BIOCHEMISTRY

Integrator restrains paraspeckles assembly by promoting isoform switching of the lncRNA *NEAT1*

Jasmine Barra[1,2,3,4,5], Gabriel S. Gaidosh[6], Ezra Blumenthal[6], Felipe Beckedorff[6], Mina M. Tayari[6], Nina Kirstein[6], Tobias K. Karakach[7,8], Torben Heick Jensen[9], Francis Impens[3,4,10], Kris Gevaert[3,4], Eleonora Leucci[5]*, Ramin Shiekhattar[6]*, Jean-Christophe Marine[1,2]*†

RNA 3′ end processing provides a source of transcriptome diversification which affects various (patho)-physiological processes. A prime example is the transcript isoform switch that leads to the read-through expression of the long non-coding RNA *NEAT1_2*, at the expense of the shorter polyadenylated transcript *NEAT1_1*. *NEAT1_2* is required for assembly of paraspeckles (PS), nuclear bodies that protect cancer cells from oncogene-induced replication stress and chemotherapy. Searching for proteins that modulate this event, we identified factors involved in the 3′ end processing of polyadenylated RNA and components of the Integrator complex. Perturbation experiments established that, by promoting the cleavage of *NEAT1_2*, Integrator forces *NEAT1_2* to *NEAT1_1* isoform switching and, thereby, restrains PS assembly. Consistently, low levels of Integrator subunits correlated with poorer prognosis of cancer patients exposed to chemotherapeutics. Our study establishes that Integrator regulates PS biogenesis and a link between Integrator, cancer biology, and chemosensitivity, which may be exploited therapeutically.

INTRODUCTION

Most human genes have multiple sites at which RNA 3′ end cleavage and polyadenylation can occur (*1*). Alternative 3′ end cleavage gives rise to transcript isoforms that differ either in their coding sequences or in their 3′ untranslated regions (UTRs) and, thus, contribute to transcriptome diversification (*1*, *2*). Remodeling of 3′ UTRs can have particularly profound phenotypic consequences; hence, transcript isoforms may differ in their relative stability, localization, translation rate, and/or function (*2*). Although it is well known that RNA 3′ end processing can be finely regulated depending on the cellular needs, the factors involved in alternative 3′ end processing are only partially characterized.

The core of the pre-mRNA 3′ end processing complex consists of four subcomplexes, namely, cleavage and polyadenylation factor (CPSF), cleavage stimulation factor (CSTF), cleavage factor I (CFI), and CFII. Few other proteins, including symplekin and polyadenylate [poly(A)] polymerase (PAP), are also involved in completing the 3′ end formation of polyadenylated RNA (*3*). In metazoans, sites of pre-mRNA polyadenylation are primarily defined by the canonical poly(A) signal AAUAAA, which is positioned ~21-nucleotides (nt) upstream of the cleavage site (*3*). This hexamer is recognized by the cotranscriptionally recruited CPSF subcomplex, which carries out the endonucleolytic cleavage event, followed by the addition of a poly(A) tail to the 5′ cleavage product by PAP. Deregulation of protein expression levels and/or activity of core 3′ end processing factors can obviously contribute to 3′ end processing rewiring, either globally or specifically, and thereby affect transcriptome diversification in response to specific environmental cues. Moreover, many other RNA binding proteins can influence RNA 3′ end processing, often depending on the binding positions within mRNA target 3′ UTRs (*4*).

Other protein complexes are involved in the 3′ end processing of nonpolyadenylated RNA species. For instance, the Integrator complex, which binds to the C-terminal domain (CTD) of the RNA polymerase II (Pol II), is responsible for the RNA 3′ end processing of uridylate-rich small nuclear RNA transcripts (UsnRNAs) (*5*). This complex has been shown to control termination of transcription and 3′ end processing at enhancer RNAs and replication-dependent histone loci (*6*, *7*). Furthermore, Integrator binding to the proximal promoter region of polyadenylated target genes negatively regulates their expression (*7*).

Alternative 3′ end processing–dependent transcriptome diversification plays key roles in various important biological processes (*4*). Individual 3′ end processing events have also been implicated under pathological conditions, including autoimmune disorders and cancer (*4*). Consistent with a reported general association between the expression of short RNA 3′ UTRs and a proliferative cellular state (*8*), most cancers express transcripts with shorter 3′ UTRs than those expressed in corresponding normal tissues (*4*). Some studies have attributed cancer-related 3′ end RNA patterns to the deregulated activity of specific 3′ end processing factors, such as CSTF2 (*9*) and CFIm25 (*10*). However, the motifs recognized by these core 3′ end processing factors do not explain the observed quantitative changes in poly(A) site usage between tumor and normal tissue samples from The Cancer Genome Atlas (TCGA) (*11*), indicating that other unknown modulators also contribute.

