment in the IP fraction. Strikingly, SUB1-associated transcripts
249 were downregulated upon *EPB41L4A-AS1* KD (**Fig. 3E**), similar to what was observed for the
250 eCLIP-based targets. Since SUB1 is also a transcriptional regulator associated with chromatin,
251 we performed CUT&RUN to uncover the genome-wide chromatin occupancy of SUB1 in MCF-7
252 cells (**Table S3**). SUB1 binding was found to be enriched around TSSs of active genes (**Fig. 3F**
253 and **3G**), and among its 835 high-confidence genomic peaks, 418 (~50%) overlapped with the
254 promoters of 411 genes. Of these genes, transcripts from only 56 were also bound by SUB1 at
255 the RNA level, suggesting that SUB1 largely targets distinct gene sets at the DNA and RNA
256 levels. Genes with promoter occupancy of SUB1 were enriched for biological processes related
257 to genome maintenance, stress response, and nuclear transport (**Fig. S7G**), which is in line with
258 previous studies[57–60]. These genes were also slightly, yet highly significantly downregulated
259 following *EPB41L4A-AS1* KD (**Fig. 3H**). Taken together, *EPB41L4A-AS1* interacts with SUB1
260 and its KD led to both a decrease in SUB1 mRNA level and the concomitant downregulation of
261 SUB1 targets on chromatin and at the RNA level, positioning this lncRNA as a modulator of
262 SUB1 activity.

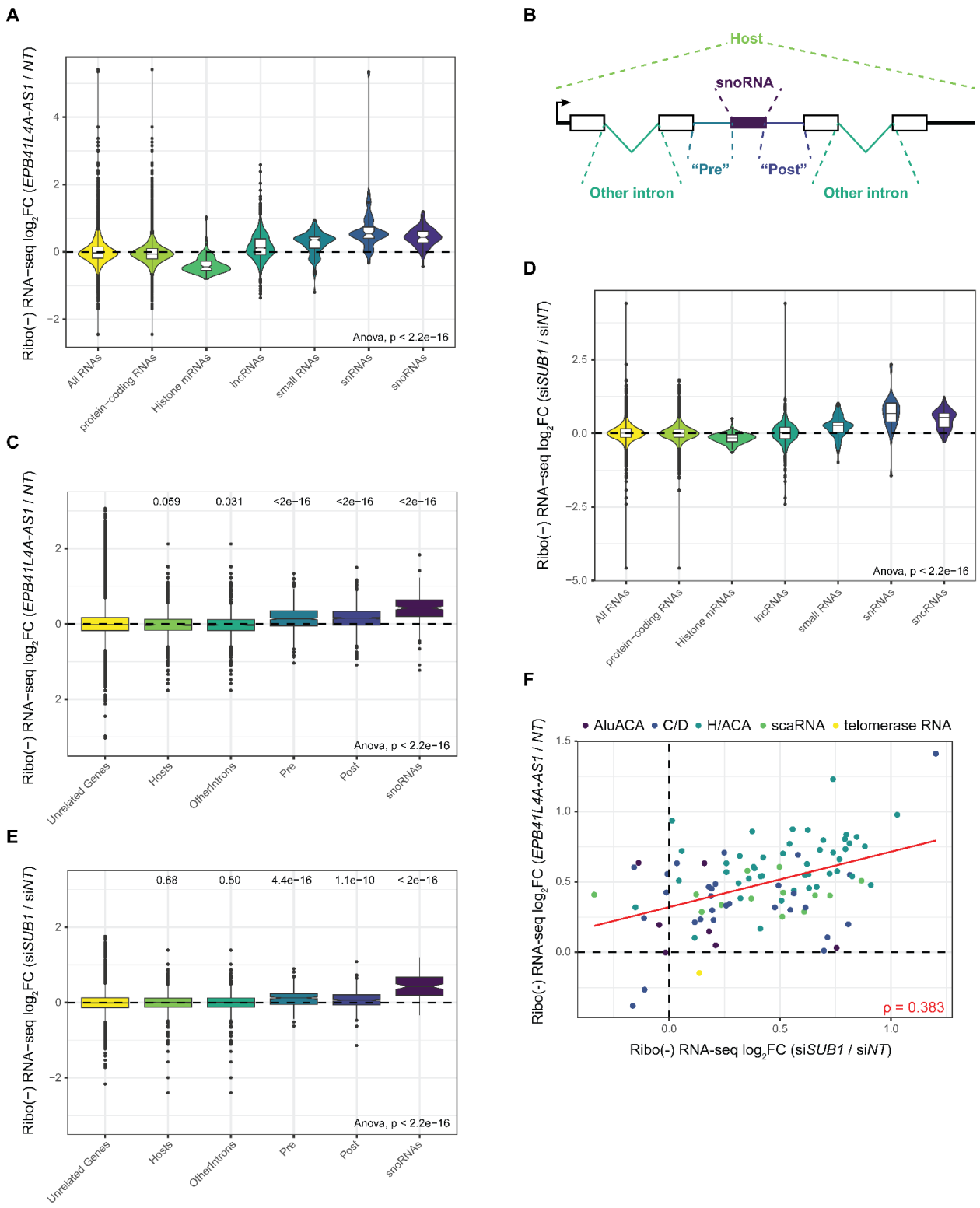
263 **Global de-regulation of snoRNAs and histone mRNA expression upon loss of *EPB41L4A-***
264 ***AS1***

265 Since many classes of noncoding RNAs—including those localizing to the nucleolus, such as
266 snoRNAs—are not polyadenylated, we next sequenced total rRNA-depleted RNAs from
267 *EPB41L4A-AS1* KD and control MCF-7 cells. As expected, using this approach we could
268 quantify various RNA species lacking polyA tails, such as small nuclear RNAs (snRNAs),

269 snoRNAs and histone mRNAs (**Fig. 4A**), and could also better quantify intronic expression.
270 Consistent with a *cis*-acting regulation at the transcriptional level, both the exonic and intronic
271 reads of *EPB41L4A* were significantly reduced after *EPB41L4A-AS1* downregulation (**Fig. S8A**).
272 When considering different transcript classes, we observed a marked upregulation of most
273 snRNAs and snoRNAs and a significant downregulation of histone mRNAs (**Fig. 4A**). To
274 quantity SNHG de-regulation at higher resolution, we created a custom annotation in which we
275 divided each SNHG into exonic regions, the snoRNA-encoding region, regions in the snoRNA-
276 containing introns that are found upstream ('pre-') and downstream ('post-') of the snoRNA, and
277 other snoRNA-less introns ('Other') of the SNHG transcripts, while considering the various
278 isoforms annotated for each SNHG (see Methods) (**Fig. 4B**). We then quantified the levels of
279 each region in control and KD conditions. Intriguingly, while SNHG exons and unrelated introns
280 were not affected by *EPB41L4A-AS1* KD, read coverage was increased over both the snoRNAs
281 and the flanking intronic regions (**Fig. 4C**). These changes were shared among all the snoRNA
282 classes (**Fig. S8B**) and were concordant within individual introns when both (pre- and post-) had
283 enough coverage to be quantified (**Fig. S8C**). Notably, the effect sizes of these changes were
284 small, and they were not evident when plotting metagenes of these regions—that show the
285 median coverage level—which is close to zero even in rRNA-depleted data (**Fig. S8D**),
286 therefore suggesting that snoRNA processing remains an efficient process even in cells
287 depleted of *EPB41L4A-AS1*. Therefore, most of the effect we observe is likely in the increased
288 accumulation of the mature and stable snoRNAs, eventually hinting that *EPB41L4A-AS1*
289 perturbations specifically affect the snoRNA production and/or stability rather than SNHG
290 transcription.

291 To test if the effects on snoRNA biogenesis are related to the changes in SUB1 expression
292 and/or activity, we sequenced the rRNA-depleted RNA from MCF-7 cells in which *SUB1* was
293 depleted by siRNAs (**Fig. S8E** and **S8F**). We identified a total of 243 upregulated and 154
294 downregulated genes (**Fig. S8G**), of which 41 (~17%) and 74 (~48%) were concordantly
295 affected by *EPB41L4A-AS1* depletion, and the magnitudes of the genome-wide responses were
296 moderately, but significantly correlated (Spearman's $R = 0.138$, $P<10^{-15}$). Strikingly, *SUB1* KD
297 led to a similar increase in snoRNA levels and a decrease in histone mRNAs (**Fig. 4D, 4E** and
298 **S8H**). Interestingly, the changes in snoRNAs expression were highly concordant with those
299 driven by the KD of *EPB41L4A-AS1* (**Fig. 4F**, $R = 0.383$) and intronic regions flanking the
300 snoRNAs were also increased, suggesting that *EPB41L4A-AS1* and SUB1 affect the
301 abundance of the same pool of snoRNAs in MCF-7 cells through likely a shared mechanism.

302 Lastly, we tested whether the effects of *EPB41L4A-AS1* perturbation could be a generic
303 outcome of a noncoding SNHG (ncSNHG) perturbation. To this end, we used GapmeRs to
304 deplete *GAS5*, another abundant SNHG which encodes 11 snoRNAs within its introns (**Fig.**
305 **S9A** and **S9B**). As expected from the various functions associated with this ncSNHG[65–73], *GAS5*
306 KD led to substantial changes in gene expression (**Fig. S9C**), including a marked suppression
307 of histone mRNA transcription, in line with its anti-proliferative function (**Fig. S9D**). However,
308 snoRNAs or their host introns were largely unaffected (**Fig. S9D-F**), suggesting that the effect of
309 *EPB41L4A-AS1* on snoRNA expression is not a common feature of ncSNHGs.



310

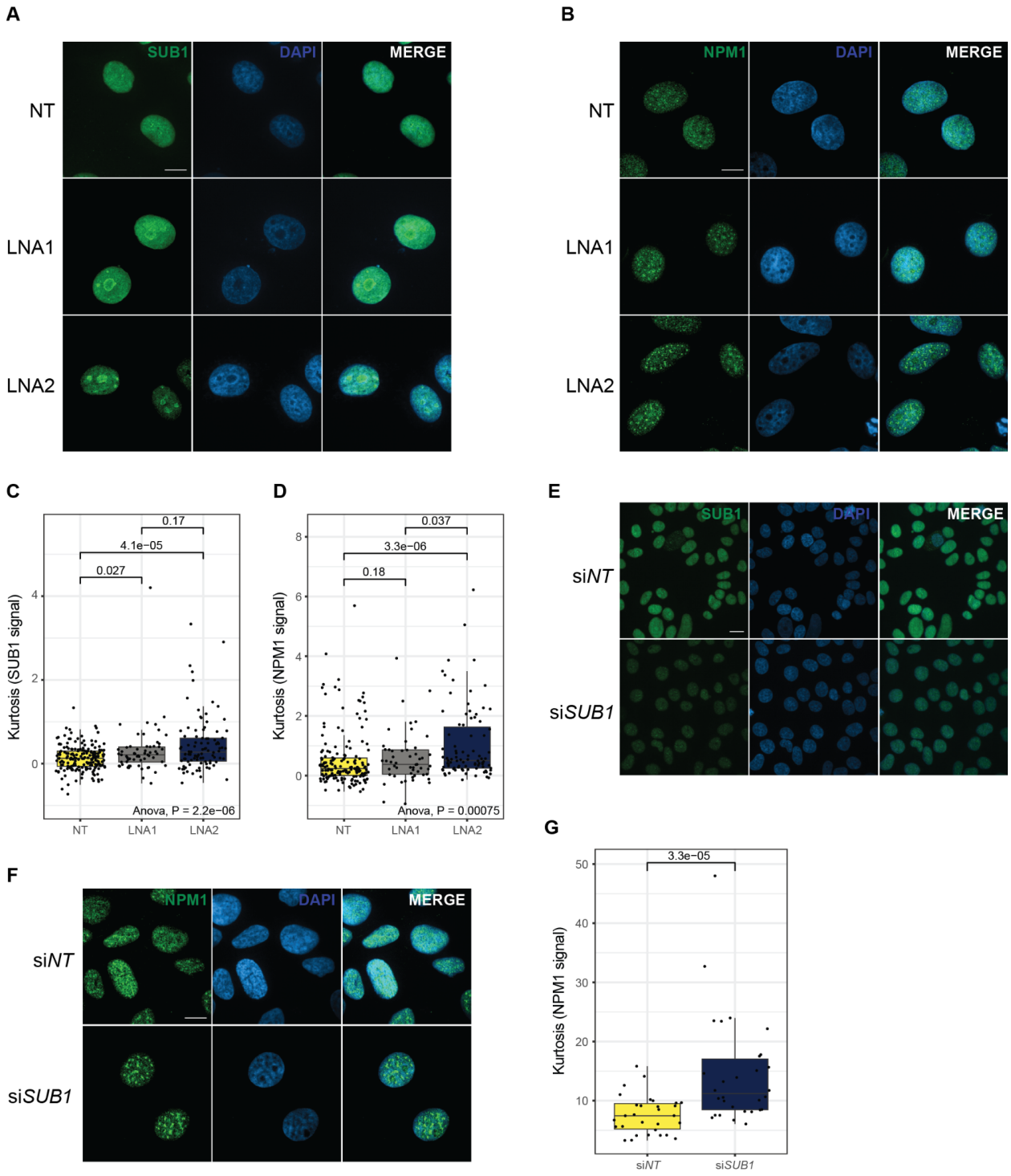
Figure 4. *EPB41L4A-AS1* and SUB1 depletion results in a widespread accumulation of mature snoRNAs. **(A)** Changes in gene expression upon *EPB41L4A-AS1* KD with GapmeRs of the indicated RNA classes. **(B)** Schematics depicting the different regions which were separately quantified in each

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.

311

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

313 The results described so far point to *EPB41L4A-AS1* being important for SUB1 function. To test
314 whether *EPB41L4A-AS1* is important for its proper intracellular distribution, we immunostained
315 for SUB1 and NPM1 in MCF-7 cells transfected with either control or *EPB41L4A-AS1*-targeting
316 GapmeRs. A substantial change in subcellular localization could be observed for both proteins
317 (**Fig. 5A** and **5B**). Specifically, SUB1 normally displays a pan-nuclear localization, sometimes
318 slightly enriching into DAPI-low areas corresponding to the nucleoli (**Fig. 5A** and **5C**). Upon
319 *EPB41L4A-AS1* depletion, SUB1 shifts towards a stronger nucleolar localization while retaining
320 a presence throughout the nucleus. NPM1 instead regularly shows enrichment inside the
321 nucleoli; however, in GapmeR-transfected cells, this is accompanied by the appearance of
322 globular aggregates within the nucleus (**Fig. 5B** and **5D**), reminiscent of the previously reported
323 nucleolar damage/stress response[74,75].



324

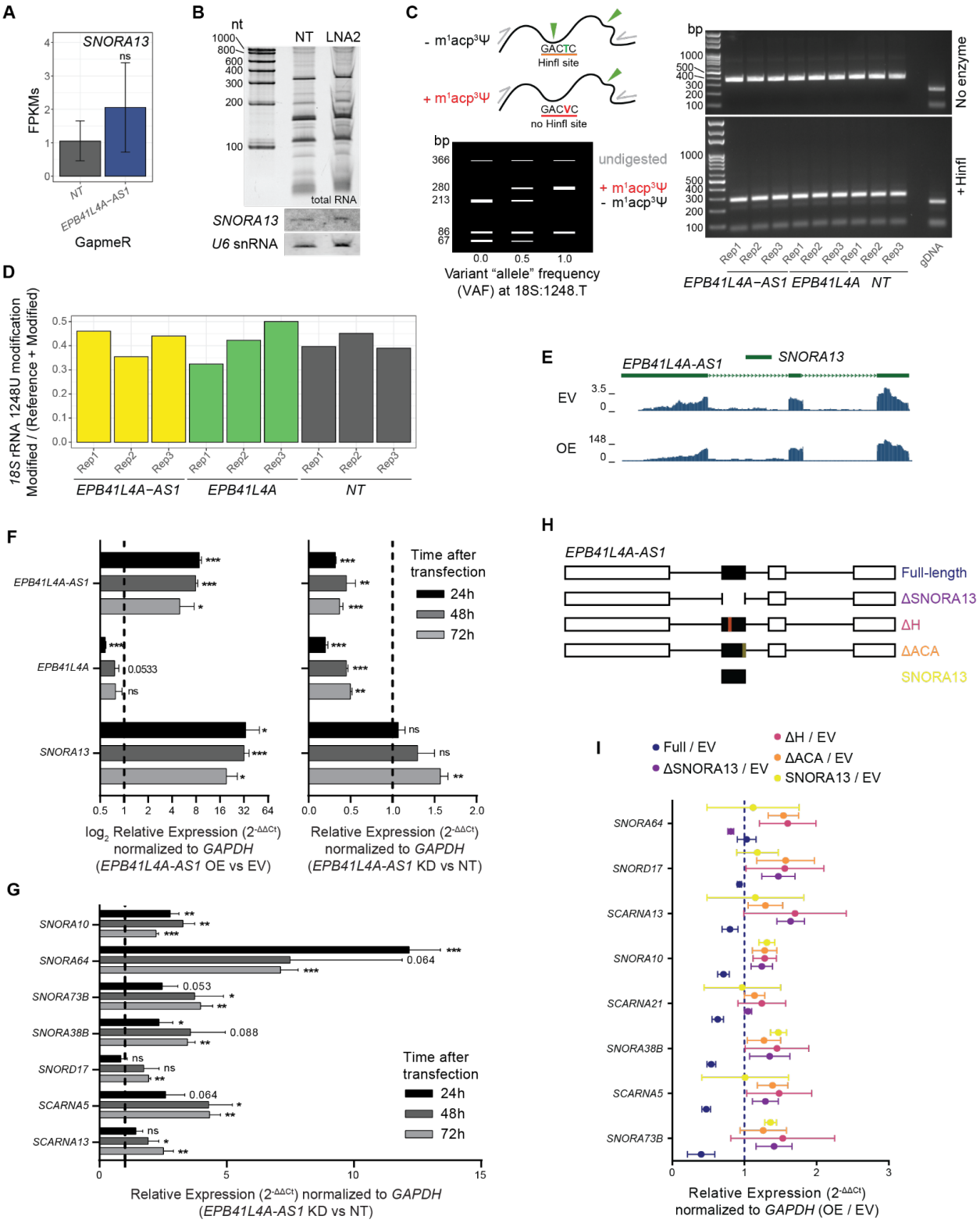
Figure 5. Loss of *EPB41L4A-AS1* results in an altered nuclear patterning of SUB1 and NPM1. **(A)** Representative immunofluorescence images for SUB1 after *EPB41L4A-AS1* depletion with two distinct GapmeRs (scale bar = 20 µm, 100X magnification). **(B)** Same as in (A), but for NPM1. **(C)** Quantification 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[76–79]. 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. S10A** and **S10B**). SUB1 localization also changed, but only at the last time point (24hrs) (**Fig. S10C** and **S10D**). 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 significantly 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 the 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 the MACP modification is not significantly 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.