Deregulation of RNA 3′ end processing at specific loci may also contribute to tumor growth as illustrated in recent findings implicating the long noncoding RNA (lncRNA) locus *NEAT1* in cancer development (*12*). This locus produces two lncRNA isoforms (*13*). The shorter isoform, *NEAT1_1* (3700 nt in length), contains a functional poly(A) site. The long *NEAT1_2* isoform (22700 nt in length), which

[1]Laboratory for Molecular Cancer Biology, Center for Cancer Biology, VIB, Leuven, Belgium. [2]Laboratory for Molecular Cancer Biology, Department of Oncology, KU Leuven, Leuven, Belgium. [3]VIB Center for Medical Biotechnology, VIB, 9000 Ghent, Belgium. [4]Department of Biomolecular Medicine, Ghent University, 9000 Ghent, Belgium. [5]Laboratory for RNA Cancer Biology, Department of Oncology, LKI, KU Leuven, Leuven, Belgium. [6]Sylvester Comprehensive Cancer Center, Department of Human Genetics, University of Miami Miller School of Medicine, 1501 NW 10th Avenue, Miami, FL 33136, USA. [7]Bioinformatics Core Laboratory, Children's Hospital Research Institute of Manitoba (CHRIM), Winnipeg, Manitoba, Canada. [8]Department of Pediatrics and Child Health, University of Manitoba, Winnipeg, Manitoba, Canada. [9]Department of Molecular Biology and Genetics, Aarhus University, Aarhus, Denmark. [10]VIB Proteomics Core, 9000 Ghent, Belgium.
*These authors contributed equally to this work.
†Corresponding author. Email: jeanchristophe.marine@kuleuven.vib.be

is not polyadenylated, is produced as a read-through transcript when the 3′ end processing of *NEAT1_1* is inefficient (*14*). The mechanisms underlying *NEAT1* isoform switching remain poorly understood. The ubiquitous heterogeneous nuclear ribonucleoprotein K (HNRNPK) has been implicated in this process, by competing with CPSF6 for the binding of NUDT21 (CFIm25) and impairing *NEAT1_1* polyadenylation (*14*). Moreover, TAR DNA binding protein 43 (TDP-43) enhances *NEAT1_1* polyadenylation in pluripotent cells (*15*). Whereas the function of *NEAT1_1* still needs to be established (*16*, *17*), *NEAT1_2* is an essential architectural component of paraspeckles (PS) (*18*), which are highly ordered and phase-separated nuclear stress bodies (*19*). Thus, PS assembly critically depends on the poorly understood *NEAT1* isoform switch. Expression of *NEAT1_2*, and thereby PS assembly, can only be detected under specific physiological conditions (i.e., lactating mammary glands) and in response to various forms of stresses, including oncogenic stress (*12*, *20*–*22*). Accordingly, PS appear in over 65% of human epithelial cancers (*12*), where they predict poor prognosis (*23*) and are either completely absent, or only sporadically detectable, in normal tissues (*12*, *24*). In a classical two-stage chemically induced skin cancer mouse model, PS are induced in skin epidermal cells exposed to oncogenic stress, while genetic ablation of *NEAT1* markedly impairs tumor initiation and progression into aggressive and invasive lesions (*12*). However, mouse skin that lacks only the short *Neat1_1* isoform does not exhibit these protective properties (*17*). Critically, specific down-regulation of *NEAT1_2* using antisense oligonucleotides sensitized a series of epithelial cancer cell lines to various clinically relevant anticancer therapeutics (*12*). Hence, these studies identified *NEAT1_2*, and by extension PS, as promising cell-specific therapeutic targets for the chemosensitization of a wide range of epithelial cancers. We therefore reasoned that a better understanding of pathways and factors/enzymes involved in the molecular mechanisms underlying *NEAT1* isoform switching and, thereby, PS biogenesis may lead to the identification of targets that are amenable to conventional therapeutics.

RESULTS

Identification of Integrator as a previously unknown *NEAT1* RNA interactor

To identify proteins that modulate *NEAT1_2* expression and consequently PS biogenesis, we adapted the RNA antisense purification (RAP) protocol we previously used to identify interactors of the melanoma-specific lncRNA *SAMMSON* (*25*). Complexes directly bound to the endogenous *NEAT1* transcript were purified from freshly isolated nuclei of MCF-7 (human breast adenocarcinoma) cells exposed to ultraviolet (UV) cross-linking, using tiling DNA-based biotinylated oligonucleotides, targeting the 5′ portion of *NEAT1* (N1_5′). In parallel, control probes (Ctrl) were designed against the melanoma-specific *LINC00698* transcript, which is not expressed in MCF-7 cells (Fig. 1A). The quality of the nuclear isolation was verified by reverse transcription quantitative polymerase chain reaction (RT-qPCR), assessing the cytoplasmic RNA encoding the 40*S* ribosomal protein *S14*, as well as *NEAT1* and *MALAT1*, both of which are exclusively nuclear transcripts (Fig. 1B). The efficiency and specificity of the N1_5′ pulldown was confirmed by RT-qPCR. Whereas a robust signal was detected for total *NEAT1* (*NEAT1*) and *NEAT1_2* transcripts in the isolated RAP extracts, neither the housekeeping *TBP* and *HPRT1* mRNAs nor the lncRNA *MALAT1*, used as negative controls, were detectable (Fig. 1C).

RAP experiments were performed in biological triplicates, and purified proteins were analyzed by label-free mass spectrometry (MS). Principal component analysis (PCA) of the replicates confirmed the clustering of the samples into two groups: the control (Ctrl) and the RAP pulldown performed with *NEAT1*-specific probes (N1_5′) (fig. S1A). We identified 34 proteins, which were significantly enriched by the N1_5′ probes [*t* test, $P < 0.05$ and fold change (FC) > 1.6], as high-confidence *NEAT1* interactors (table S1). Two of these were known PS proteins (Fig. 1D) (*43*). Gene ontology analysis using the Search Tool for Recurring Instances of Neighboring Genes (https://string-db.org) indicated that the remaining *NEAT1* interactors are mainly involved in key aspects of RNA biogenesis and processing (table S2). Among these were multiple subunits of the Integrator and mRNA 3′ end processing complexes, including INTS1, INTS3, INTS6, CSTF1, CSTF2, CSTF2T, CSTF3, CPSF1, WDR33, SYMPK, and FIP1L1 (Fig. 1E, top). One-third (11 of 34) of all high-confidence interactors belonged to these two multicomponent protein complexes. Three additional previously unknown *NEAT1* interactors were also identified, namely, the F-box protein FBXO11 and its binding partner CUL1, as well as the transcription factor TCF7L2 (Fig. 1E, bottom).

These findings were next validated by RAP Western blotting experiments. Using the N1_5′ probes, the interactions between *NEAT1* and FBXO11, TCF7L2, CPSF2 (component of the mRNA 3′ end processing machinery), INTS3, INTS11 (catalytic subunit of Integrator complex), INIP (auxiliary component of the complex), H3 (used as negative control), and the PS proteins PSF, PSPC1, NONO, and TDP-43 (used as a positive control) were confirmed (Fig. 1F).

Integrator limits PS biogenesis by promoting *NEAT1* isoform switching

The identification of several components of the Integrator complex as high-confidence *NEAT1* RNA interactors raises the possibility that the complex contributes to the regulation of *NEAT1* isoform switching. To test this hypothesis, we first checked for interaction between *NEAT1* transcript and INTS11, the catalytic subunit of Integrator. To this end, we performed enhanced cross-linking and immunoprecipitation (eCLIP), a well-established and comprehensive procedure for the identification of RNA binding protein targets (*26*), in HeLa cells using two distinct INTS11-specific antibodies (Fig. 2A). Quantification of the eCLIP signal (size-matched coverage) relative to the control immunoglobulin G (IgG) showed an enrichment of INTS11 binding to the *NEAT1_1* transcript (Fig. 2B). *RNU11* and other noncoding RNAs are shown as positive and negative controls, respectively. Note that the binding between integrator and *NEAT1_1* could also be validated in MCF-7 cells by RNA immunoprecipitation (RIP) qPCR (Fig. 2C and fig. S1B). In agreement with previous findings (*7*, *27*), eCLIP detects two major peaks at the 5′ end of the transcript that may be implicated in premature transcriptional termination. Another peak is detected immediately upstream of the 3′ end of *NEAT1_1* (Fig. 2D).