356

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. *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 then 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

382 any rescue of the snoRNAs abundance (**Fig. 6I**). This suggests that the entire *EPB41L4A-AS1*
383 transcript is required to counteract *EPB41L4A-AS1* knockdown, most likely because the intact
384 *SNORA13* sequence is necessary to undergo the characteristic processing steps of a SNHG.

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

386 To distinguish between the effects of *EPB41L4A-AS1* depletion on RNA synthesis and RNA
387 decay, we performed SLAM-seq[80] using ribo-depleted total RNA in *EPB41L4A-AS1* KD cells
388 (**Table S4**). Following 48 hours after control or *EPB41L4A-AS1*–targeting GapmeRs
389 transfection, we collected RNA at 0, 4, 8 and 24 hours post-addition of 4sU to the media, while
390 replenishing fresh 4sU at each time point (see Methods) (**Fig. 7A**). *EPB41L4A* showed
391 decreased synthesis rate and unchanged half-life (**Fig. 7B**), as expected from a gene regulated
392 at the level of transcription. As a group, genes upregulated after *EPB41L4A-AS1* KD in the
393 regular steady-state RNA-seq data had substantially higher synthesis rates, and those
394 downregulated had lower synthesis rates, when compared to unchanged ones ($P=3.25x10^{-14}$ for
395 the up-regulated genes and P=0.002 for the down-regulated genes, Mann-Whitney test, **Fig.**
396 **7C**), whereas only a minor difference was observed in terms of half-lives (P=0.011 for the up-
397 regulated genes and $P=1.82x10^{-8}$ for the down-regulated genes, Mann-Whitney test, **Fig. 7D**).
398 This suggests that the *trans*-regulated genes are also mostly regulated at the level of their
399 transcription. Additionally, when comparing the effects across genes in the two neighboring
400 TADs (“Cis”), other TADs on the same chromosome (“Chr5”), or other chromosomes (“Trans”),
401 the genes proximal to *EPB41L4A-AS1* specifically exhibit a reduced synthesis rate compared to
402 the others (**Fig. S11A**), but no changes in their half-lives (**Fig. S11B**).

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

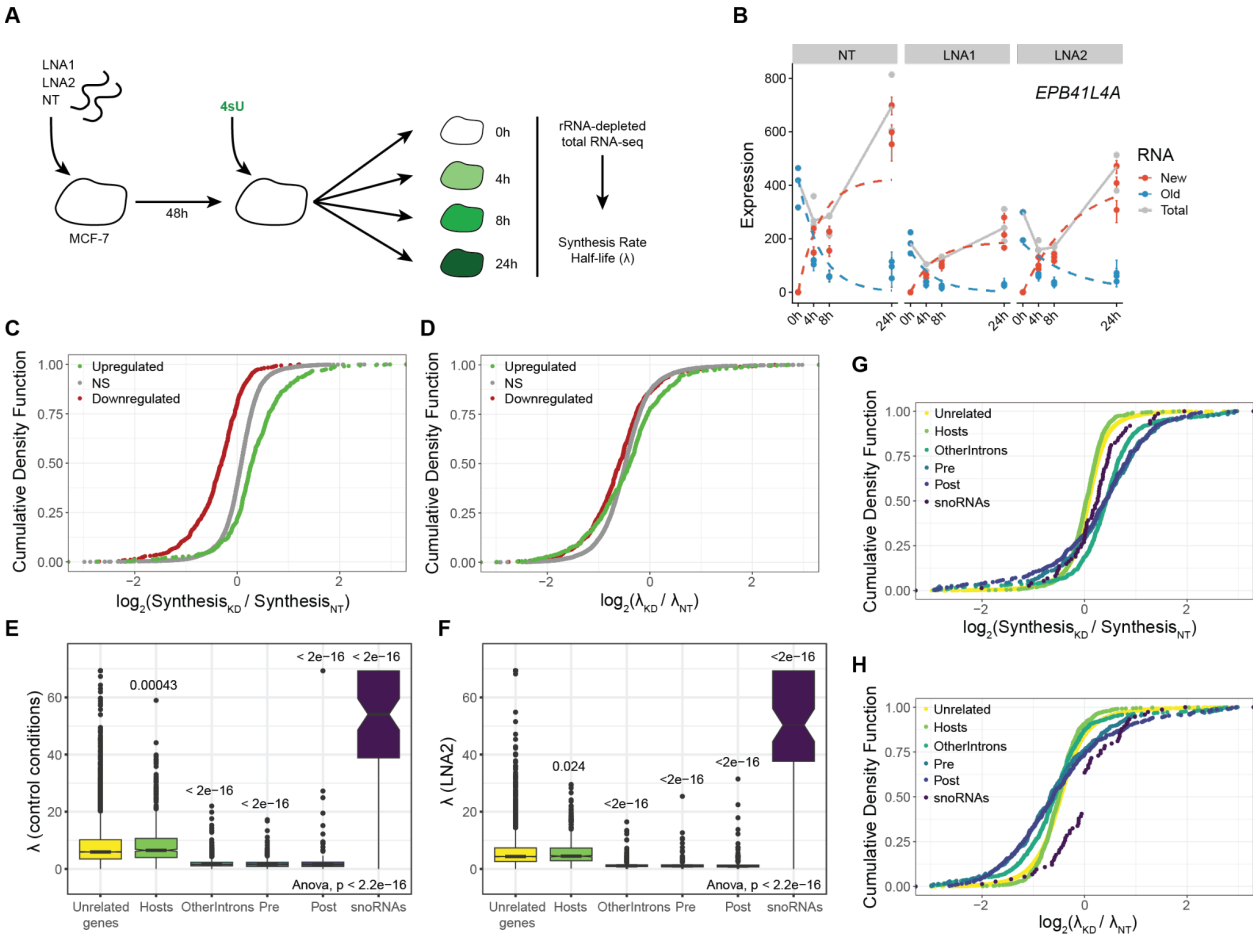
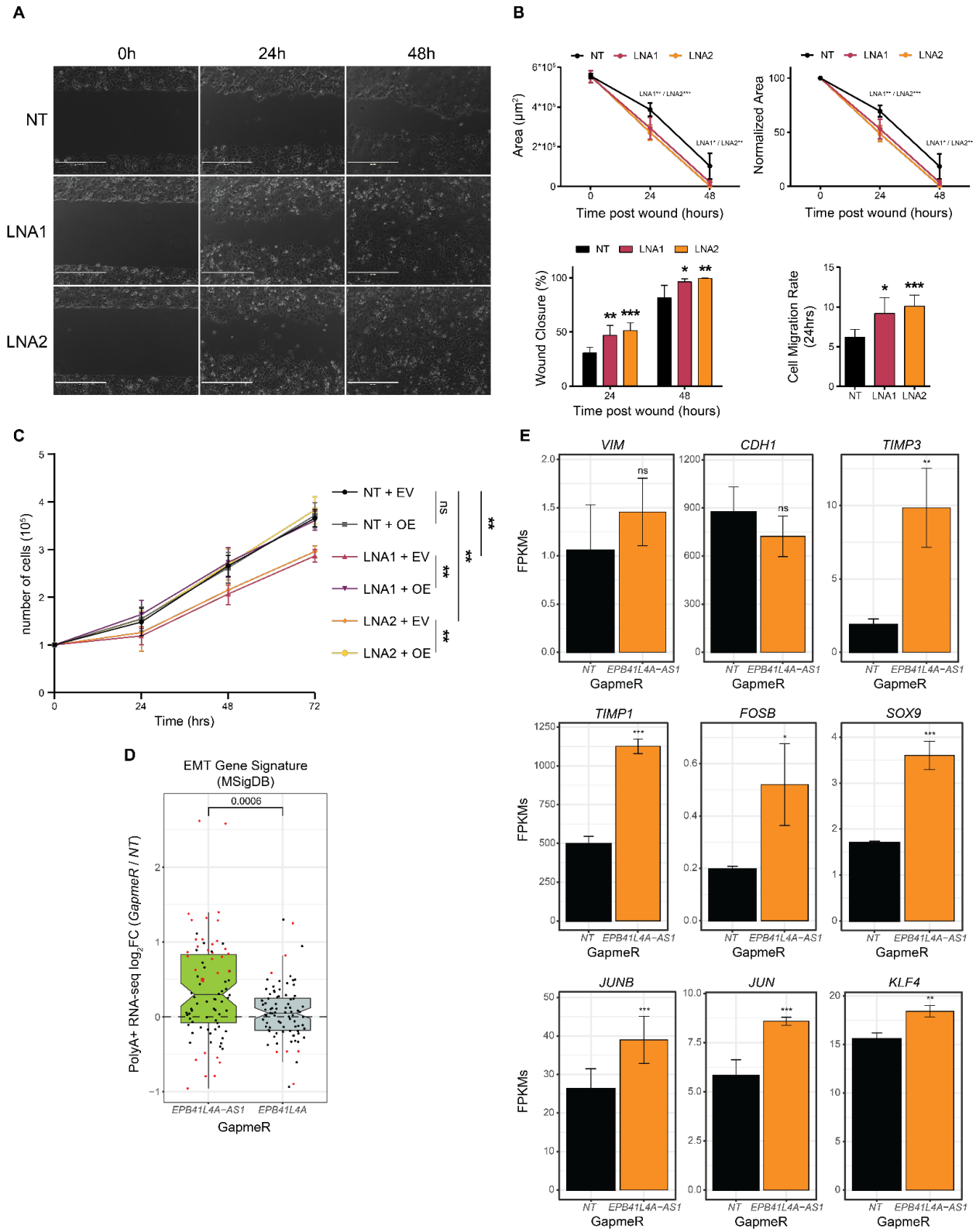


Figure 7. The increased abundance of snoRNAs is primarily due to their hosts' increased transcription and stability. **(A)** Workflow of the SLAM-seq experiment. MCF-7 cells were transfected with the indicated GapmeRs for 48 hours, after which the media was replaced with media containing 4sU and the 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. P-values refer to the comparison with the unrelated genes. **(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**

417 Lastly, we aimed to assess the phenotypic consequences of *EPB41L4A-AS1* loss. As
418 mentioned above, the genes affected by the KD were found to be enriched for GO terms related
419 to cell adhesion, movement and division (**Fig. 2I**). Thus, we measured cell migration and
420 proliferation by performing wound healing and cell counting assays in cells with reduced
421 *EPB41L4A-AS1* levels. Cells transfected with *EPB41L4A-AS1* targeting GapmeRs exhibited
422 significantly higher migration rates, compared to controls (**Fig. 8A** and **8B**), with cells bordering
423 the wounded region displaying a flattened and translucent morphology (**Fig. 8A**). This higher
424 migration is not due to higher cellular proliferation/division, as cells with reduced lncRNA levels
425 proliferate less (**Fig. 8C**), and this deficiency could be rescued by co-transfecting a vector
426 encoding the GapmeR-resistant, full-length *EPB41L4A-AS1* (**Fig. 8C**), suggesting that this
427 modulation takes place in *trans*.

428 These findings suggest that proper *EPB41L4A-AS1* expression is required for cellular
429 proliferation, whereas its deficiency results in the onset of more aggressive and migratory
430 behavior, likely linked to the increase of the gene signature of epithelial to mesenchymal
431 transition (EMT) (**Fig. 8D**). Because EMT is not characterized by a unique gene expression
432 program and rather involves distinct and partially overlapping gene signatures[81], we checked
433 the expression level of marker genes linked to different types of EMTs (**Fig. 8E**). The most
434 upregulated gene in Fig. 8D is *TIMP3*, a matrix metallopeptidase inhibitor associated with a
435 particular EMT signature that is less invasive and more profibrotic (EMT-T2)[81]. Interestingly, we
436 observed a stark upregulation of other genes linked to EMT-T2, such as *TIMP1*, *FOSB*, *SOX9*,
437 *JUNB*, *JUN* and *KLF4*, whereas *MPP* genes (linked to EMT-T1, which is highly proteolytic and
438 invasive) are generally downregulated or not expressed. This suggests that the downregulation
439 of *EPB41L4A-AS1* is primarily linked to a specific EMT program (EMT-T2), and future studies
440 aimed at uncovering the exact mechanisms and relevance will shed light upon a possible
441 therapeutic potential of this lncRNA.



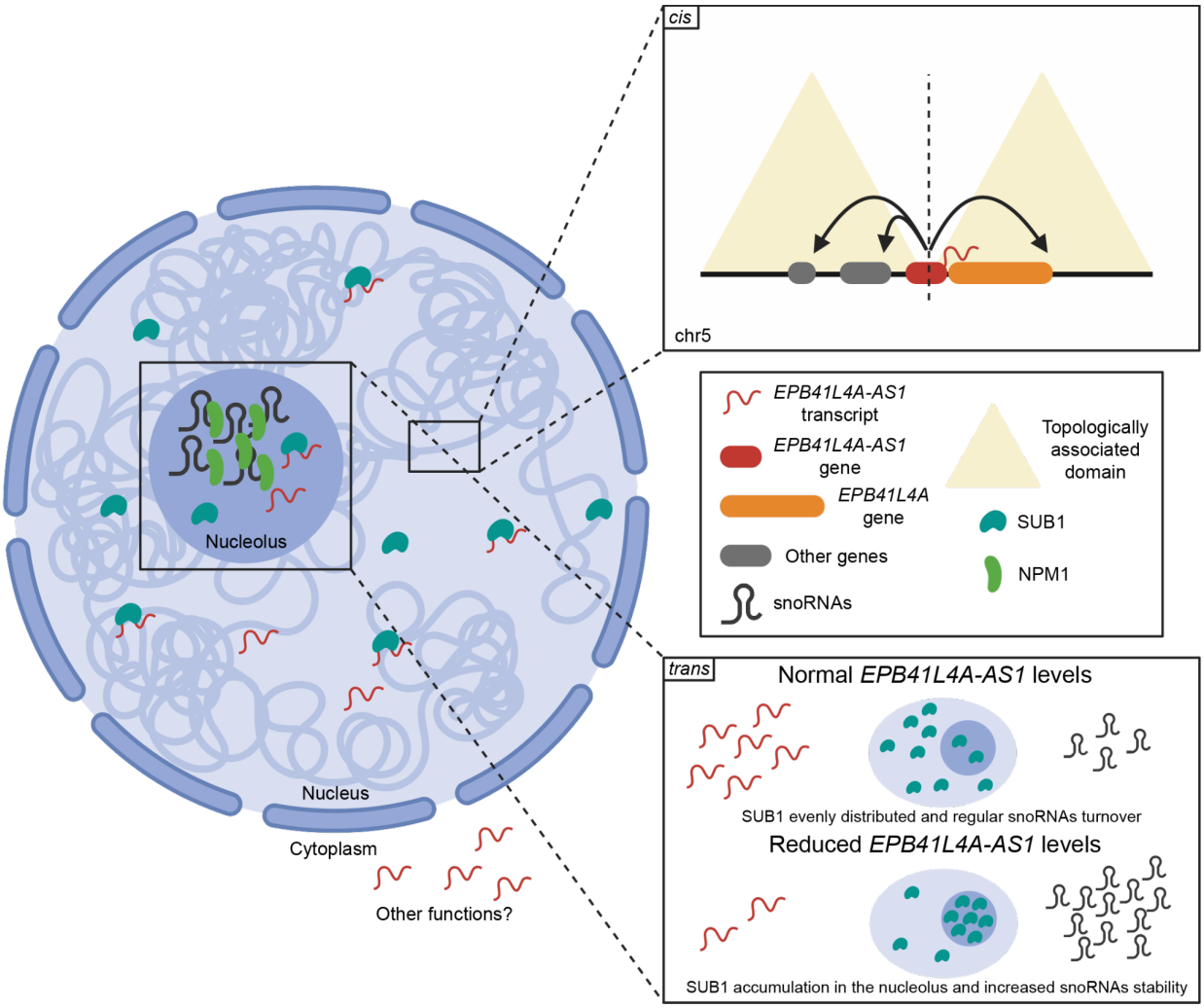
442

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. **(E)** Average expression levels (in FPKMs) following *EPB41L4A-AS1* KD with GapmeRs of selected genes that have been previously linked to EMT, with the error bars representing the standard deviation across the n = 3 replicates. DESeq2 adjusted P-values compared to control GapmeR are also reported. 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 (**Fig. 9**). At the site of its own transcription, which overlaps a strong TAD boundary, *EPB41L4A-AS1* is required to maintain expression of several adjacent genes, which are 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[47,82]. 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. S11C**). We also found that the KD of *EPB41L4A-AS1* for 72 hours—which depletes the RNA product, but may also affect the nascent RNA transcription[53,54]—reduces the spatial contacts between the TAD boundary and the *EPB41L4A* promoter (**Fig. 2E** and **Fig. S4J**). 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.



465

Figure 9. Proposed model of the *cis* and *trans* functions of the *EPB41L4A-AS1* lncRNA. Under normal conditions, the *EPB41L4A-AS1* locus is transcribed and ensures the proper expression of several genes located in *cis*. In *trans*, it interacts with SUB1 which eventually is evenly distributed throughout the nucleus. When there are reduced levels of *EPB41L4A-AS1*, the expression of the genes in *cis* is reduced, and SUB1 accumulates in the nucleolus. As a consequence, snoRNAs abundance increases, likely due to their increased stability, and MCF-7 cells acquire a more invasive *in vitro* phenotype. Created with Biorender.

466

467 A recent study found that *SNORA13* negatively regulates ribosome biogenesis in TERT-
468 immortalized human fibroblasts (BJ-HRAS^{G12V}), by decreasing the incorporation of RPL23 into
469 the maturing 60S ribosomal subunits, eventually triggering p53-mediated cellular senescence[27].
470 In their experimental setting, *SNORA13* was constitutively reduced by either CRISPRi of the
471 *EPB41L4A-AS1* locus for 7 days or its KO leaving the lncRNA sequence and expression
472 unaltered, whereas our approach relies solely on transient knockdown for 48–72hrs that does not
473 appreciably affect *SNORA13* (**Fig. 1B, 2B, 6A** and **6B**), or the levels of the RNA modification that
474 it guides (**Fig. 6C** and **6D**). Therefore, while both studies implicate *EPB41L4A-AS1* in nucleolar
475 biology and its perturbation in nucleolar stress, there is an apparent discrepancy between our

476 results on the role that *SNORA13* plays. In our hands, the differential expression of snoRNAs and
477 cell proliferation can be rescued by the expression of *EPB41L4A-AS1* with the snoRNA, and not
478 by separate expression of *SNORA13*, suggesting the snoRNA alone is not playing a central role
479 in the observed phenotypes. Also, the study by Cheng et al. showed that activity of the p53
480 pathway was less active in *SNORA13* knockout cells, with or without induction of oncogenic
481 stress, whereas we see an opposite trend of some up-regulation of the p53 pathway upon
482 transient *EPB41L4A-AS1* KD (**Fig. S11D**). We cannot however completely rule out potential
483 contributions of this snoRNA to the phenotypes observed in this study. For instance, the
484 phenotype of snoRNAs expression is partially rescued by *EPB41L4A-AS1* overexpression in *trans*
485 only when a full-length, unspliced and *SNORA13*-containing vector is used, and not when the
486 snoRNA is deleted, mutated or expressed alone (**Fig. 6I**). This would suggest that the processing
487 of *EPB41L4A-AS1* is involved in its mode of action, although further experiments will be needed
488 to address whether *SNORA13* or its DNA sequence are indeed required to regulate the
489 expression of the snoRNAs. We note that compensatory effects are possible in cells that grow for
490 extended periods of time without *SNORA13*.

491 Beyond its site of transcription, *EPB41L4A-AS1* associates with SUB1, an abundant protein
492 linked to various functions, and these two players are required for the proper distribution of at
493 least two nuclear proteins. Their dysregulation results in large-scale changes in gene
494 expression, including up-regulation of snoRNA expression, mostly through increased
495 transcription of their hosts, and possibly through a somewhat impaired snoRNA processing
496 and/or stability. To further hinder our efforts in discerning between these two possibilities, the
497 exact molecular pathways involved in snoRNAs biogenesis, maturation and decay are still not
498 completely understood. One of the genes that most consistently was affected by *EPB41L4A-*
499 *AS1* KD with GapmeRs is *MTREX* (also known as *MTR4*), that becomes downregulated at both
500 the RNA and protein levels (**Fig. S12A-C**). Interestingly, MTREX it is part of the NEXT and
501 PAXT complexes[83], that target several short-lived RNAs for degradation, and the depletion of
502 either *MTREX* or other complex members leads to the upregulation of such RNAs, that include
503 PROMPTs, uaRNAs and eRNAs, among others. It is therefore tempting to hypothesize a role
504 for *MTREX*-containing complexes in trimming and degrading those introns, and releasing the
505 mature snoRNAs. Future studies specifically aimed at uncovering novel players in mammalian
506 snoRNA biology will both conclusively elucidate whether *MTREX* is indeed involved in these
507 processes.

508 Several features of the response to *EPB41L4A-AS1* resemble nucleolar stress, including altered
509 distribution of NPM1[74,75]. SUB1 was shown to be involved in many nuclear processes, including
510 transcription[58], DNA damage response[84,85], telomere maintenance[61], and nucleolar processes
511 including rRNA biogenesis[63,86]. Our results suggest a complex and multi-faceted relationship
512 between *EPB41L4A-AS1* and SUB1, as *SUB1* mRNA levels are reduced by the transient (72
513 hours) KD of the lncRNA (**Fig. 3B**), the distribution of the protein in the nucleus is altered (**Fig.**
514 **5A** and **5C**), while the protein itself is the most prominent binder of the mature *EPB41L4A-AS1*
515 in ENCODE eCLIP data (**Fig. 3A**). The most striking connection between *EPB41L4A-AS1* and
516 SUB1 is the similar phenotype triggered by their loss (**Fig. 4**). We note that a recent study has
517 shown that SUB1 is required for Pol I-mediated rDNA transcription in the nucleolus[63]. In the
518 presence of nucleolar SUB1, rDNA transcription proceeds as expected, but when SUB1 is

519 depleted or its nucleolar localization is affected—by either sodium butyrate treatment or
520 inhibition of KAT5-mediated phosphorylation at its lysine 35 (K35)—the levels of the 47S pre-
521 rRNA are significantly reduced. In our settings, SUB1 enriches into the nucleolus following
522 *EPB41L4A-AS1* KD; thus, we might expect to see a slightly increased rDNA transcription or no
523 effect at all, given that SUB1 localizes in the nucleolus in baseline conditions as well. It is
524 however difficult to determine which of the connections between these two genes is the most
525 functionally relevant and which may be indirect and/or feedback interactions. For example, it is
526 possible that *EPB41L4A-AS1* primarily acts as a transcriptional regulator of *SUB1* mRNA, or
527 that its RNA product is required for proper stability and/or localization of the SUB1 protein, or
528 that *EPB41L4A-AS1* acts as a scaffold for the formation of protein-protein interactions in which
529 SUB1 is involved.

530 In the context of cancer, a systematic study of TCGA data, which is consistent with our analysis
531 (**Fig. S2**) found that *EPB41L4A-AS1* is downregulated in tumors compared to tumor-adjacent
532 normal tissues[29], in accordance with both the downregulation of the TIGA1 protein[29] and the
533 reduction in the *SNORA13*-guided rRNA modification[26]. The association with poor prognosis[29] is
534 consistent with the observed increased invasiveness that we find in MCF-7 cells. Effects of
535 *EPB41L4A-AS1* on migration were also reported in colorectal cancer cells[32]. The notion that
536 cells with reduced *EPB41L4A-AS1* levels exhibit lower proliferation (**Fig. 8C**), yet increased
537 invasion (**Fig. 8A** and **8B**) is compatible with a function as an oncogene by promoting EMT (**Fig.**
538 **8D** and **8E**). Cells undergoing this process may reduce or even completely halt proliferation/cell
539 division, until they revert back to an epithelial state[87,88]. Notably, downregulated genes following
540 *EPB41L4A-AS1* KD are enriched in GO terms related to cell proliferation and cell cycle
541 progression (**Fig. 2I**), whereas those upregulated for terms linked to EMT processes. Thus,
542 while we cannot rule out a potential function as tumor suppressor gene, our data better fits the
543 idea that this lncRNA promotes invasion, and thus, primarily functions as an oncogene.

544 We note that this is not the first instance in which this lncRNA has been investigated. Several
545 recent studies have described the functions of *EPB41L4A-AS1* in various biological systems,
546 highlighting its dysregulation in cancer[29,32,89], diabetes[31], aging[90,91], and Alzheimer's disease
547 (AD)[91]. Both p53 and PGC-1α were shown to act upstream of *EPB41L4A-AS1* and regulate its
548 expression[29]. In these different systems, different pathways were described as regulated by this
549 lncRNA, including glycolysis, glutamine metabolism[29], and autophagy[91]. Mechanistically, it was
550 reported that this lncRNA regulates gene expression via its association with and upregulation of
551 KAT2A (GCN5/GNC5L2)[31,91]. These prior studies generally relied on candidate-gene or co-
552 expression approaches to identify the genes regulated by *EPB41L4A-AS1*, and did not
553 systematically examine its *cis*- and *trans*-acting effects, with the exception of a *cis*-acting
554 regulation, that was briefly hinted in a preprint posted four years ago[92], but not published so far
555 to the best of our knowledge. These studies also did not systematically investigate the *trans*-
556 regulated targets of *EPB41L4A-AS1*, did not separately deplete the nascent and the mature
557 forms of the lncRNA, and did not attempt to decouple its functions from those of *SNORA13*.

558

559

560

561

562 # Methods

563 Cell culture, transfection and treatments

564 MCF-7 cells were cultured in DMEM supplemented with 4.5 g/L of glucose, 4 mM L-Glutamine,
565 10% FBS, 1% PenStrep solution, grown in incubators maintained at 37°C with 5% CO_2, and
566 were detached for passaging using a Trypsin (0.05%) – EDTA (0.02%) solution. DNA plasmid,
567 siRNAs (final concentration 25 nM) and GapmeRs (50 nM) transfections were performed by
568 using Lipofectamine 3000 (ThermoFisher Scientific), according to the manufacturer's protocol
569 protocol. In all cases, cells were plated the day before the transfection in an antibiotic-free
570 medium which was subsequently changed prior to adding the transfection mix, and again after
571 24 hr to maximize cell viability. All transfections were carried out for 72 hours, unless specified
572 otherwise. The complete list with the sequences of both siRNAs and LNA/GapmeRs is reported
573 in **Table S5**.

574 Cell synchronization was performed using a standard double-thymidine block. Briefly, thymidine
575 (Sigma) was added to the culture medium at a final concentration of 2 mM, and the cells were
576 incubated for 18 hours. The medium was then replaced with standard complete medium for 9
577 hours, followed by the addition of fresh thymidine-containing medium. After an additional 15
578 hourS, the medium was replaced with standard complete medium to release the cells from the
579 G1/S boundary into the S phase.

580 Lipopolysaccharide (LPS) was purchased from Sigma, diluted in DNase/RNase-free H_2O and
581 added to the cell culture media at the indicated final concentrations. Hydrogen peroxide (H_2O_2)
582 was purchased from Sigma and added directly to the cell culture media at the indicated final
583 concentrations. Thapsigargin and Etoposide (Sigma) were diluted in tissue culture-grade DMSO
584 and added to the cell culture media at the indicated final concentrations. Cells were cultured
585 with these stressors for a total of three days, and cells supplemented with an equal volume of
586 the respective solvent were used as controls. CX-5461 (Sigma) was diluted in sterile-filtered
587 50 mM NaH_2PO_4 at a stock concentration of 10mM and added to the cell culture media at a final
588 concentration of 0.5, 1 or 5 μM. As a control, cells were supplemented with an equal volume of
589 50 mM NaH_2PO_4.

590 Plasmids construction

591 The sequence corresponding to the cDNA of human *EPB41L4A-AS1* (ENST00000688370.2_3)
592 was synthesized and cloned by Hylabs inside a pcDNA3.1(+) vector. The vector to express the
593 *EPB41L4A* ORF (NM_022140.5) was purchased from Genecopoeia (EX-A8889-Lv201). The
594 vector expressing a full length and unspliced *EPB41L4A-AS1* (pLV[ncRNA]-Puro-CMV) was
595 purchased from Vectorbuilder, and the target site of LNA2 was mutated by introducing four
596 mismatches (AACTTAAAAGCAGCGT → A<u>C</u>CTTA<u>G</u>AAG<u>T</u>AG<u>A</u>GT) via restriction-free cloning.

597 To generate the Δ*SNORA13*, ΔH, and ΔACA vectors, we ordered *EPB41L4A-AS1* gene
598 fragments with the desired mutations from TWIST. These fragments corresponded to the
599 sequence between the unique SalI and MluI restriction sites and were inserted via restriction
600 digestion. The pLKO.1 *SNORA13* plasmid[27] was a gift from Prof. Mendell (UT Southwestern).
601 All gRNAs were ordered as single-stranded oligos (Sigma) and cloned inside pKLV-Puro
602 vectors as previously described[93]. The complete list of gRNAs and their sequences is available
603 in **Table S5.** All plasmids were sent for Sanger and whole plasmid (Plasmidsaurus) sequencing
604 prior to use.

605 Wound healing assay and analysis

606 Wound healing experiments were carried out as previously described[94]. To perform the wound
607 healing assay, we employed the µ-Plate 24 Well system (Ibidi). Specifically, we transfected the
608 targeting and control GapmeRs via reverse transfection, diluting the transfection mixes directly
609 into the cell mixtures while seeding the cells. After reaching confluence overnight (O/N), the
610 culture inserts were gently removed with sterile forceps and the wells were washed with 1X
611 PBS. Brightfield images were taken using an EVOS Cell Imaging System (Thermo Fisher
612 Scientific) at different time points. Data analysis was then performed using the
613 Wound_healing_size_tool[95] plugin for ImageJ, which calculates the total (μm^2) and percentage
614 (%) of the wounded area. From these data, we calculated the rate of cell migration (R_M) as:

615
$$R_M = \frac{W_i - W_t}{t}$$

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

618
$$Wound\ Closure\ (\%) = \frac{A_{t=0} - A_{t=\Delta t}}{A_{t=0}}$$

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

621 Gene expression analysis

622 RNA was extracted by using TRI-Reagent as previously described[94], according to the
623 manufacturer's instructions. The isolated RNA was then resuspended in DNase/RNase-free
624 H_2O and DNase-treated using the Baseline-ZERO™ DNase (Biosearch Technologies), following
625 the manufacturer's recommendations and incubating the RNAs for 1 hour at 37°C. RNA was
626 then retrotranscribed using the qScript™ Flex cDNA Synthesis Kit (Quanta Bio), following
627 standard conditions as suggested by the manufacturer. The resulting cDNA was finally diluted
628 1:5-1:10 with DNase/RNase-free H_2O and used as input for Quantitative Real-Time PCR using

629 the Fast SYBR™ Green Master Mix (Applied Biosystems). Each reaction was amplified using
630 the following PCR program: 20'' at 95°C, 40 cycles by denaturing for 1" at 95°C followed by
631 annealing/extension for 20'' at 60°C. Results were analyzed using the ΔΔCt method. The
632 complete list of primers and their sequences is available in **Table S5**.

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

634 The aminocarboxylpropyl reverse transcription (aRT)-PCR assay to detect and quantify the
635 1248.m1acp3Ψ modification was performed as previously described[26], with minor modifications.
636 Briefly, total RNA was extracted as described above, and 1 µg per sample was DNase-treated
637 using TURBO DNase (Invitrogen), according to the manufacturer's protocol. Retrotrascription
638 was then carried out using SuperScript III (Invitrogen), following the manual's recommendations
639 and by using only random hexamers. The cDNA was then diluted 1:5 with DNase/RNase-free
640 H_2O and used for end-point PCR using the Q5 High-Fidelity DNA Polymerase (NEB), according
641 to the standard protocol and using the primers listed in the original protocol[26] and available in
642 **Table S5**. From each PCR reaction, 5 µL were digested in duplicate with HinFI (NEB) in
643 rCutSmart buffer (NEB) and an additional 5 µL were incubated without the enzyme as a
644 negative control. Lastly, the resulting fragments were then run on a 2% agarose gel.

645 Northern Blot

646 Northern Blot to detect *SNORA13* was performed as previously described[27], with minor
647 modifications. A total of 20 µg of RNA per sample was loaded onto an 8% TBE-UREA
648 polyacrylamide gel, which had been pre-equilibrated by running at low voltage in 1X TBE. After
649 electrophoresis, the gel was stained for 15 minutes in 0.5X TBE with a few drops of ethidium
650 bromide while rocking, to assess the quality of the run. The RNA was then transferred to an
651 HyBond-NX membrane (GE Healthcare) in 0.5X TBE at room temperature (RT) using a
652 constant current of 0.25 A for 2 hours. Following the transfer, the membrane was rinsed in 2X
653 SSC, crosslinked at 1200 J/cm² in a UV oven, and pre-hybridized for 30 minutes at 42°C in 10
654 mL of pre-warmed ULTRAhyb™ Ultrasensitive Hybridization Buffer (ThermoFisher Scientific)
655 while rotating. IR fluorescent probes targeting *U6* and *SNORA13* were then added to a final
656 concentration of 10 nM, and the membrane was incubated O/N at 42°C while rotating. The next
657 day, the membrane was washed twice in 2X SSC, 0.1% SDS at 42°C for 5 minutes while
658 rotating, followed by two additional washes in 0.1X SSC, 0.1% SDS, and then finally imaged.
659 The complete list of probes is available in **Table S5**.

660 Protein extraction and Western Blot

661 Proteins were extracted using RIPA as previously described[94]. Briefly, cells were detached
662 using a 0.05% Trypsin – 0.02% EDTA solution, pelleted by centrifugation, resuspended in ice-
663 cold RIPA buffer, and incubated on ice for 20 minutes with intermittent vortexing. The lysates
664 were then centrifuged at 15,000g for 15 minutes at 4°C, after which the supernatants were
665 collected and quantified. For Western blots, equal amounts of protein were loaded onto a
666 polyacrylamide gel and run in an SDS running buffer. Proteins were then transferred onto a pre-
667 activated PVDF membrane in a cold Tris-Glycine buffer supplemented with 20% methanol, with
668 a constant current of 0.30 A for 2 hours. Membranes were then blocked in 5% milk in 1X PBS-

669 0.1% Tween 20 for 1 hour, followed by O/N incubation at 4°C with primary antibodies diluted in
670 the same blocking solution. The next day, membranes were washed three times with 1X PBS-
671 0.1% Tween 20, incubated for 2 hours at room temperature with the appropriate secondary
672 antibodies, and washed three more times before image acquisition. The complete list of primary
673 and secondary antibodies and relative dilutions is available in **Table S5**.

674 Immunostaining

675 Immunofluorescence was performed as previously described[94]. Cells were plated and
676 transfected on either sterile glass coverslips or 8-well chambers (Ibidi). They were first washed
677 once with 1X PBS, then fixed with cold 4% paraformaldehyde (PFA) in 1X PBS for 15 minutes
678 at room temperature. After two 5-minute washes with 1X PBS, cells were permeabilized by
679 incubating them in a permeabilization/blocking solution containing 5% horse serum, 1 mg/mL
680 BSA, and 0.1% Triton X-100 in 1X PBS for 30 minutes at room temperature. Following three
681 washes with 1X PBS, cells were incubated O/N at 4°C with primary antibodies diluted in
682 blocking solution (5% horse serum, 1 mg/mL BSA in 1X PBS). The next day, cells were washed
683 three times with 1X PBS and incubated for 2 hours at room temperature with secondary
684 antibodies in blocking solution, protected from light. After three additional washes with 1X PBS,
685 cells were stained with DAPI and mounted using ProLong™ Glass Antifade Mountant
686 (Invitrogen). The slides were left to cure O/N in the dark at room temperature before imaging
687 with a Zeiss Spinning Disk confocal microscope. The complete list of primary and secondary
688 antibodies and relative dilutions is available in **Table S5**.