To determine whether Integrator contributes to the 3′ end maturation of *NEAT1_1*, we performed RNA sequencing (RNA-seq) in HeLa cells expressing a doxycycline-inducible short hairpin RNA (shRNA) construct targeting INTS11 or green fluorescent protein (GFP) as a control (shCtrl) (Fig. 2E). We observed an accumulation of the long *NEAT1_2* isoform in the INTS11 knockdown (KD) cells (Figs. 2E and 3A and fig. S1C). In agreement with the total RNA-seq data, small RNA-seq (smRNA-seq) analysis, which captures cleavage

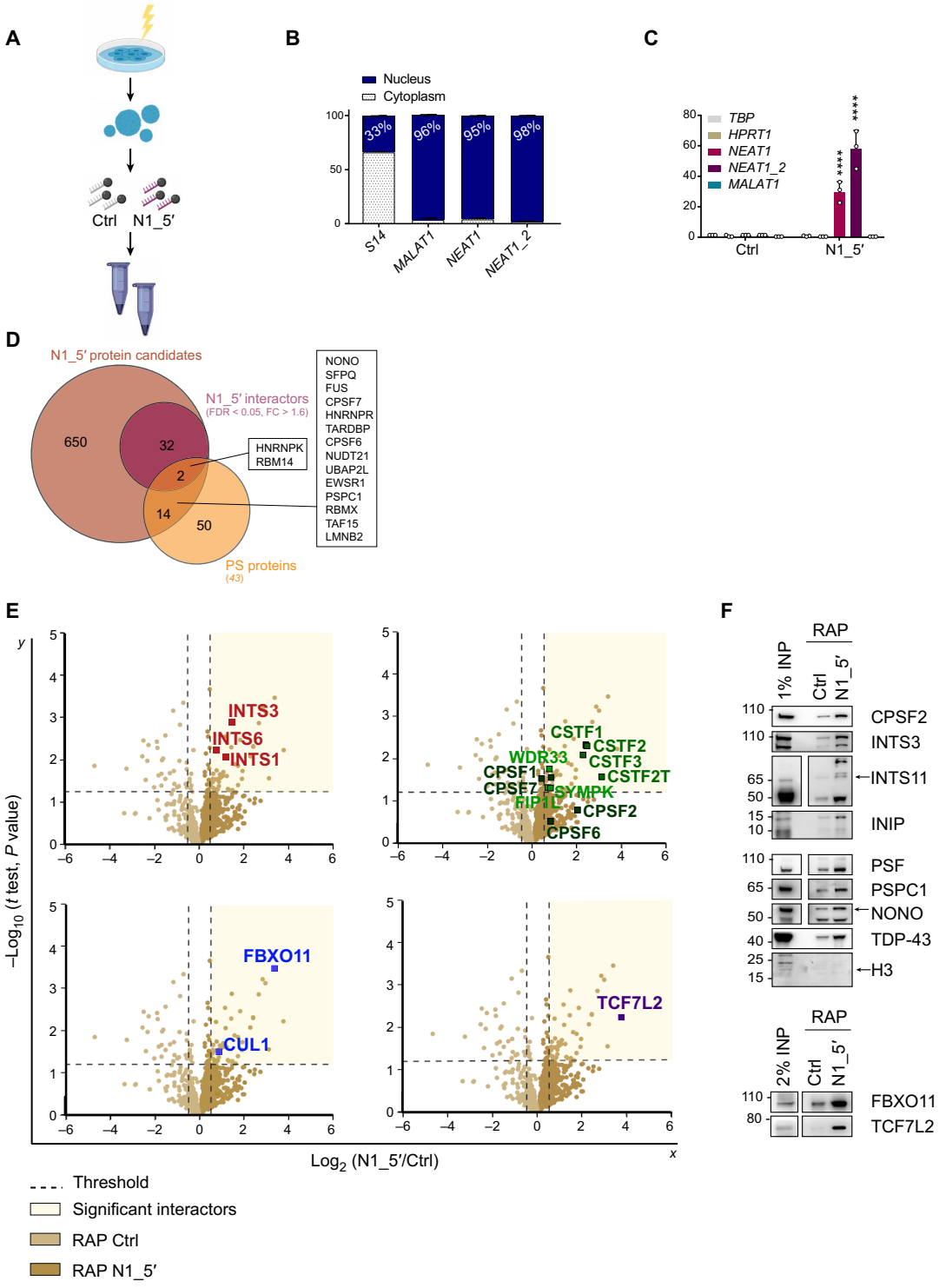

Fig. 1. *NEAT1* interactome is enriched in 3′ end processing factors. (**A**) RAP–mass spectrometry (MS) strategy to study *NEAT1* interactome. Probes targeting either the melanoma-specific *LINC00698* (Ctrl) or the 5′ end of *NEAT1* short and long isoforms (N1_5′) were used. (**B**) Relative abundance by RT-qPCR of the *S14*, *NEAT1*, and *MALAT1* in MCF-7 cytoplasmic and nuclear fractions. (**C**) Both form-specific (*NEAT1*) and long form-specific (*NEAT1_2*) RT-qPCR to evaluate the efficiency of the RAP. Three abundant coding and noncoding transcripts (*MALAT1*, *TBP*, and *HPRT1*) are shown as negative controls. Error bars represent means ± SD. *P* values were calculated by two-way analysis of variance (ANOVA), using three biological replicates. ****$P < 0.0001$. (**D**) Venn diagram showing overlap between *NEAT1* previously identified partners (orange) and candidates identified in this study (brown). In purple, the interactors considered for further analysis. (**E**) Volcano plots indicating the interactors significantly enriched (*t* test, $P < 0.05$ and FC > 1.6) by N1_5′ probes. Highlighted candidates are color-coded by protein complexes. The *x* axis indicates the ratio N1_5′/Ctrl in $\log_2$ scale. The *y* axis is the $-\log_{10}$ (*t* test, *P* value). (**F**) RAP Western blot validation of novel interactors. PS proteins PSF, PSPC1, NONO, and TDP-43 were used as positive controls, and H3 was used as negative control. Input (INP) was either 1 or 2% of the total nuclear lysate.