689 Subcellular fractionation

690 To fractionate cells and isolate the cytoplasmic, nucleoplasmic and chromatin RNA fractions we
691 employed a published protocol[96] with minor modifications. Confluent 10 cm plates were
692 harvested with a Trypsin (0.05%) – EDTA (0.02%) solution, pelleted by centrifugation, washed
693 once with ice-cold 1X PBS and resuspended in 380 µL of ice-cold HLB buffer (10 mM Tris-HCl,
694 pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 0.3% NP-40, 10% Glycerol) freshly supplemented with
695 RNase Inhibitors (EURX). The samples were then incubated on ice for 20' with occasional
696 vortexing, spun down for 3' at 4°C / 1000 g, the supernatant was then transferred to a new tube,
697 centrifuged again for 10' at 4°C / 1000 g, after which the supernatant (cytoplasmic fraction) was
698 moved to a new tube, supplemented with 1 mL of RNA precipitation solution (RPS, 150mM
699 Sodium Acetate in 100% ethanol) and stored at -20°C for at least 1 h. The semi-pure nuclei
700 were then washed three times with ice-cold HLB and centrifugations for 2' at 4°C / 500 g, after
701 which the pellets were resuspended in 380 µL of ice-cold MWS (10 mM Tris-HCl, pH 7.0, 4 mM
702 EDTA, pH 8.0, 0.3 M NaCl, 1 M Urea, 1% NP-40) freshly supplemented with RNase Inhibitors,
703 quickly vortexed and incubated on ice for 5 minutes, vortexed again and incubated for an
704 additional 10 minutes. The tubes were then centrifuged for 3' at 4°C / 1000 g, the supernatant
705 (nucleoplasmic / nuclear soluble fraction) was saved to a new tube, supplemented with 1 mL of
706 RPS and stored at -20°C for at least 1 h. The chromatin pellets were finally washed three times
707 with ice-cold MWS and centrifugations for 2' at 4°C / 500 g, after which 1 mL of TRI-Reagent
708 was added and samples were stored at -20°C. After 1+ hours at -20°C, the tubes with the
709 cytoplasmic and nucleoplasmic fractions in RPS were briefly vortexed, centrifuged for 15' at 4°C

710 / 15000 g, the pellets were washed with 70% ethanol by vortexing and centrifuged again for 15'
711 at 4°C / 15000 g, after which they were air-dried and resuspended in 1 mL of TRI-Reagent. To
712 each sample in 1 mL of TRI-Reagent, 10 µL of 0.5 M EDTA, pH 8.0 were added followed by an
713 incubation with vortexing for ~10 minutes at 65°C, or until the RNA pellets fully dissolved. The
714 RNA was then extracted following the standard TRI-Reagent protocol.

715 Native RNA Immunoprecipitation (RIP)

716 To perform Native RNA Immunoprecipitation, we have employed a slightly modified version of
717 the ENCODE IP protocol[97]. Briefly, $5x10^6$ MCF-7 cells were harvested by trypsinization, washed
718 once with 1X PBS, and flash frozen as pellets in 50 µL of 1X PBS. The pellets were then thawed
719 on ice, resuspended in 450 µL of ice-cold lysis buffer (50 mM Tris-HCl, pH 7.4; 100 mM NaCl;
720 1% NP-40; 0.1% SDS; 0.5% sodium deoxycholate) supplemented with EDTA-free Protease
721 Inhibitor Cocktail (APExBIO)) and RNase Inhibitor (EURX), and incubated on ice for 20' with
722 occasional vortexing to allow complete lysis. The lysates were then centrifuged at max speed
723 for 15' at 4°C and the supernatants were collected into a new tube. A total of 20 µL/sample of
724 Protein A/G beads were then washed twice with ice-cold lysis buffer and resuspended in 250 µL
725 of lysis buffer, of which 50 µL were used to pre-clear the lysate by incubation at RT for 30' with
726 gentle rotation, and the remaining 200 µL were instead incubated for 1 h/RT with 3 µg of
727 antibody. After 30' of preclearing, the clear lysate was moved to a new tube in ice, and 25 µL
728 (5%) of it was saved as INPUT sample. After 1 h, the antibody/beads were washed three times
729 with ice-cold lysis buffer and incubated with the pre-cleared lysate O/N at 4°C while gently
730 rotating. The following day, 25 µL (5%) of the supernatant was saved as UNBOUND sample,
731 and the beads (BOUND sample) were gently washed twice with ice-cold wash buffer (20 mM
732 Tris-HCl pH 7.4, 10 mM MgCl2, 0.2% Tween-20) freshly supplemented with RNase Inhibitors
733 (EUR_X). Finally, 500 µL of TRI-Reagent was added to all three fractions and RNA isolation was
734 carried out as previously described. The complete list of antibodies and relative dilutions is
735 available in **Table S5**.

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

737 To visualize the subcellular localization of EPB41L4A-AS1 in vivo, we performed single-
738 molecule fluorescence in situ hybridization (smFISH) using HCR™ amplifiers. Probe sets (n =
739 30 unique probes) targeting *EPB41L4A-AS1* and *GAPDH* (positive control) were designed and
740 ordered from Molecular Instruments. We followed the Multiplexed HCR v3.0 protocol with minor
741 modifications. MCF-7 cells were plated in 8-well chambers (Ibidi) and cultured O/N as described
742 above. The next day, cells were fixed with cold 4% PFA in 1X PBS for 10 minutes at RT and
743 then permeabilized O/N in 70% ethanol at -20°C. Following permeabilization, cells were washed
744 twice with 2X SSC buffer and incubated at 37°C for 30 minutes in hybridization buffer (HB). The
745 HB was then replaced with a probe solution containing 1.2 pmol of *EPB41L4A-AS1* probes and
746 0.6 pmol of *GAPDH* probes in HB. The slides were incubated O/N at 37°C. To remove excess
747 probes, the slides were washed four times with probe wash buffer at 37°C for 5 minutes each,
748 followed by two washes with 5X SSCT at RT for 5 minutes. The samples were then pre-
749 amplified in amplification buffer for 30 minutes at RT and subsequently incubated O/N in the
750 dark at RT in amplification buffer supplemented with 18 pmol of the appropriate hairpins. Finally,

751 excess hairpins were removed by washing the slides five times in 5X SSCT at RT. The slides
752 were mounted with ProLong™ Glass Antifade Mountant (Invitrogen), cured O/N in the dark at
753 RT, and imaged using a Nikon CSU-W1 spinning disk confocal microscope. In order to estimate
754 the RNA copy number, we imaged multiple distinct fields, extracted the number of *EPB41L4A-*
755 *AS1*/*GAPDH* molecules in each field using the "Find Maxima" tool in ImageJ/Fiji, and divided
756 them by the number of cells (as assessed by DAPI staining).

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

758 CUT&RUN reactions were performed following the V3 of the protocol[98], with minor modifications
759 which were described previously[94]. An equal number of cells was harvested and washed once
760 with 1X PBS at room temperature. After three washes with Wash Buffer (20 mM HEPES-NaOH
761 pH 7.5, 150 mM NaCl, 0.5 mM Spermidine supplemented with EDTA-free Protease Inhibitor
762 Cocktail (APExBIO)) at RT, cells were resuspended in 1 mL Wash Buffer and incubated for 10
763 minutes while gently rotating with 20 µL of pre-activated Concanavalin A-coated magnetic
764 beads (EpiCypher). The buffer was then removed, and the beads were resuspended in 150 µL
765 of Antibody Buffer (Wash Buffer with 0.1% Digitonin (Sigma) and 2 mM EDTA, pH 8) containing
766 the antibody of interest, and left gently rotating O/N with a 180° angle at 4°C. On the next day,
767 cells were washed two times with ice-cold Dig-Wash Buffer (Wash Buffer with 0.1% Digitonin),
768 resuspended in 150 µL of Dig-Wash Buffer supplemented with 1 µL of custom-made pA/G-
769 MNase every mL and left rotating for 1h at 4°C. After that, cells were washed again two times
770 with Dig-Wash Buffer, resuspended in 100 µL of the same buffer, and chilled into a thermoblock
771 sitting on ice. To initiate the cleavage reaction, 2 µL of 100 mM $CaCl_2$ was added to each tube,
772 and the tubes were left at 0°C for 30 minutes. To halt the reaction, 100 µL of a 2X STOP Buffer
773 (340 mM NaCl, 20 mM EDTA pH 8, 4 mM EGTA pH 8, 0.05% Digitonin, 100 µg/mL RNAse A,
774 50 µg/mL Glycogen) were added to each tube, and cleaved DNA fragments were released by
775 incubating the samples for 30' at 37°C. The tubes were then centrifuged for 5 minutes at 4°C /
776 16,000g and the supernatant was collected. Finally, the DNA was purified by standard
777 Phenol/Chloroform extraction using the 5PRIME Phase Lock Gel Heavy tubes (QuantaBio), and
778 the success of the CUT&RUN reaction was assessed by running the positive control
779 (H3K27me3 or H3K4me3) on a Tapestation (Agilent Technologies) using a High Sensitivity
780 D1000 ScreenTape (Agilent Technologies). The complete list of primary and secondary
781 antibodies and relative dilutions is available in **Table S5**.

782 RNA-seq and CUT&RUN Library Preparation

783 RNA-seq libraries were prepared using the CORALL mRNA-Seq Kit V2 (Lexogen), following the
784 manufacturer's instructions. Prior to library generation, 1 µg of total RNA was either PolyA-
785 enriched using the Poly(A) RNA Selection Kit V1.5 (Lexogen), or depleted of rRNAs using the
786 RiboCop rRNA depletion kit for HMR V2 (Lexogen). To prepare the CUT&RUN libraries, we
787 followed the original protocol[98], with slight modifications which were described previously[94]. The
788 libraries were then quality-checked by both dsDNA Qubit (Thermo Fisher Scientific) and
789 Tapestation (Agilent Technologies). All libraries were sequenced on an Illumina NovaSeq 6000
790 or Novaseq X instrument, aiming for either 10M (CUT&RUN) or 15M (RNA-seq) reads per
791 sample.

792 RNA-seq data analysis

793 Analysis of the RNA-seq libraries was performed as previously described[94]. Raw FASTQ files
794 were processed by a customized Lexogen CORALL analysis pipeline script
795 (https://github.com/Lexogen-Tools/corall_analysis). Briefly, adaptors were trimmed by using
796 Cutadapt[99], and the resulting trimmed FASTQ files were used as input for mapping to the hg19
797 human genome with STAR aligner[100], using the `--quantMode GeneCounts` option in order to
798 count the number of reads mapping to each feature provided by an appropriate GTF file. To call
799 for differential expression, we used DESeq2[101], considering a gene to be differentially expressed
800 when adjusted P<0.05 and absolute $\log_2$FC >0.41 (~33% increase or decrease). Prior to any
801 subsequent analysis, we filtered out poorly expressed genes, pseudogenes, and genes with an
802 exonic length <200nt. To perform GO enrichment analysis and GSEA, we used the
803 Clusterprofiler R package[102] by considering a term to be enriched when the adjusted P<0.05.
804 Whenever several redundant terms were found, we employed the *simplify* function with a
805 cutoff=0.7 to reduce the redundancy.

806 CUT&RUN data analysis

807 Raw FASTQ files were aligned to the hg19 reference genome with Bowtie2[103], using the -p 8, -X
808 2000 and --no-unal options. Peaks were called using the MACS2 callpeak function[104]. Peak
809 quantification and differential expression were performed using the HOMER tool suite[105], using
810 the *annotatePeaks.pl* and *getDiffExpression.pl* functions, respectively. Lastly, to plot the
811 metagene profiles and coverage heatmaps, we employed deepTools[106].

812 Computational screen to identify putative *cis*-acting lncRNAs

813 We reasoned that lncRNAs that are likely to act in *cis* will be occasionally co-regulated with their
814 target genes and also found in spatial proximity to the target promoters. We therefore sought
815 such lncRNA-target pairs using the GeneHancer database[22] that connects between regulatory
816 elements and their targets by combining tissue co-expression data, ChIP-seq of enhancer-
817 enriched histone modifications and transcription factor binding, Hi-C data and eQTLs within
818 putative enhancer regions. We complemented this by analyzing several RNA-seq datasets of
819 MCF-7 breast cancer cells exposed to different conditions[23–25], to infer genome-wide
820 transcriptional alterations across a variety of different experimental settings. We then focused
821 on those cases in which a lncRNA 1) overlaps a high-confidence GeneHancer (GH) element, 2)
822 is differentially expressed in at least one of the conditions analyzed, and 3) the targets of the
823 aforementioned GH element are also concomitantly differentially expressed. Because this
824 approach identified over 600 candidate lncRNA-target pairs (**Table S1**), we first discarded
825 lncRNAs which have been already reported to act in *cis*, and prioritized those cases in which the
826 lncRNA is well expressed and/or the target genes are biologically interesting. This led us to
827 focus on four lncRNA-target pairs connected in GeneHancer and in which both genes were
828 significantly either upregulated or downregulated in the treatment condition, compared to
829 controls (**Fig. S1A-D**).

830 Chromosome Conformation Capture (3C)

831 Between 5 and $10x10^6$ cells for each condition/replicate were collected and washed once with
832 1X PBS at RT. Cells were then resuspended in 10 mL of ice-cold 1X PBS with 10% FBS,
833 supplemented with another 10 mL of 1X PBS with 10% FBS and 4% Formaldehyde, and left
834 gently rotating for 5' at RT. Quenching was achieved by adding 1020 µL 2.5 M Glycine to each
835 tube and placing them on ice to cool down, after that the cells were centrifuged for 8' at 4°C /
836 300 g and washed once with ice-cold 1X PBS. The pellet was then resuspended in 10 mL of
837 Nuclear Permeabilization Buffer (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 0.5% NP-40)
838 supplemented with EDTA-free Protease Inhibitor Cocktail (APExBIO)), gently rotated for 1 h in a
839 cold room and centrifuged for 5' at 4°C / 600 g. After discarding the supernatant, the nuclei were
840 resuspended in ~1 mL of ice-cold 1X PBS, moved to a 1.5 mL centrifuge tube, centrifuged and
841 washed again with 1mL ice-cold 1X PBS, and eventually frozen as a pellet at -80 °C.

842 After 1+ days at -80 °C, each pellet was thawed on ice and gently resuspended in 175.5 µL of
843 nuclease-free H_2O, after which 24.5 µL of 10X DpnII Reaction Buffer was added and the tubes
844 were moved to a thermomixer set at 37°C. Each reaction was then supplemented with 3 µL of
845 pre-warmed 20% SDS, incubated for 1h at 37°C / 900g, supplemented again with 20 µL of pre-
846 warmed 20% Triton-X100 and incubated for an additional 1h at 37°C / 900g, after which 15 µL
847 of DpnII (NEB, 50.000 U/mL) were added and the reaction was finally incubated O/N at 37°C /
848 900g. The following day, the restriction reaction was removed by centrifugation for 5' at 4°C /
849 600 g, and the pellets were washed once with ice-cold 1X PBS, eventually leaving a small
850 volume (~30 µL) on top of the pellet. The tubes were then incubated for 20' at 65°C, to
851 inactivate DpnII activity, followed by thorough resuspension in 445 µL of nuclease-free H_2O and
852 were transferred to ice. After chilling, 50 µL of 10X Ligation Buffer (NEB), 0.5 µg BSA and 5 µL
853 of Quick Ligase (NEB) were added to each tube, and left incubating O/N at 16°C. After the O/N
854 incubation, the ligation reaction was removed by centrifugation for 5' at 4°C / 600 g, the pellets
855 were washed once with ice-cold 1X PBS and eventually resuspended in 250 µL 10 mM Tris-
856 HCl, pH 7.5. A total of 30 µg of RNase A were then added to each sample, which were
857 subsequently incubated for 45' at 37°C / 750g, after which 15 µL of PCR-grade Proteinase K
858 were added and incubated O/N at 65°C / 750g. The following day, the de-crosslinked reactions
859 (3C template) were cleaned with 0.5X Sera-Mag Size Selection Beads (Cytiva Life Sciences),
860 according to manufacturer's instructions, eluted in 200 µL Elution Buffer (Qiagen) and quantified
861 by dsDNA Qubit.

862 <u>UMI Circular Chromosome Conformation Capture (UMI-4C)</u>

863 Quantitative UMI-4C was performed as previously described[45], with minor modifications. An
864 equal amount of 3C template (5-10 µg) for each sample/replicate was diluted to a final volume
865 of 200 µL with Elution Buffer and sonicated to an extent the fragments size ranges between 250
866 and 1000bp. To the sonicated fragments, 20 µL of 10X End-Repair Buffer (NEB) and 10 µL of
867 End Repair Mix (NEB) were added and the samples were incubated for 30' at 20°C, after which
868 they were cleaned with 2.2X Sera-Mag Selection Beads and eluted in 78 µL 10 mM Tris-HCl,
869 pH 8.0. To 76 µL of the cleaned DNA, 10 µL 10X NEBuffer 2, 10 µL 10 mM dATP, and 4 µL
870 Klenow (3'-5' exo, NEB) were added, and the reaction was incubated for 30' at 37°C. After this,
871 the enzymes were inactivated by incubating the samples for 20' at 75°C, then 2 µL of FastCIP
872 (NEB) were added and the reactions were further incubated for 1 h at 50°C, followed by

873 purification with 2X Sera-Mag Size Selection Beads and elution in 67 µL of 10mM Tris-HCl, pH
874 8.0. Illumina-compatible adapters were ligated by adding 80 µL Quick Ligase Buffer (NEB), 10
875 µL Quick Ligase (NEB) and 4 µL of 15 µM Y-Shaped Adaptors (IDT) to 66 µL of the eluted DNA
876 from the previous step, followed by an incubation at 25°C for 15'. After that, the samples were
877 denatured for 2' at 95°C, placed on ice, brought to volume by adding 160 µL of nuclease-free
878 H_2O, cleaned with 1X Sera-Mag Size Selection Beads and eluted in 40 µL 10 mM Tris-HCl, pH
879 8.0. To quantify the resulting 4C template, 1 µL was diluted in 4 µL of 10mM Tris-HCl, pH 8.0,
880 denatured for 5' at 95°C, placed on ice and analyzed by ssDNA Qubit.

881 Libraries were then prepared using a nested PCR approach, preparing five reactions for each
882 sample and employing three bait primers for each target (EPB41L4A-AS1 TSS, EPB41L4A
883 TSS, STARD4 TSS, see next for design criteria) to increase complexity of the 4C profiles. For
884 the first PCR (PCR1), upstream bait primers were designed close (<80-85 bp) to a DpnII
885 restriction site (hg19 genome build); for the second PCR (PCR2), the downstream bait primers
886 were designed to follow – but not overlap – the upstream ones, ending ideally <20bp from the
887 DpnII site ("pad" sequence), and containing a sequence consisting of the Universal Illumina P5
888 on their 5' end. Both primers were designed to have a standard annealing temperature of
889 ~60°C, and their sequences are available in **Table S5**. PCR1 consisted of 200 ng of 4C
890 template, 10 µL of 10X Phusion Buffer (NEB), 1 µL of 10 mM dNTPs, 2.5 µL of pooled upstream
891 bait primers with a final concentration of 10 µM, 2.5 µL of 10 µM Universal Illumina P7 primer,
892 0.5 µL Phusion Polymerase (NEB), and nuclease-free H_2O up to a final volume of 50 µL. The
893 reactions were incubated for 30'' at 98°C, amplified with 20 cycles consisting of 10'' at 98°C, 30''
894 at 56°C and 60'' at 72°C, with a final extension of 10' at 72°C, individually purified with 0.8X
895 Sera-Mag Size Selection Beads and eluted in 20 µL of 10 mM Tris-Hcl, pH 8.0. PCR2 consisted
896 of 15 µL of purified template from PCR1, 10 µL of 10X Phusion Buffer (NEB), 1 µL of 10 mM
897 dNTPs, 2.5 µL of pooled downstream bait primers with a final concentration of 10 µM, 2.5 µL of
898 10µM Universal Illumina P7 primer, 0.5 µL Phusion Polymerase (NEB), and nuclease-free H_2O
899 up to a final volume of 50 µL. The reactions were incubated with the same program of PCR1 but
900 with 18 cycles of amplification instead of 20, all five replicates of each sample were then pooled
901 together, purified with 0.8X Sera-Mag Size Selection Beads and eluted in 40 µL of Elution
902 Buffer. The libraries were then quality-checked by both dsDNA Qubit and Tapestation, and
903 eventually pooled according to the relative number of reads needed. All libraries were
904 sequenced on an Illumina NovaSeq 6000 instrument, aiming at ~5M reads per sample.

905 <u>UMI-4C Data Analysis</u>

906 To analyze the UMI-4C data we have employed the umi4cpackage R package
907 (https://github.com/tanaylab/umi4cpackage). After setting up the configuration file as described,
908 they were loaded with the *p4cLoadConfFiles* function, and the raw fastq files were analyzed
909 using the *p4cCreate4CseqTrack* function. The 4C profiles were then built and plotted using the
910 *p4cNewProfile* and plot functions, respectively, specifying the *cis* window and overlaying two
911 profiles when needed. The p-values and log_2FC between conditions and across a genomic
912 region of interest were finally calculated using the *p4cIntervalsMean* function.

913 <u>Metabolic labeling, SLAM-seq and data analysis</u>

914 For half-life and synthesis rate estimation upon the KD of *EPB41L4A-AS1*, cells were
915 transfected with control and targeting GapmeRs as described above. Metabolic labeling and
916 SLAM-seq were then performed as previously described[107], with minor modifications. All steps
917 were performed in the dark until the chemical conversion with iodoacetamide (IAA). After 48h
918 hours, the growth medium was replaced with fresh media supplemented with 500 µM 4-
919 Thiouridine (4sU, Sigma), and replaced again with fresh 4sU after every time point. In total, we
920 collected RNA from cells at time 0 (unlabeled), and after 4, 8 and 24h hours post-4sU addition.
921 RNA extraction was carried out as described above with the addition of DTT (0.1 mM final
922 concentration) to keep the samples under reducing conditions, and IAA conversion was
923 performed as described previously[80,107]. Ribosomal RNAs were then depleted using the
924 NEBNext rRNA Depletion Kit V2 (NEB) and libraries were generated with the NEBNext Ultra II
925 Directional Library Kit (NEB), both according the manufacturer's instructions. Estimation of half-
926 lives and synthesis rates was performed using GRAND-SLAM/grandR[108], and are listed in **Table**
927 **S4**. Notably, the estimation assumes steady state expression in control samples, but not in the
928 KD ones.

Data availability

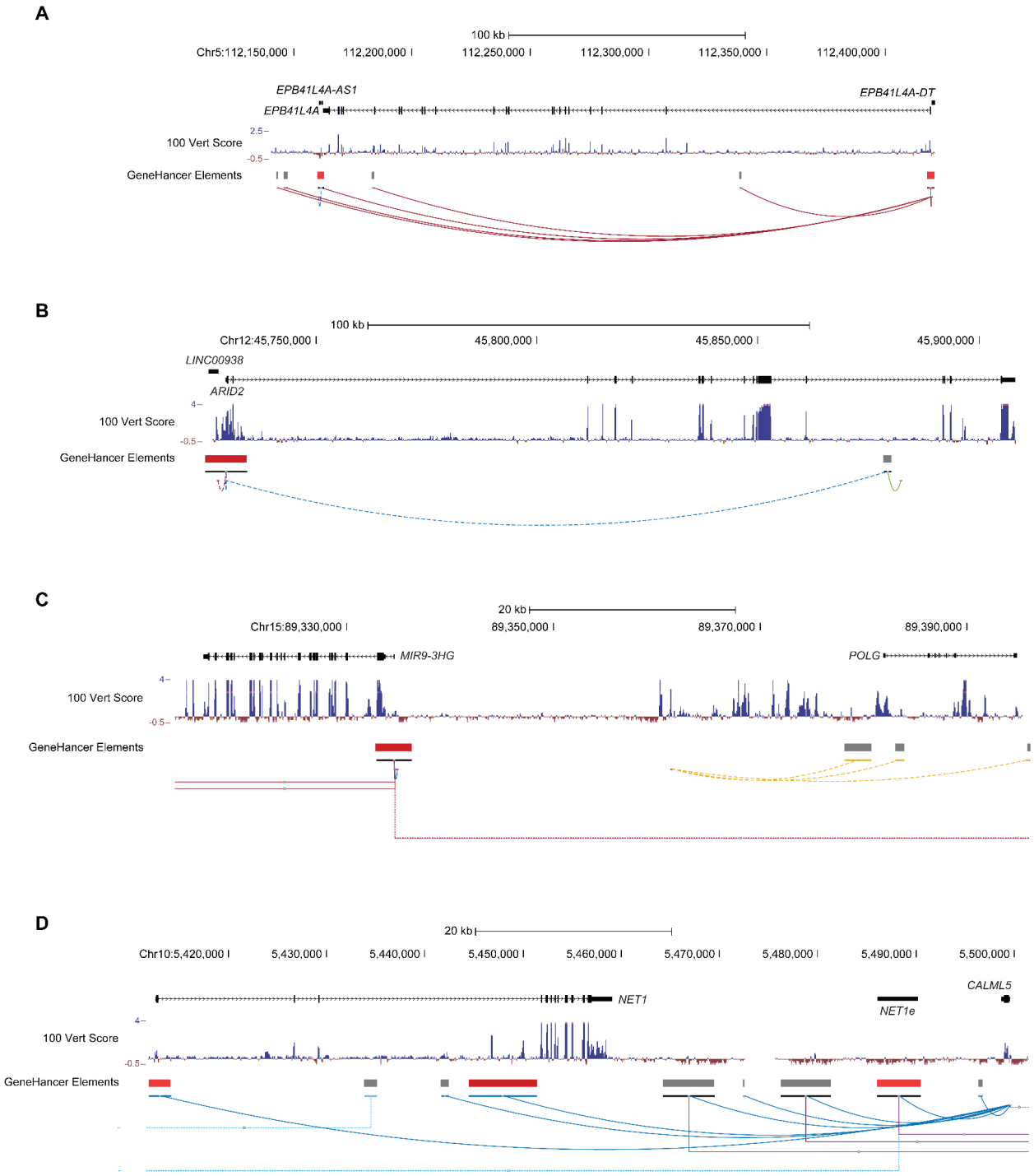
All the RNA-seq, RIP-seq, UMI-4C, and CUT&RUN datasets were deposited to the GEO database under the accession GSE292272.

Author contributions

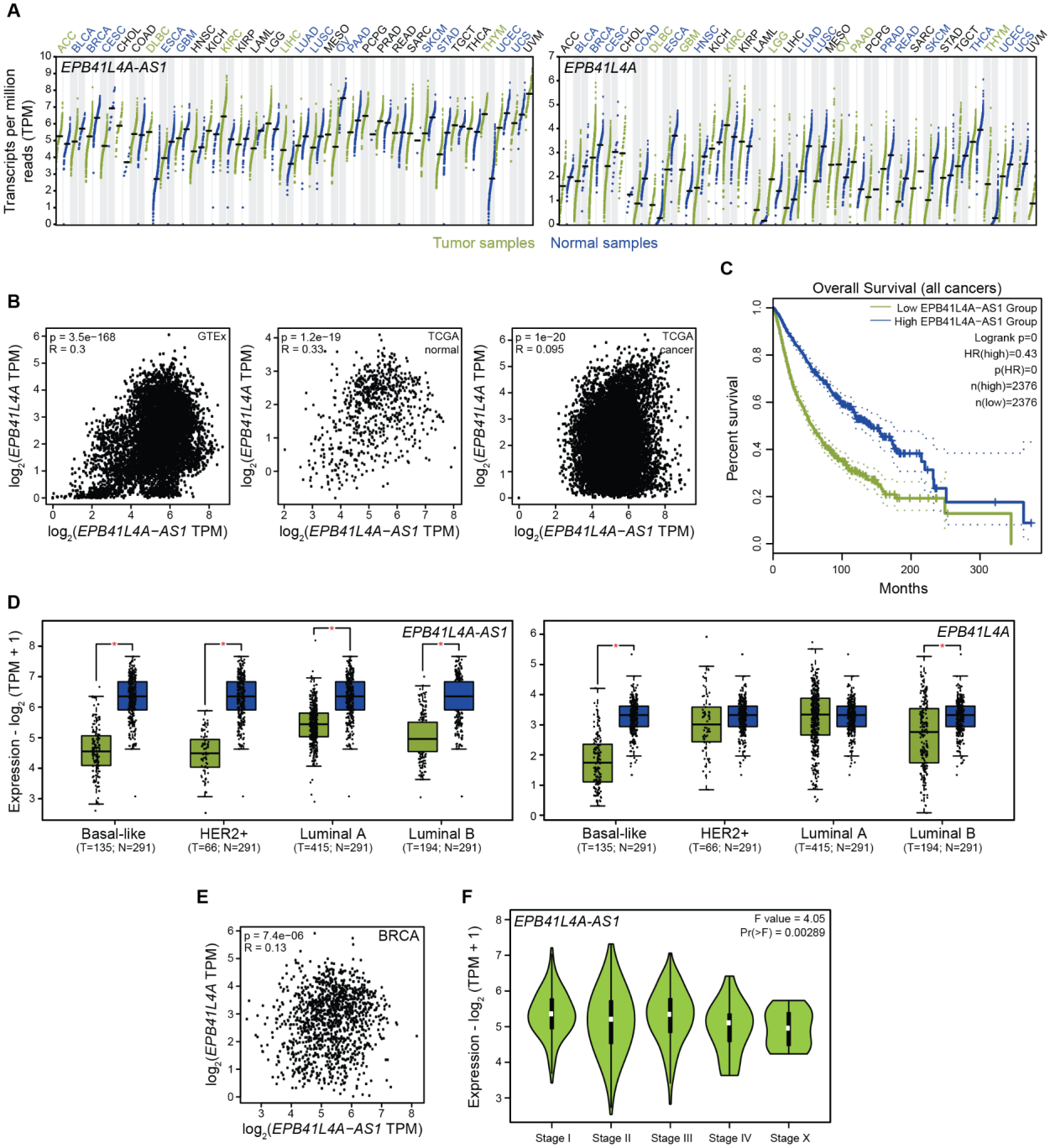
A.M. and I.U. conceived the study. A.M. carried out most experiments and analyzed the data. J.P.U. helped with experimental designs, UMI-4C experiments and provided insightful input. T.C. assisted with stress induction experiments. A.M. and I.U. wrote the manuscript with input from the other authors.

Acknowledgements

We thank members of the Ulitsky lab for insightful discussions and comments on the manuscript, and in particular Rotem Nevo for his help with the smFISH analysis. 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.



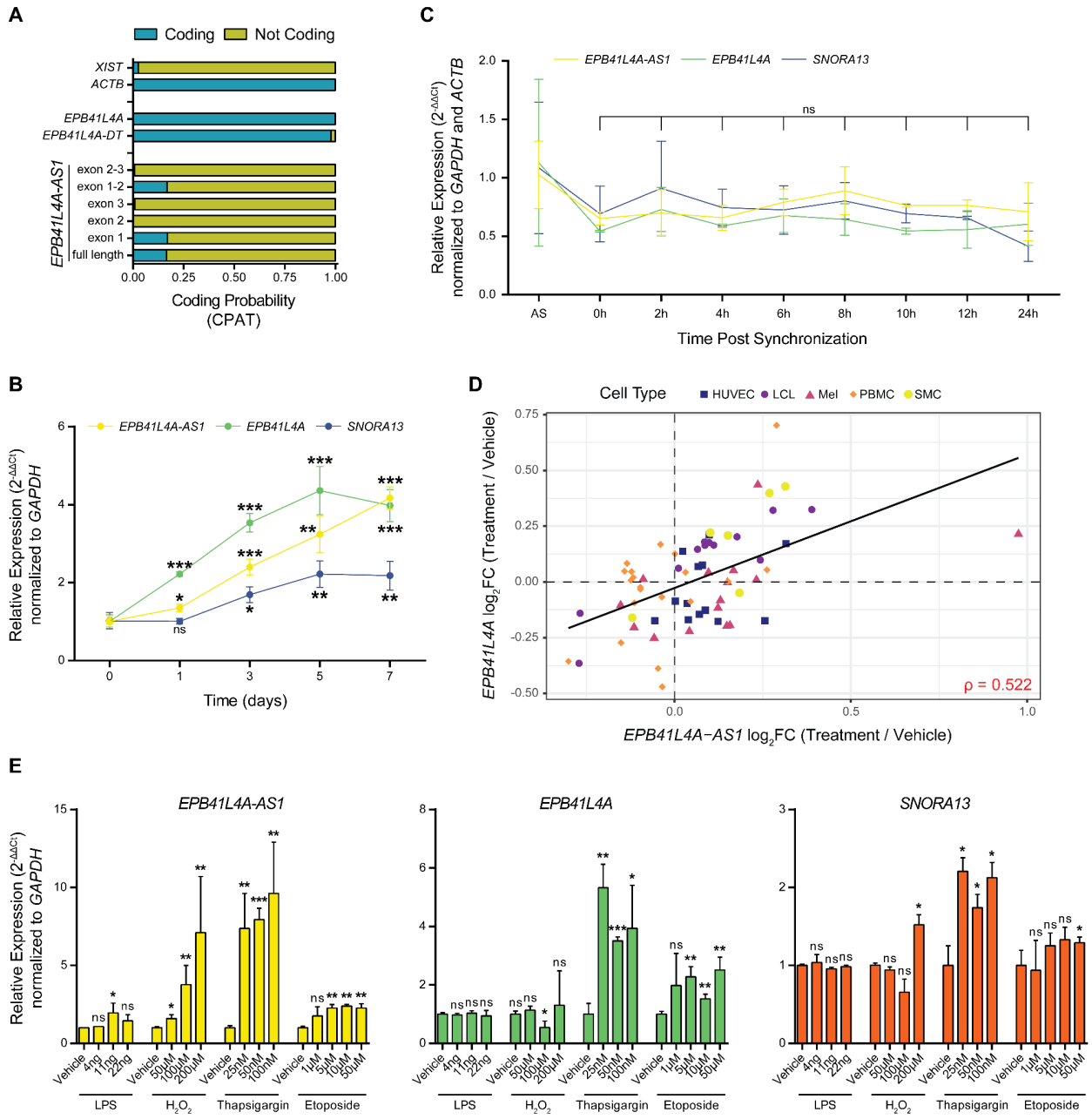
Supplementary Figure 1. **UCSC Genome Browser view of the loci containing the selected lncRNAs and target genes for validation**. View of the *EPB41L4A-AS1* - *EPB41L4A* **(A)**, *LINC00938* - *ARID2* **(B)**, *MIR9-3HG* - *POLG* **(C)** and *NET1e* - *NET1* - *CALML5* **(D)** loci. In each view, the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



946

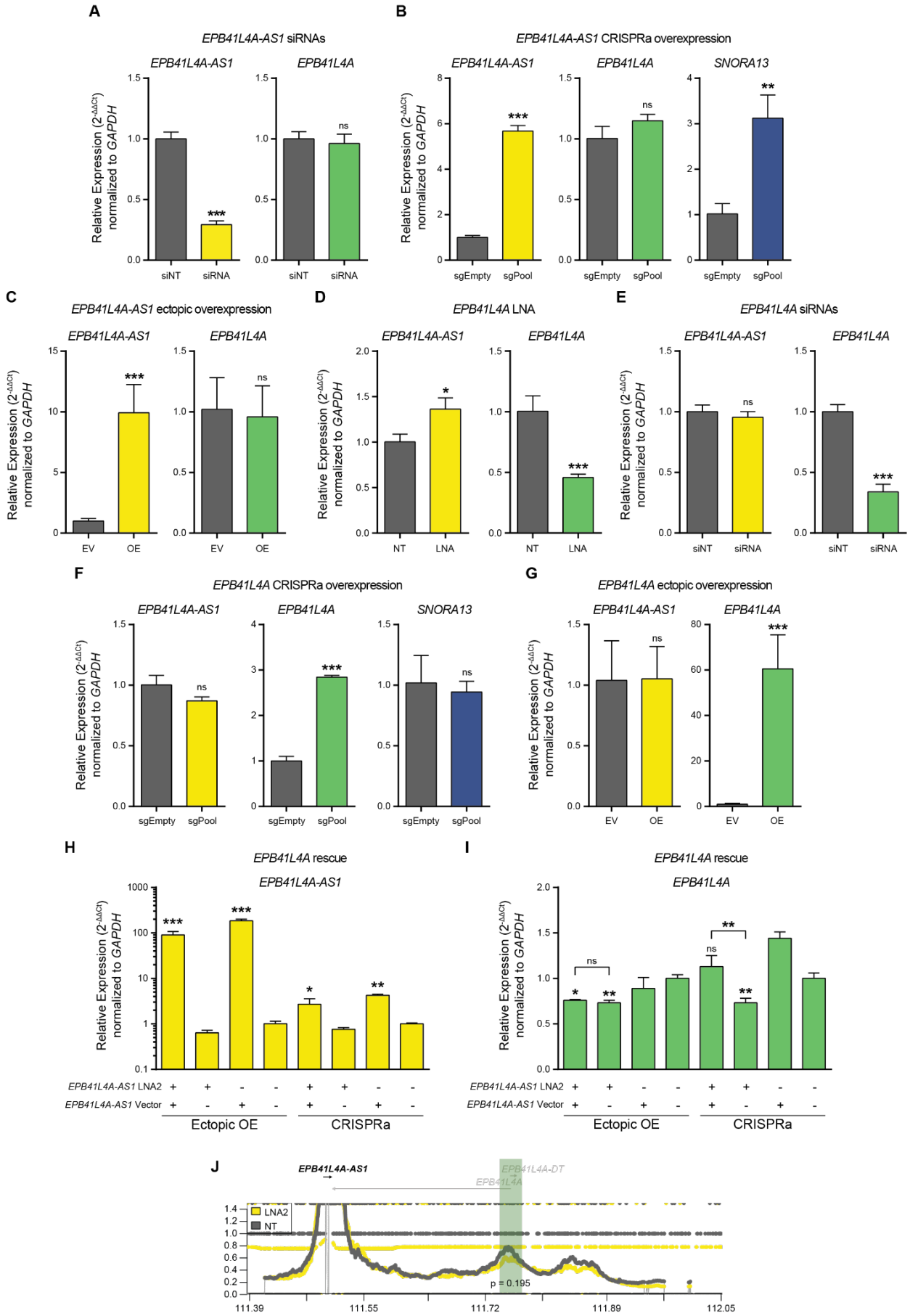
Supplementary Figure 2. ***EPB41L4A-AS1* is dysregulated in several cancer types and correlates with survival**. **(A)** Expression level of *EPB41L4A-AS1* (left) and *EPB41L4A* (right) across the TCGA cohort. The thick lines and each point represent the median and an individual sample, respectively, 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[109].