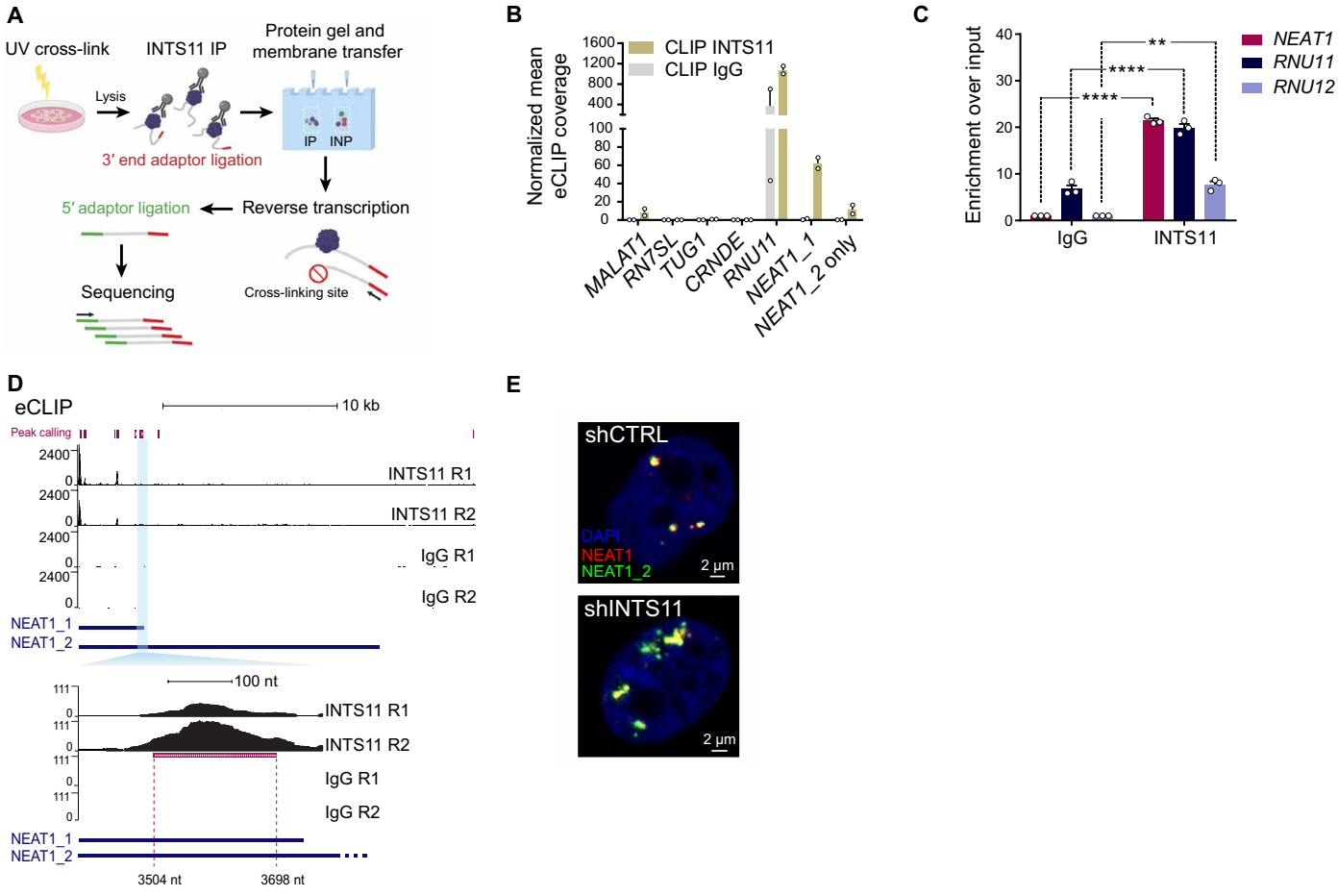

Fig. 2. Integrator is a novel *NEAT1* interactor. (**A**) eCLIP protocol (HeLa). IP, immunoprecipitation; INP, input. (**B**) Quantification of INTS11 (and size-matched IgG) eCLIP signal at indicated loci relative to expression levels (RNA-seq–based). (**C**) RIP RT-qPCR for INTS11 in MCF-7. Error bars represent means ± SD. *P* values were calculated by two-way ANOVA, using three biological replicates. $^{**}P < 0.01$ and $^{****}P < 0.0001$. (**D**) INTS11 eCLIP at *NEAT1* locus. Binding site on *NEAT1* is in pink. In the insert, the 3′ end peak and corresponding position on *NEAT1* transcript. (**E**) FISH for *NEAT1* (both transcripts, red) and *NEAT1_2* (long transcript, green) in HeLa upon INTS11 KD (shINTS11) or control (shCtrl).

products of INTS11 catalysis (RNA species smaller than 75 nt), indicated a decrease in the small RNA cleavage product of *NEAT1_1* (Fig. 3A, bottom). Likewise, 3′ mRNA-seq, which detects selectively polyadenylated transcripts, confirmed a decrease in the *NEAT1_1* 3′ ends upon Integrator KD (fig. S1D).

This phenotype was not an off-target effect, as it was rescued by concomitant expression of a wild-type (WT) form of INTS11 but not a catalytic dead mutant (E203Q) (Fig. 3B and fig. S1E), indicating that the *NEAT1* isoform switching is dependent on INTS11 enzymatic activity. Similar results were obtained for the snRNA *RNU11*, a well-known Integrator target (fig. S1, F and G).