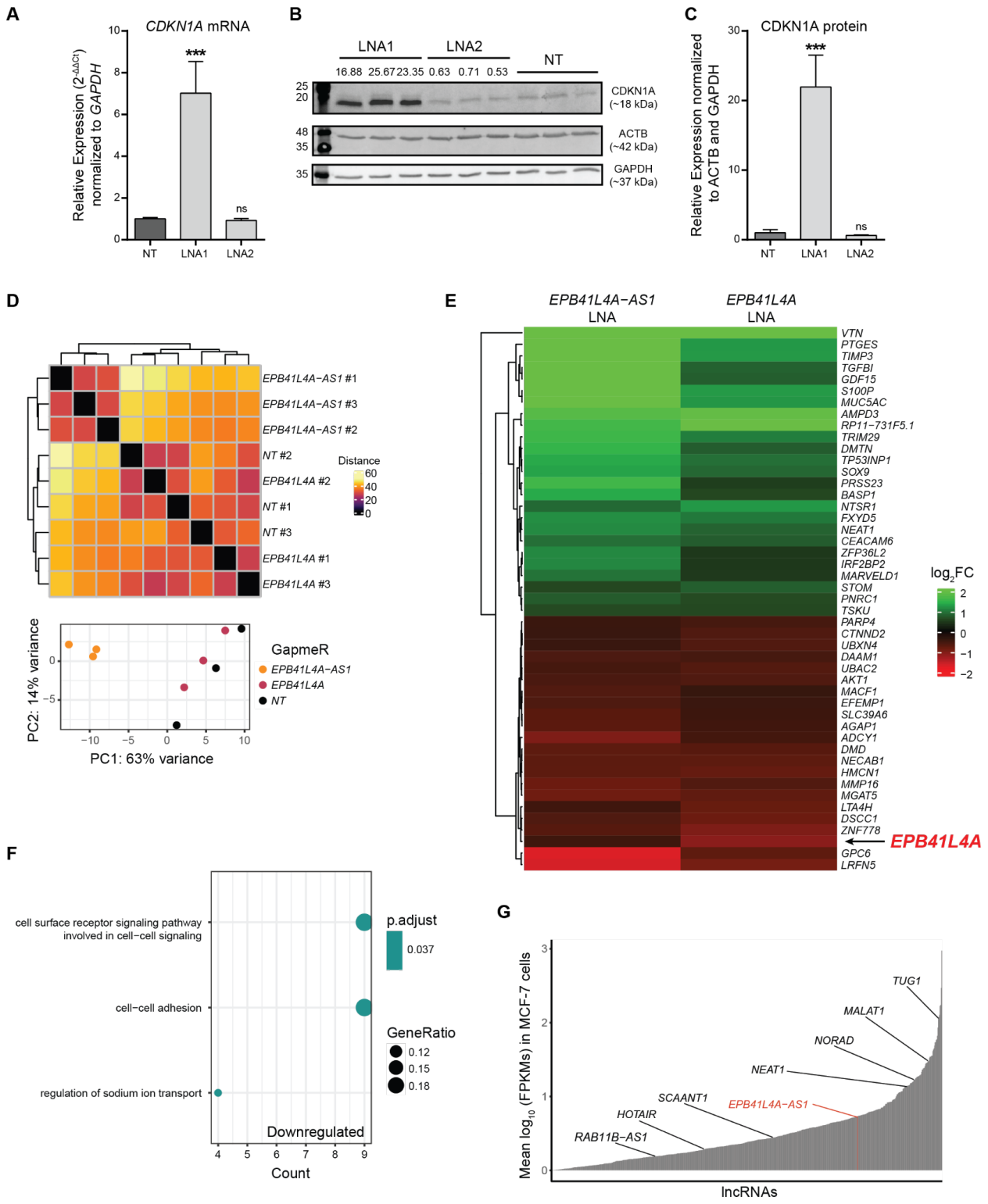
Supplementary Figure 3. ***EPB41L4A-AS1* expression is altered upon multiple stimuli**. **(A)** CPAT[41] analysis of the indicated transcripts. As controls for a noncoding and a coding transcript, *XIST* and *ACTB* are reported. Changes in gene expression for the indicated genes **(B)** during a time 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).

949


A
EPB41L4A-AS1 siRNAs
EPB41L4A-AS1
EPB41L4A
Relative Expression (2^-ΔΔCt) normalized to GAPDH
siNT
siRNA
***
ns
B
EPB41L4A-AS1 CRISPRa overexpression
EPB41L4A-AS1
EPB41L4A
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. In both panels (Ectopic OE and CRISPRa) the "-" samples represent those transfected with the Empty Vector or sgControl. Asterisks indicate significance relative to the –/– control (transfected with both the control GapmeR and vector). **(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).

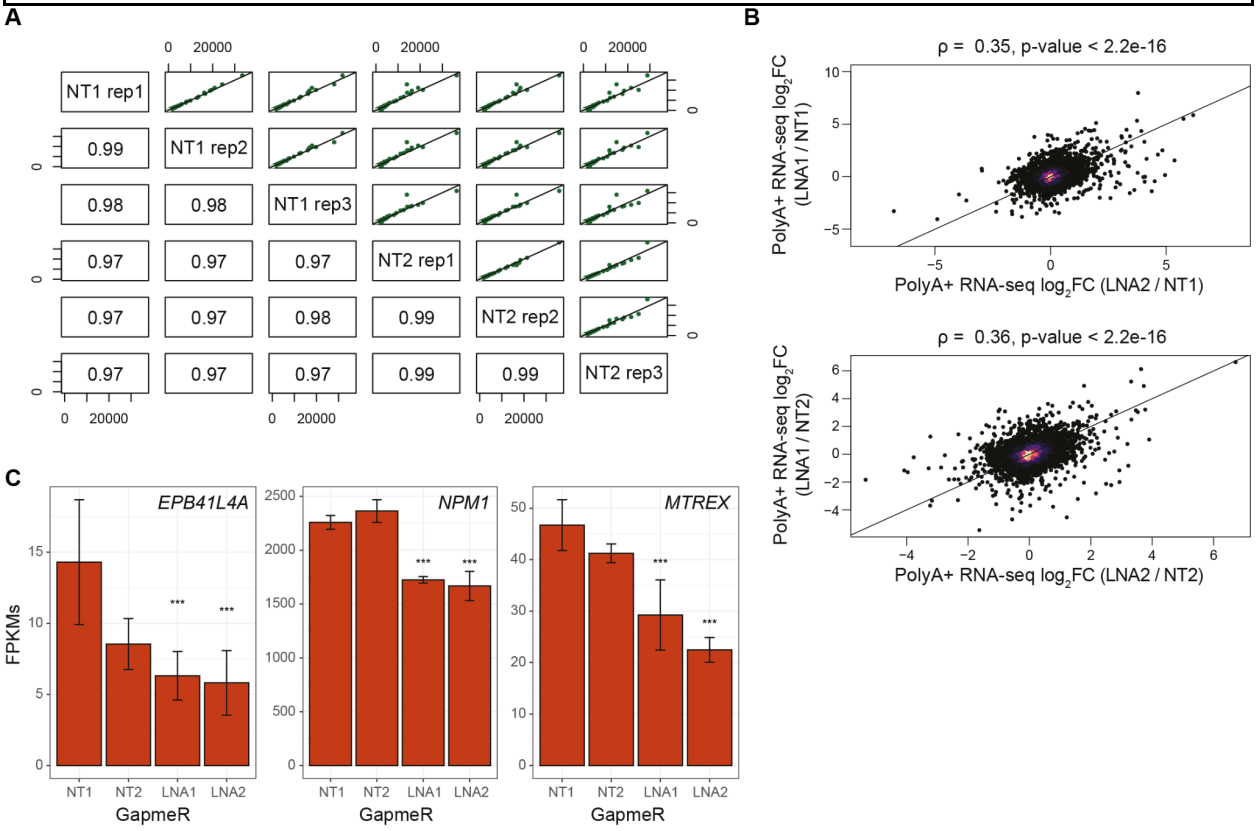
951



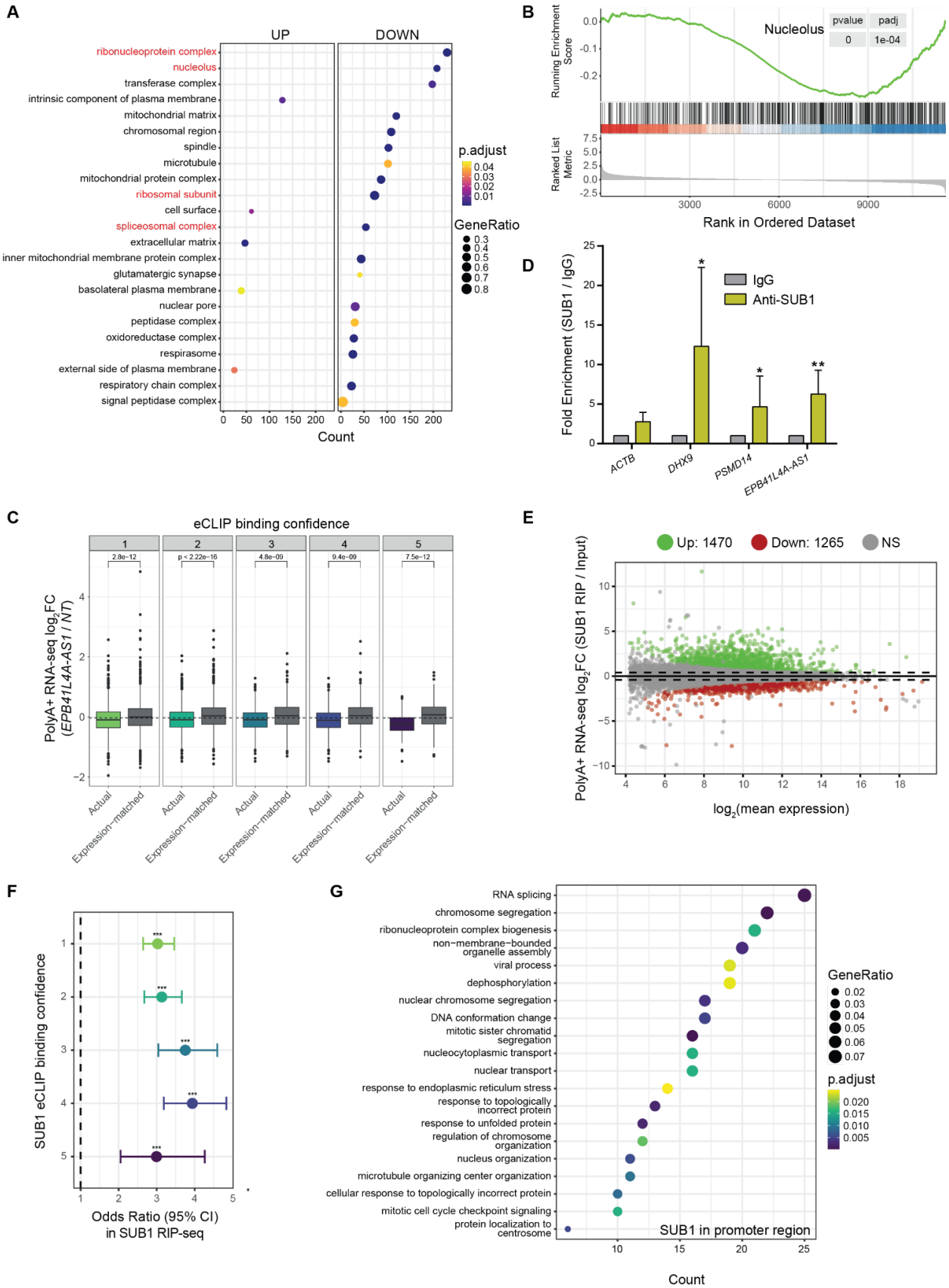
952

Supplementary Figure 5. **Most of the transcriptomic changes following *EPB41L4A-AS1* downregulation are not explained by *EPB41L4A*. (A)** RT-qPCR and **(B)** Western blot for *CDKN1A* mRNA and protein following *EPB41L4A-AS1* KD with LNA1 and LNA2. **(C)** Quantification of CDKN1A

protein levels from (B). **(D)** 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. **(E)** Heatmap of the differentially expressed genes shared between *EPB41L4A-AS1* and *EPB41L4A* depletions. Color intensity reflects the changes in gene expression (log_2FC). **(F)** GO enrichment analysis of the genes downregulated following *EPB41L4A* depletion. **(G)** 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, 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_2Fold-change| >0.41 (corresponding to a change of 33%).



Supplementary Figure 6. **LNA1 and LNA2 induce a similar transcriptional response. (A)** Correspondence of the gene expression levels (in FPKMs) between each individual replicate of the two control GapmeRs. The values depict the Spearman's correlation coefficients of each pairwise correlation. **(B)** Correspondence between changes in gene expression after *EPB41L4A-AS1* KD using either LNA1 or LNA2, normalized on either NT1 (top) or NT2 (bottom). **(C)** Average expression levels (in FPKMs) of *EPB41L4A*, *NPM1* and *MTREX* in cells transfected with the two GapmeRs targeting *EPB41L4A-AS1*, with the error bars representing the standard deviation across the n = 3 replicates. DESeq2 adjusted P-values compared to control GapmeR #1 (NT1) are also reported. All experiments were performed in n=3 biological replicates, with the error bars in the barplots representing the standard deviation. ns - P>0.05; * - P<0.05; ** - P<0.01; *** - P<0.001.


A
UP
DOWN
ribonucleoprotein complex
nucleolus
transferase complex
intrinsic component of plasma membrane
mitochondrial matrix
chromosomal region
spindle
microtubule
mitochondrial protein complex
ribosomal subunit
cell surface
spliceosomal complex
extracellular matrix
inner mitochondrial membrane protein complex
glutamatergic synapse
basolateral plasma membrane
nuclear pore
peptidase complex
oxidoreductase complex
respirasome
external side of plasma membrane
respiratory chain complex
signal peptidase complex
Count
p.adjust
GeneRatio
B
Nucleolus
pvalue
padj
0
1e-04
Running Enrichment Score
Ranked List Metric
Rank in Ordered Dataset
C
eCLIP binding confidence
PolyA+ RNA-seq log2FC (EPB41L4A-AS1 / NT)
Actual
Expression-matched
D
IgG
Anti-SUB1
Fold Enrichment (SUB1 / IgG)
ACTB
DHX9
PSMD14
EPB41L4A-AS1
E
Up: 1470
Down: 1265
NS
PolyA+ RNA-seq log2FC (SUB1 RIP / Input)
log2(mean expression)
F
SUB1 eCLIP binding confidence
Odds Ratio (95% CI) in SUB1 RIP-seq
G
RNA splicing
chromosome segregation
ribonucleoprotein complex biogenesis
non-membrane-bounded organelle assembly
viral process
dephosphorylation
nuclear chromosome segregation
DNA conformation change
mitotic sister chromatid segregation
nucleocytoplasmic transport
nuclear transport
response to endoplasmic reticulum stress
response to topologically incorrect protein
response to unfolded protein
regulation of chromosome organization
nucleus organization
microtubule organizing center organization
cellular response to topologically incorrect protein
mitotic cell cycle checkpoint signaling
protein localization to centrosome
SUB1 in promoter region
Count

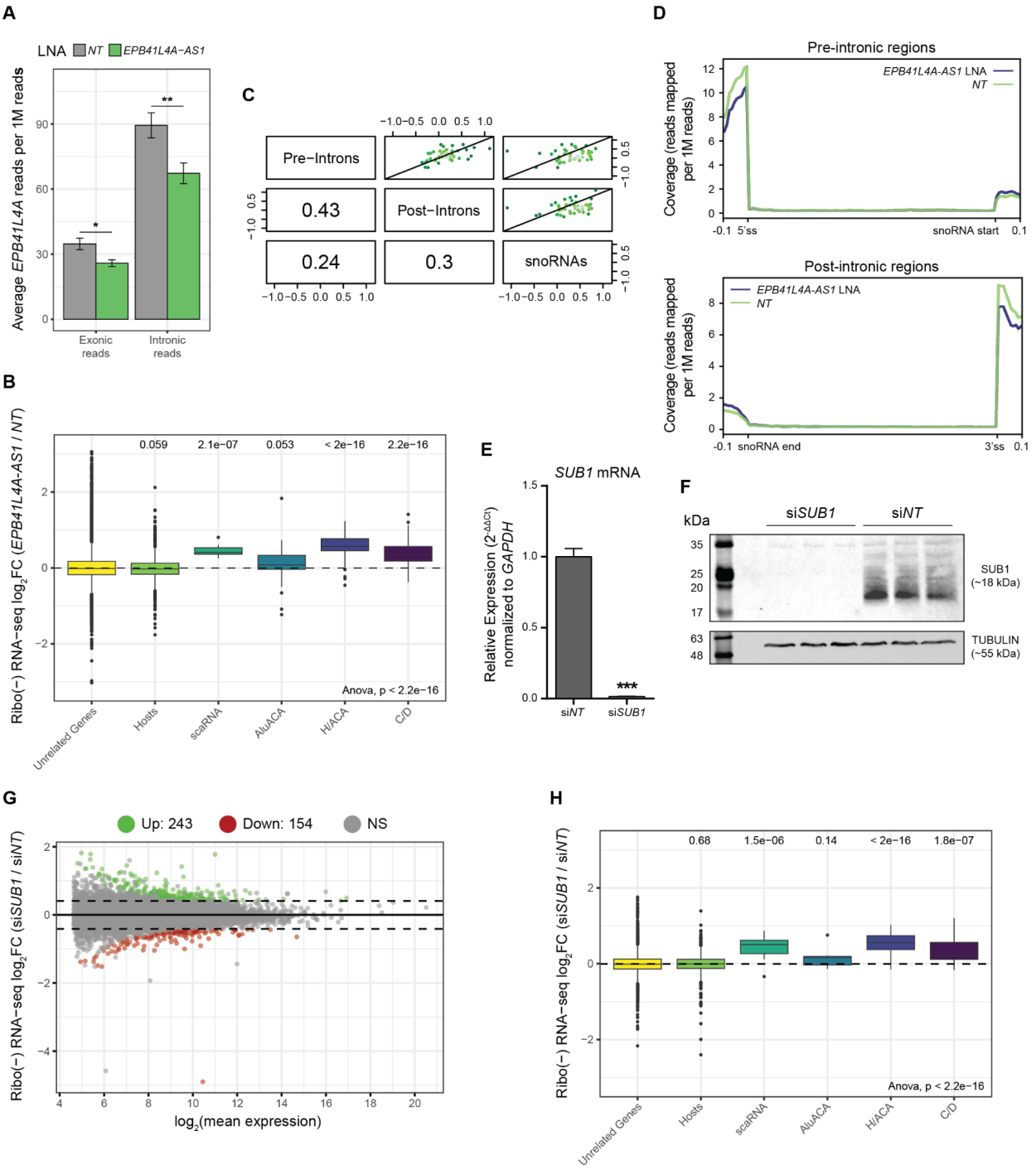
955

Supplementary Figure 7. **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)** Changes in gene expression upon *EPB41L4A-AS1* KD of the genes ranked by the SUB1 eCLIP binding confidence, for both the actual targets (as in Fig. 3C) and expression-matched controls (in grey). **(D)** RT-qPCR for the indicated genes after RIP with either an anti-SUB1 or control IgG antibody. **(E)** MA plot showing the relative enrichment in RIP-seq after using either an anti-SUB1 or control IgG antibody. **(F)** 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. **(G)** 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). 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. 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)).