The phenotype was not cell type specific, as it could be recapitulated in the breast cancer cell line MCF-7 in which INTS11 was silenced by small interfering RNA (siRNA) (Fig. 3C and fig. S1H). As expected, the levels of the 3′ end extended product of two well-established Integrator targets, *RNU11* and *RNU12*, as indicated by the percentage of long transcript relative to the gene body (Fig. 3C). Together, these data further supported a direct contribution of Integrator in the regulation of *NEAT1* isoform switching and indicated that the catalytic activity of Integrator is required for the correct processing of the *NEAT1* transcript.

The observed increase in *NEAT1_2* levels upon silencing of INTS11 raised the possibility that Integrator activity limits the formation of PS. Consistent with this possibility, an increase in the number and size of *NEAT1_2* foci was observed by RNA fluorescence in situ hybridization (FISH) in MCF-7 cells depleted for INTS11 (Fig. 3D). RNA FISH coupled to immunofluorescence (IF) revealed that the foci colocalized with the PS-specific protein PSPC1, thus demonstrating an increase in PS assembly in INTS11-depleted cells (Fig. 3E and fig. S1I). This observation indicated that Integrator restrains the formation of PS nuclear bodies by promoting *NEAT1_1* expression, at the detriment of *NEAT1_2*, under steady-state conditions.

Stress does not disrupt *NEAT1*-Integrator interaction and promotes accumulation of Integrator at PS

Various forms of stress stimulate PS formation (*28*, *29*). It was recently shown that exposure of cells to hydroxyurea (HU), which inhibits deoxyribonucleotide synthesis and thus causes DNA replication stress, induces the formation of large PS (*17*). Accordingly, increased PS formation was detected by RNA FISH and RNA FISH combined to IF for PSPC1 in HU-exposed MCF-7 cells (Fig. 4A). The increase in PS formation was the result of the transcriptional up-regulation

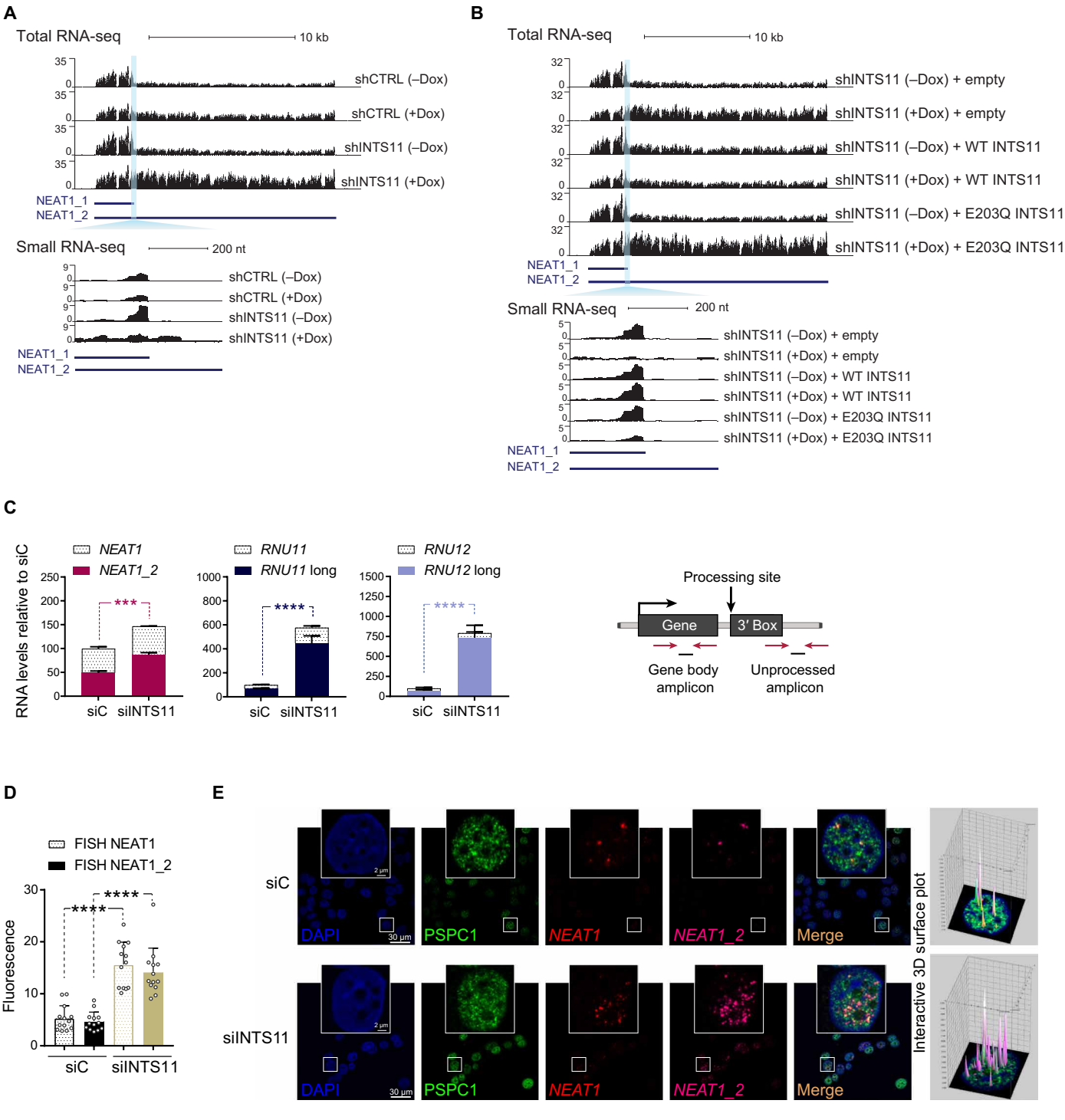

Fig. 3. Integrator restrains PS biogenesis. (**A**) RNA-seq (total and small) of HeLa shINTS11 or shCtrl at *NEAT1* locus. The blue shade indicates *NEAT1* termination site. Dox, doxycycline. (**B**) RNA-seq (total and small) of HeLa expressing INTS11 catalytic dead mutant (E203Q), empty vector (empty), or WT INTS11 at *NEAT1* locus. (**C**) RT-qPCR for *NEAT1* and *NEAT1_2* expression (relative to siC) in MCF-7 upon siINTS11. *RNU11* and *RNU12* and their read-through are positive controls. Error bars represent means ± SD. *P* values were calculated by two-tailed *t* test, using three biological replicates. ***$P < 0.001$ and ****$P < 0.0001$. (**D**) FISH quantification for *NEAT1* and *NEAT1_2*. Error bars represent means ± SD. *P* values were calculated by paired two-tailed *t* test. ****$P < 0.0001$. (**E**) Immunofluorescence (IF) for the PS protein PSPC1 (green) and RNA FISH for *NEAT1* in MCF-7 cells siINTS11 (total *NEAT1*, red; *NEAT1_2*, magenta; nuclei, blue). Right: 3D quantification peaks represent signal intensities at colocalization site.

of the *NEAT1* locus (particularly *NEAT1_2*) as demonstrated by the fact that treatment with actinomycin D (RNA Pol II inhibitor) abolished HU-induced *NEAT1* up-regulation (Fig. 4B). We reasoned that stress-induced *NEAT1_2* expression and PS formation may be caused, at least in part, by a decrease in the recruitment of Integrator to the *NEAT1* transcript. To test this hypothesis, we performed RAP-MS experiments on freshly isolated nuclei of MCF-7 cells exposed to HU (fig. S2, A and B). As expected, the recovery of *NEAT1_2* RNA in this assay was higher in HU-treated than in dimethyl sulfoxide (DMSO)–treated cells (fig. S2A). Unexpectedly, our RAP-MS data revealed that most of the previously enriched candidates from unstimulated cells were also recovered in stimulated cells (Fig. 4C, fig. S2C, and table S3). Integrator subunits and mRNA 3′ end processing factors were also