956

957

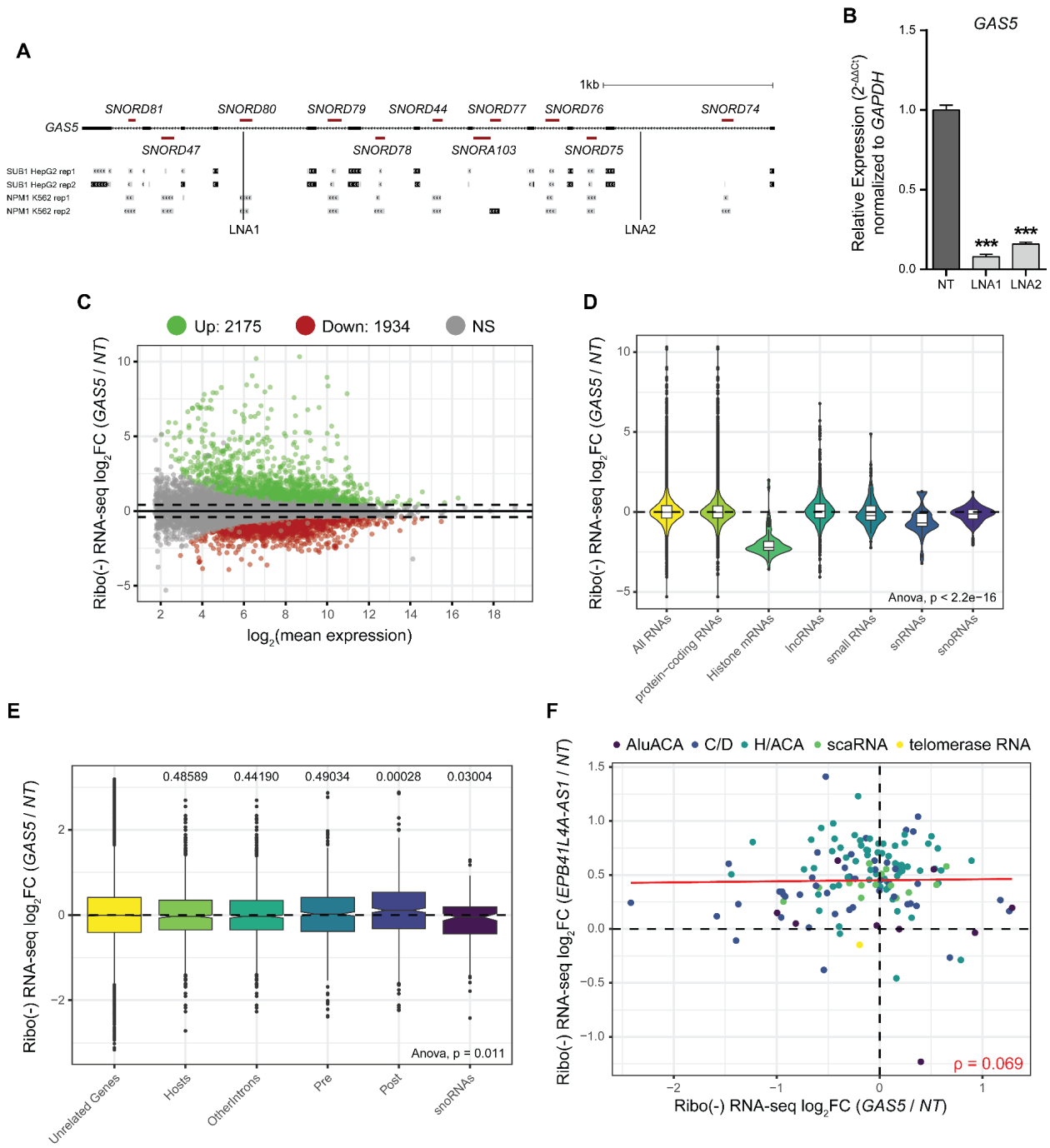
958



959

Supplementary Figure 8. ***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_2FC) 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 $|\log_2 \text{Fold-change}| > 0.41$ (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).

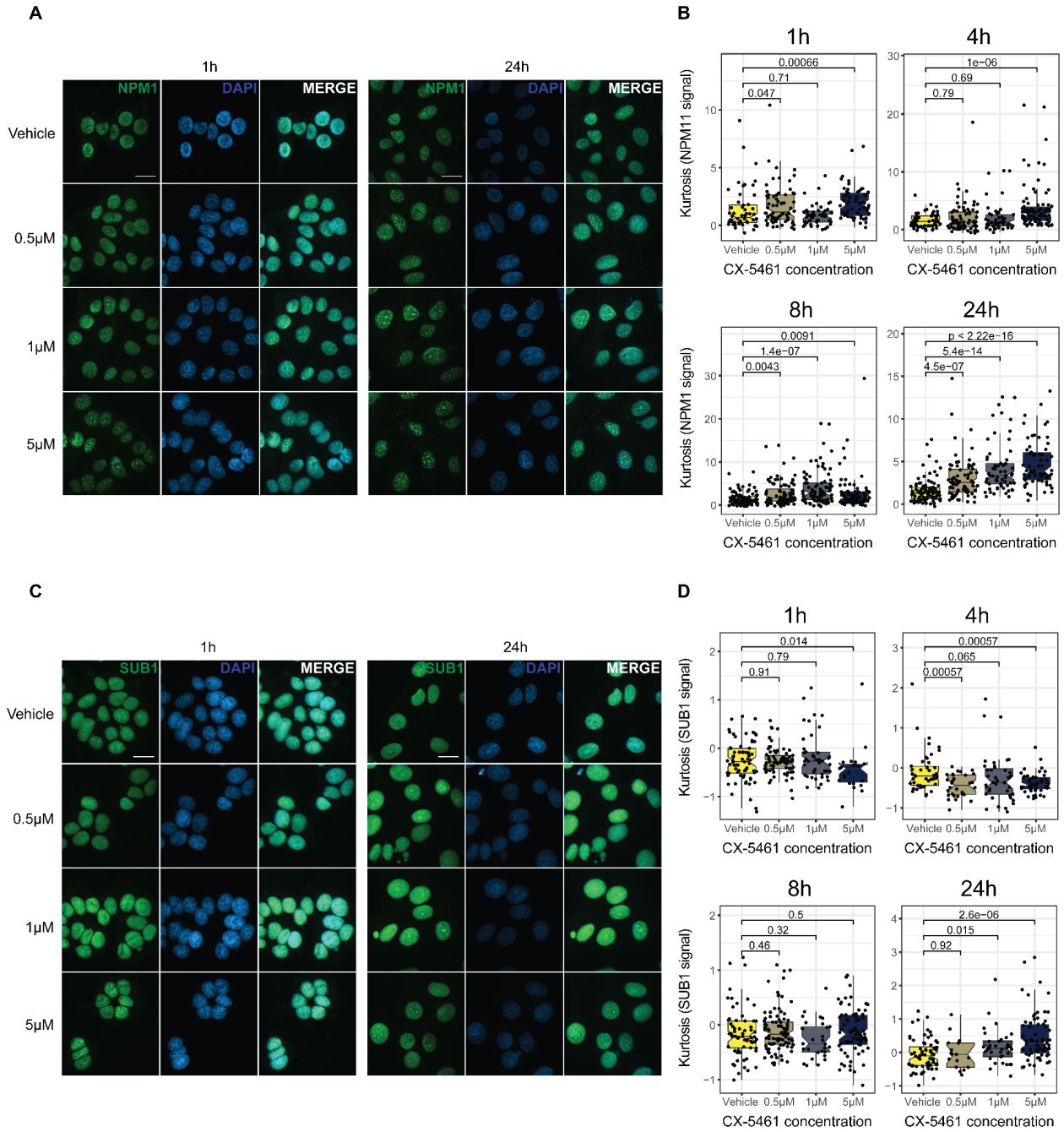


960

Supplementary Figure 9. ***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 ($\log_2$FC) after *GAS5* depletion for the indicated RNA classes. **(E)** Boxplot of gene expression changes ($\log_2$FC) 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 $|\log_2 \text{Fold-change}| > 0.41$ (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).

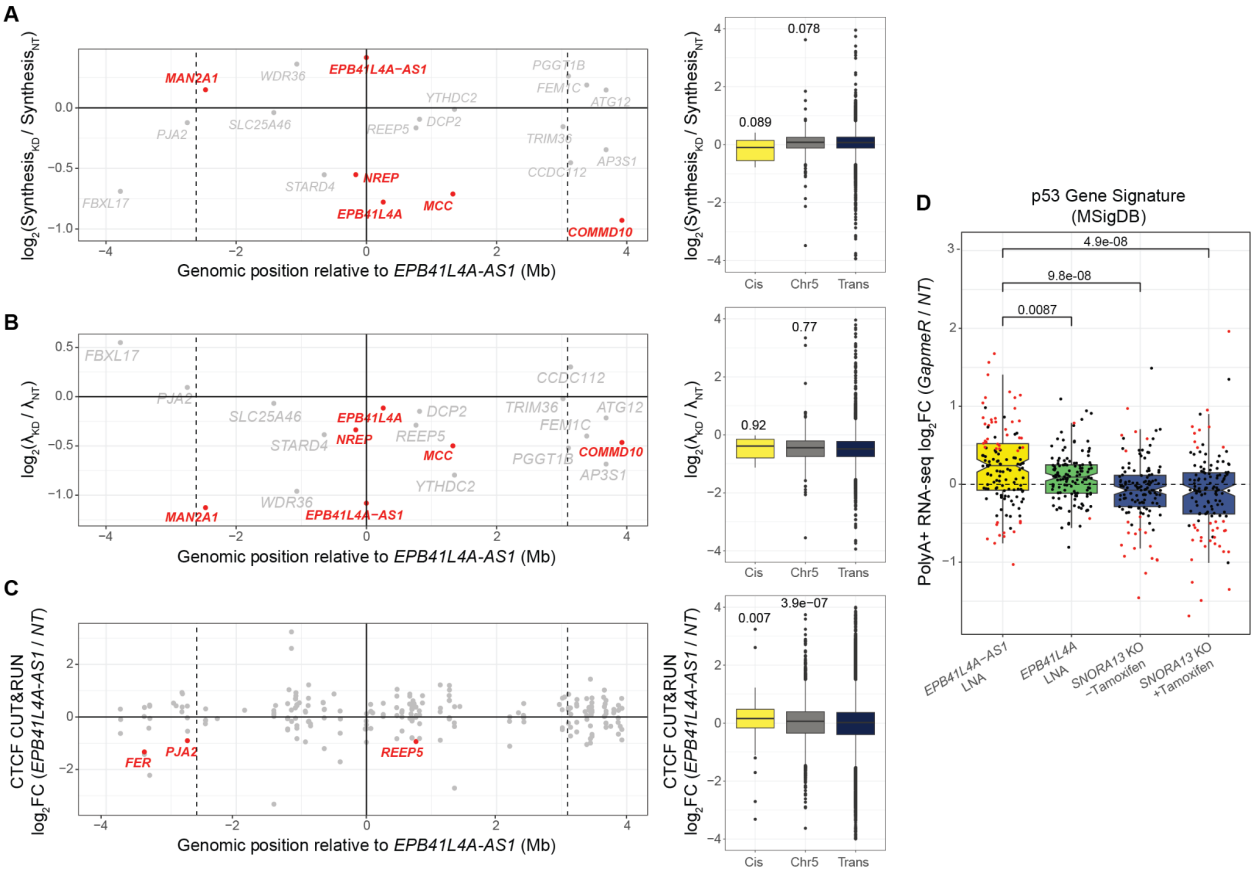
961



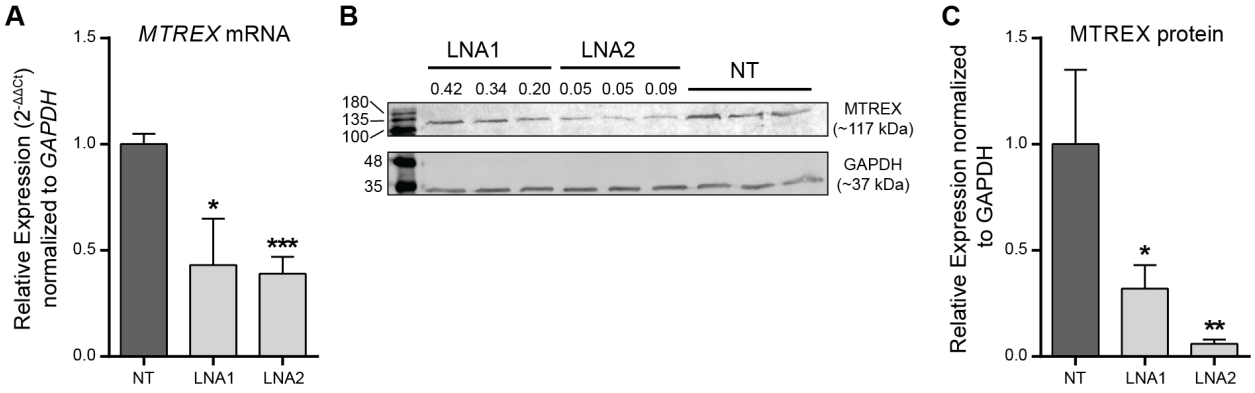
962

Supplementary Figure 10. **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 11. ***EPB41L4A-AS1* affects local RNA metabolism. (A)** Changes in synthesis rate in KD cells of the genes in the two flanking 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.



Supplementary Figure 12. ***EPB41L4A-AS1* depletion consistently represses MTREX. (A)** RT-qPCR and **(B)** Western blot for *MTREX* mRNA and protein following *EPB41L4A-AS1* KD with LNA1 and LNA2. **(C)** Quantification of MTREX protein levels from (B). All experiments were performed in n=3 biological replicates, 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).

References

1. Statello, L., Guo, C.-J., Chen, L.-L. & Huarte, M. Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* **22**, 96–118 (2021).
2. Wu, H., Yang, L. & Chen, L.-L. The Diversity of Long Noncoding RNAs and Their Generation. *Trends Genet.* **33**, 540–552 (2017).
3. Ransohoff, J. D., Wei, Y. & Khavari, P. A. The functions and unique features of long intergenic non-coding RNA. *Nat. Rev. Mol. Cell Biol.* **19**, 143–157 (2018).
4. Mattick, J. S. *et al.* Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* **24**, 430–447 (2023).
5. Hangauer, M. J., Vaughn, I. W. & McManus, M. T. Pervasive transcription of the human genome produces thousands of previously unidentified long intergenic noncoding RNAs.

982 *PLoS Genet.* **9**, e1003569 (2013).
983 6. Carninci, P. *et al.* The transcriptional landscape of the mammalian genome. *Science* **309**,
984 1559–1563 (2005).
985 7. ENCODE Project Consortium *et al.* Identification and analysis of functional elements in 1%
986 of the human genome by the ENCODE pilot project. *Nature* **447**, 799–816 (2007).
987 8. Kapranov, P. *et al.* RNA maps reveal new RNA classes and a possible function for
988 pervasive transcription. *Science* **316**, 1484–1488 (2007).
989 9. Mudge, J. M. *et al.* GENCODE 2025: reference gene annotation for human and mouse.
990 *Nucleic Acids Res.* (2024) doi:10.1093/nar/gkae1078.
991 10. Ruiz-Orera, J., Messeguer, X., Subirana, J. A. & Alba, M. M. Long non-coding RNAs as a
992 source of new peptides. *Elife* **3**, e03523 (2014).
993 11. Chen, J. *et al.* Pervasive functional translation of noncanonical human open reading
994 frames. *Science* **367**, 1140–1146 (2020).
995 12. Barczak, W. *et al.* Long non-coding RNA-derived peptides are immunogenic and drive a
996 potent anti-tumour response. *Nat. Commun.* **14**, 1078 (2023).
997 13. Kesner, J. S. *et al.* Noncoding translation mitigation. *Nature* **617**, 395–402 (2023).
998 14. Chothani, S. P. *et al.* A high-resolution map of human RNA translation. *Mol. Cell* **82**, 2885–
999 2899.e8 (2022).
1000 15. Housman, G. & Ulitsky, I. Methods for distinguishing between protein-coding and long
1001 noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs.
1002 *Biochim. Biophys. Acta* **1859**, 31–40 (2016).
1003 16. Paralkar, V. R. *et al.* Unlinking an lncRNA from Its Associated cis Element. *Mol. Cell* **62**,
1004 104–110 (2016).
1005 17. Ali, T. & Grote, P. Beyond the RNA-dependent function of LncRNA genes. *Elife* **9**, (2020).
1006 18. Beucher, A. *et al.* The HASTER lncRNA promoter is a cis-acting transcriptional stabilizer of
1007 HNF1A. *Nat. Cell Biol.* **24**, 1528–1540 (2022).

1008 19. De Santa, F. *et al.* A large fraction of extragenic RNA pol II transcription sites overlap
1009 enhancers. *PLoS Biol.* **8**, e1000384 (2010).
1010 20. Gil, N. & Ulitsky, I. Regulation of gene expression by cis-acting long non-coding RNAs. *Nat.*
1011 *Rev. Genet.* **21**, 102–117 (2020).
1012 21. Monziani, A. & Ulitsky, I. Noncoding snoRNA host genes are a distinct subclass of long
1013 noncoding RNAs. *Trends Genet.* (2023) doi:10.1016/j.tig.2023.09.001.
1014 22. Fishilevich, S. *et al.* GeneHancer: genome-wide integration of enhancers and target genes
1015 in GeneCards. *Database* **2017**, (2017).
1016 23. Sun, M., Gadad, S. S., Kim, D.-S. & Kraus, W. L. Discovery, Annotation, and Functional
1017 Analysis of Long Noncoding RNAs Controlling Cell-Cycle Gene Expression and
1018 Proliferation in Breast Cancer Cells. *Mol. Cell* **59**, 698–711 (2015).
1019 24. Nagarajan, S. *et al.* ARID1A influences HDAC1/BRD4 activity, intrinsic proliferative capacity
1020 and breast cancer treatment response. *Nat. Genet.* **52**, 187–197 (2020).
1021 25. Janky, R. 's *et al.* iRegulon: from a gene list to a gene regulatory network using large motif
1022 and track collections. *PLoS Comput. Biol.* **10**, e1003731 (2014).
1023 26. Babaian, A. *et al.* Loss of m1acp3Ψ Ribosomal RNA Modification Is a Major Feature of
1024 Cancer. *Cell Rep.* **31**, 107611 (2020).
1025 27. Cheng, Y. *et al.* A non-canonical role for a small nucleolar RNA in ribosome biogenesis and
1026 senescence. *Cell* **187**, 4770–4789.e23 (2024).
1027 28. Hegre, S. A. *et al.* Joint changes in RNA, RNA polymerase II, and promoter activity through
1028 the cell cycle identify non-coding RNAs involved in proliferation. *Sci. Rep.* **11**, 18952
1029 (2021).
1030 29. Liao, M. *et al.* LncRNA EPB41L4A-AS1 regulates glycolysis and glutaminolysis by
1031 mediating nucleolar translocation of HDAC2. *EBioMedicine* **41**, 200–213 (2019).
1032 30. Zhu, Y. *et al.* Overexpression of lncRNA EPB41L4A-AS1 Induces Metabolic
1033 Reprogramming in Trophoblast Cells and Placenta Tissue of Miscarriage. *Mol. Ther.*

1034 *Nucleic Acids* **18**, 518–532 (2019).
1035 31. Liao, W. *et al.* Persistent high glucose induced EPB41L4A-AS1 inhibits glucose uptake via
1036 GCN5 mediating crotonylation and acetylation of histones and non-histones. *Clin. Transl.*
1037 *Med.* **12**, e699 (2022).
1038 32. Bin, J. *et al.* Long noncoding RNA EPB41L4A-AS1 functions as an oncogene by regulating
1039 the Rho/ROCK pathway in colorectal cancer. *J. Cell. Physiol.* **236**, 523–535 (2021).
1040 33. Wang, Z. *et al.* Downregulation of lncRNA EPB41L4A-AS1 Mediates Activation of MYD88-
1041 Dependent NF-κB Pathway in Diabetes-Related Inflammation. *Diabetes Metab. Syndr.*
1042 *Obes.* **14**, 265–277 (2021).
1043 34. Ishiguro, H. *et al.* Isolation and characterization of human NBL4, a gene involved in the
1044 beta-catenin/tcf signaling pathway. *Jpn. J. Cancer Res.* **91**, 597–603 (2000).
1045 35. Comiskey, D. F., Jr *et al.* Characterizing the function of EPB41L4A in the predisposition to
1046 papillary thyroid carcinoma. *Sci. Rep.* **10**, 19984 (2020).
1047 36. Guo, Y., Christine, K. S., Conlon, F., Gessert, S. & Kühl, M. Expression analysis of
1048 epb41l4a during Xenopus laevis embryogenesis. *Dev. Genes Evol.* **221**, 113–119 (2011).
1049 37. Didiot, M.-C. *et al.* Nuclear Localization of Huntingtin mRNA Is Specific to Cells of Neuronal
1050 Origin. *Cell Rep* **24**, 2553–2560.e5 (2018).
1051 38. Ly, S. *et al.* Mutant huntingtin messenger RNA forms neuronal nuclear clusters in rodent
1052 and human brains. *Brain Commun* **4**, fcac248 (2022).
1053 39. Bahar Halpern, K. *et al.* Nuclear Retention of mRNA in Mammalian Tissues. *Cell Rep* **13**,
1054 2653–2662 (2015).
1055 40. Lubelsky, Y. & Ulitsky, I. Sequences enriched in Alu repeats drive nuclear localization of
1056 long RNAs in human cells. *Nature* **555**, 107–111 (2018).
1057 41. Wang, L. *et al.* CPAT: Coding-Potential Assessment Tool using an alignment-free logistic
1058 regression model. *Nucleic Acids Res* **41**, e74 (2013).
1059 42. Yabuta, N. *et al.* Isolation and characterization of the TIGA genes, whose transcripts are

1060 induced by growth arrest. *Nucleic Acids Res.* **34**, 4878–4892 (2006).
1061 43. Moyerbrailean, G. A. *et al.* High-throughput allele-specific expression across 250
1062 environmental conditions. *Genome Res.* **26**, 1627–1638 (2016).
1063 44. Fafard-Couture, É., Bergeron, D., Couture, S., Abou-Elela, S. & Scott, M. S. Annotation of
1064 snoRNA abundance across human tissues reveals complex snoRNA-host gene
1065 relationships. *Genome Biol* **22**, 172 (2021).
1066 45. Schwartzman, O. *et al.* UMI-4C for quantitative and targeted chromosomal contact profiling.
1067 *Nat. Methods* **13**, 685–691 (2016).
1068 46. Vian, L. *et al.* The energetics and physiological impact of cohesin extrusion. *Cell* **175**, 292–
1069 294 (2018).
1070 47. Salnikov, P. *et al.* Structural variants in the Epb41l4a locus: TAD disruption and Nrep gene
1071 misregulation as hypothetical drivers of neurodevelopmental outcomes. *Sci Rep* **14**, 5288
1072 (2024).
1073 48. Krietenstein, N. *et al.* Ultrastructural Details of Mammalian Chromosome Architecture. *Mol.*
1074 *Cell* **78**, 554–565.e7 (2020).
1075 49. Singh, R. & Berger, B. Deciphering the species-level structure of topologically associating
1076 domains. *bioRxiv* (2021) doi:10.1101/2021.10.28.466333.
1077 50. Liang, X.-H., Vickers, T. A., Guo, S. & Crooke, S. T. Efficient and specific knockdown of
1078 small non-coding RNAs in mammalian cells and in mice. *Nucleic Acids Res.* **39**, e13
1079 (2011).
1080 51. Hagedorn, P. H. *et al.* Locked nucleic acid: modality, diversity, and drug discovery. *Drug*
1081 *Discov. Today* **23**, 101–114 (2018).
1082 52. Maranon, D. G. & Wilusz, J. Mind the gapmer: Implications of co-transcriptional cleavage
1083 by antisense oligonucleotides. *Mol. Cell* **77**, 932–933 (2020).
1084 53. Lee, J.-S. & Mendell, J. T. Antisense-Mediated Transcript Knockdown Triggers Premature
1085 Transcription Termination. *Mol. Cell* **77**, 1044–1054.e3 (2020).

1086 54. Lai, F., Damle, S. S., Ling, K. K. & Rigo, F. Directed RNase H Cleavage of Nascent
1087 Transcripts Causes Transcription Termination. *Mol. Cell* **77**, 1032–1043.e4 (2020).
1088 55. Dhaka, B. *et al.* Functional identification of cis-regulatory long noncoding RNAs at
1089 controlled false discovery rates. *Nucleic Acids Res.* **52**, 2821–2835 (2024).
1090 56. Van Nostrand, E. L. *et al.* A large-scale binding and functional map of human RNA-binding
1091 proteins. *Nature* **583**, 711–719 (2020).
1092 57. Garavís, M. & Calvo, O. Sub1/PC4, a multifaceted factor: from transcription to genome
1093 stability. *Curr. Genet.* **63**, 1023–1035 (2017).
1094 58. Conesa, C. & Acker, J. Sub1/PC4 a chromatin associated protein with multiple functions in
1095 transcription. *RNA Biol.* **7**, 287–290 (2010).
1096 59. Das, C. *et al.* Transcriptional coactivator PC4, a chromatin-associated protein, induces
1097 chromatin condensation. *Mol. Cell. Biol.* **26**, 8303–8315 (2006).
1098 60. Hou, Y. *et al.* G-quadruplex inducer/stabilizer pyridostatin targets SUB1 to promote
1099 cytotoxicity of a transplatinum complex. *Nucleic Acids Res.* **50**, 3070–3082 (2022).
1100 61. Dubois, J.-C. *et al.* The single-stranded DNA-binding factor SUB1/PC4 alleviates replication
1101 stress at telomeres and is a vulnerability of ALT cancer cells. *Proc Natl Acad Sci U S A*
1102 **122**, e2419712122 (2025).
1103 62. Salgado, S. *et al.* Human PC4 supports telomere stability and viability in cells utilizing the
1104 alternative lengthening of telomeres mechanism. *EMBO Rep* **25**, 5294–5315 (2024).
1105 63. Kaypee, S. *et al.* Positive coactivator PC4 shows dynamic nucleolar distribution required for
1106 rDNA transcription and protein synthesis. *Cell Commun Signal* **23**, 283 (2025).
1107 64. Nachmani, D. *et al.* Germline NPM1 mutations lead to altered rRNA 2'-O-methylation and
1108 cause dyskeratosis congenita. *Nat. Genet.* **51**, 1518–1529 (2019).
1109 65. Mourtada-Maarabouni, M., Hedge, V. L., Kirkham, L., Farzaneh, F. & Williams, G. T.
1110 Growth arrest in human T-cells is controlled by the non-coding RNA growth-arrest-specific
1111 transcript 5 (GAS5). *J. Cell Sci.* **121**, 939–946 (2008).

1112 66. Schneider, C., King, R. M. & Philipson, L. Genes specifically expressed at growth arrest of
1113 mammalian cells. *Cell* **54**, 787–793 (1988).
1114 67. Mourtada-Maarabouni, M., Pickard, M. R., Hedge, V. L., Farzaneh, F. & Williams, G. T.
1115 GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast
1116 cancer. *Oncogene* **28**, 195–208 (2009).
1117 68. Yacqub-Usman, K., Pickard, M. R. & Williams, G. T. Reciprocal regulation of GAS5 lncRNA
1118 levels and mTOR inhibitor action in prostate cancer cells. *Prostate* **75**, 693–705 (2015).
1119 69. Kino, T., Hurt, D. E., Ichijo, T., Nader, N. & Chrousos, G. P. Noncoding RNA gas5 is a
1120 growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci.*
1121 *Signal.* **3**, ra8 (2010).
1122 70. Zhang, Z. *et al.* Negative regulation of lncRNA GAS5 by miR-21. *Cell Death Differ.* **20**,
1123 1558–1568 (2013).
1124 71. Hu, G., Lou, Z. & Gupta, M. The long non-coding RNA GAS5 cooperates with the
1125 eukaryotic translation initiation factor 4E to regulate c-Myc translation. *PLoS One* **9**,
1126 e107016 (2014).
1127 72. Sun, D. *et al.* LncRNA GAS5 inhibits microglial M2 polarization and exacerbates
1128 demyelination. *EMBO Rep.* **18**, 1801–1816 (2017).
1129 73. He, X. *et al.* An Lnc RNA (GAS5)/SnoRNA-derived piRNA induces activation of TRAIL gene
1130 by site-specifically recruiting MLL/COMPASS-like complexes. *Nucleic Acids Res.* **43**, 3712–
1131 3725 (2015).
1132 74. Potapova, T. A. *et al.* Distinct states of nucleolar stress induced by anticancer drugs. (2023)
1133 doi:10.7554/eLife.88799.
1134 75. Yang, K. *et al.* A redox mechanism underlying nucleolar stress sensing by nucleophosmin.
1135 *Nat. Commun.* **7**, 13599 (2016).
1136 76. Haddach, M. *et al.* Discovery of CX-5461, the first direct and selective inhibitor of RNA
1137 polymerase I, for cancer therapeutics. *ACS Med. Chem. Lett.* **3**, 602–606 (2012).

1138 77. Xu, H. *et al.* CX-5461 is a DNA G-quadruplex stabilizer with selective lethality in BRCA1/2
1139 deficient tumours. *Nat. Commun.* **8**, 14432 (2017).
1140 78. Mars, J.-C. *et al.* The chemotherapeutic agent CX-5461 irreversibly blocks RNA
1141 polymerase I initiation and promoter release to cause nucleolar disruption, DNA damage
1142 and cell inviability. *NAR Cancer* **2**, zcaa032 (2020).
1143 79. Hilton, J. *et al.* Results of the phase I CCTG IND.231 trial of CX-5461 in patients with
1144 advanced solid tumors enriched for DNA-repair deficiencies. *Nat. Commun.* **13**, 3607
1145 (2022).
1146 80. Herzog, V. A. *et al.* Thiol-linked alkylation of RNA to assess expression dynamics. *Nat*
1147 *Methods* **14**, 1198–1204 (2017).
1148 81. Youssef, K. K. *et al.* Two distinct epithelial-to-mesenchymal transition programs control
1149 invasion and inflammation in segregated tumor cell populations. *Nat. Cancer* **5**, 1660–1680
1150 (2024).
1151 82. Open2C *et al.* Cooltools: Enabling high-resolution Hi-C analysis in Python. *PLoS Comput.*
1152 *Biol.* **20**, e1012067 (2024).
1153 83. Contreras, X. *et al.* PAPγ associates with PAXT nuclear exosome to control the abundance
1154 of PROMPT ncRNAs. *Nat Commun* **14**, 6745 (2023).
1155 84. Yu, L., Ma, H., Ji, X. & Volkert, M. R. The Sub1 nuclear protein protects DNA from oxidative
1156 damage. *Mol. Cell. Biochem.* **412**, 165–171 (2016).
1157 85. Mortusewicz, O. *et al.* Recruitment of RNA polymerase II cofactor PC4 to DNA damage
1158 sites. *J. Cell Biol.* **183**, 769–776 (2008).
1159 86. Tafforeau, L. *et al.* The complexity of human ribosome biogenesis revealed by systematic
1160 nucleolar screening of Pre-rRNA processing factors. *Mol. Cell* **51**, 539–551 (2013).
1161 87. Dongre, A. & Weinberg, R. A. New insights into the mechanisms of epithelial-mesenchymal
1162 transition and implications for cancer. *Nat Rev Mol Cell Biol* **20**, 69–84 (2019).
1163 88. Brabletz, T., Kalluri, R., Nieto, M. A. & Weinberg, R. A. EMT in cancer. *Nat Rev Cancer* **18**,

1164 128–134 (2018).
1165 89. Wang, M. *et al.* Integrated analysis of lncRNA-miRNA-mRNA ceRNA network identified
1166 lncRNA EPB41L4A-AS1 as a potential biomarker in non-small cell lung cancer. *Front.*
1167 *Genet.* **11**, 511676 (2020).
1168 90. Yang, T. *et al.* Down-regulation of EPB41L4A-AS1 mediated the brain aging and
1169 neurodegenerative diseases via damaging synthesis of NAD+ and ATP. *Cell Biosci.* **11**,
1170 192 (2021).
1171 91. Wang, Z. *et al.* EPB41L4A-AS1 is required to maintain basal autophagy to modulates Aβ
1172 clearance. *NPJ Aging* **10**, 24 (2024).
1173 92. Samdal, H. *et al.* The lncRNA EPB41L4A-AS1 regulates gene expression in the nucleus
1174 and exerts cell type-dependent effects on cell cycle progression. *bioRxiv*
1175 2021.02.10.430566 (2021) doi:10.1101/2021.02.10.430566.
1176 93. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-
1177 Cas9 complex. *Nature* **517**, 583–588 (2015).
1178 94. Monziani, A., Ben-Tov Perry, R., Hezroni, H. & Ulitsky, I. BAHCC1 promotes gene
1179 expression in neuronal cells by antagonizing SIN3A-HDAC1. *bioRxiv* (2025)
1180 doi:10.1101/2025.01.24.634719.
1181 95. Suarez-Arnedo, A. *et al.* An image J plugin for the high throughput image analysis of in vitro
1182 scratch wound healing assays. *PLoS One* **15**, e0232565 (2020).
1183 96. Gagnon, K. T., Li, L., Janowski, B. A. & Corey, D. R. Analysis of nuclear RNA interference
1184 in human cells by subcellular fractionation and Argonaute loading. *Nat. Protoc.* **9**, 2045–
1185 2060 (2014).
1186 97. Sundararaman, B. *et al.* Resources for the Comprehensive Discovery of Functional RNA
1187 Elements. *Mol. Cell* **61**, 903–913 (2016).
1188 98. Meers, M. P., Bryson, T. D., Henikoff, J. G. & Henikoff, S. Improved CUT&RUN chromatin
1189 profiling tools. *Elife* **8**, (2019).

1190 99. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
1191 *EMBnet J.* **17**, 10 (2011).
1192 100. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
1193 101. Love, M., Anders, S. & Huber, W. Differential analysis of count data--the DESeq2 package.
1194 *Genome Biol.* **15**, 550 (2014).
1195 102. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data.
1196 *Innovation (Camb)* **2**, 100141 (2021).
1197 103. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*
1198 **9**, 357–359 (2012).
1199 104. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).
1200 105. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-
1201 regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589
1202 (2010).
1203 106. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data
1204 analysis. *Nucleic Acids Res.* **44**, W160–5 (2016).
1205 107. Zuckerman, B., Ron, M., Mikl, M., Segal, E. & Ulitsky, I. Gene Architecture and Sequence
1206 Composition Underpin Selective Dependency of Nuclear Export of Long RNAs on NXF1
1207 and the TREX Complex. *Mol Cell* **79**, 251–267.e6 (2020).
1208 108. Jürges, C., Dölken, L. & Erhard, F. Dissecting newly transcribed and old RNA using
1209 GRAND-SLAM. *Bioinformatics* **34**, i218–i226 (2018).
1210 109. Tang, Z., Kang, B., Li, C., Chen, T. & Zhang, Z. GEPIA2: an enhanced web server for
1211 large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* **47**, (2019).

1212