efficiently pulled down by the RAP N1_5′ probes under these experimental conditions. These interactions were further validated by RAP Western blotting analysis (Fig. 4D). IF for the PS protein PSPC1, in intact cells exposed to HU, was performed to confirm the colocalization of CPSF1 and CPSF2 (as well as FBXO11 and TCF7L2) with PSPC1 (fig. S2D) and thus their recruitment to PS. Moreover, stochastic optical reconstruction microscopy (STORM) showed a significant colocalization of the Integrator subunit INTS1 with the PS protein NONO in cells exposed to HU (Fig. 4E). Together, these data demonstrated that DNA damage–induced stress is not sufficient to disrupt the interaction between *NEAT1* and Integrator. These experiments also showed that Integrator accumulates to PS nuclear bodies in DNA-damaged cells.

DNA damage–induced PS formation is, at least partly, a consequence of activation of the p53 transcription factor (fig. S2B), which, in turn, enhances *NEAT1* promoter activity (*12*, *22*). Accordingly, exposure to the MDM2 antagonist Nutlin-3a, which causes stabilization of p53 without inducing cellular stress responses, was sufficient to increase *NEAT1_2* transcription and enlarged PS (Fig. 5, A and B). Consistently, similar to HU, exposure to Nutlin-3a did not disrupt *NEAT1*-Integrator association as demonstrated by RAP-MS (Fig. 5C, fig. S2E, and table S4). PCA of all the RAP-MS experiments under control and stressed conditions confirmed the consistency of these results (fig. S2F). RAP Western blotting (Fig. 5D), confocal microscopy, and super-resolution microscopy (Fig. 5E and fig. S2G) further confirmed that activation of p53 by Nutlin-3a is not sufficient to disrupt the interaction between *NEAT1* and Integrator and that Integrator accumulates to PS nuclear bodies under these experimental conditions.

Note that exposure to HU or Nutlin-3a did not significantly alter the expression levels of various Integrator subunits (Fig. 6A). The processing of several well-known Integrator targets, including histones (*7*), was severely compromised upon induction of DNA damage or p53 activation (Fig. 6, B and C). Further evidence of impaired Integrator activity was also obtained in MCF-7 treated with Nutlin-3a, with HU, or transfected with siINT11, using a reporter construct that directs expression of GFP upon read-through of RNU7 (fig. S3, A to C) (*30*). These data indicated that Integrator activity is compromised in cells exposed to stress, possibly as a consequence of its recruitment to PS. Together with the observation that stress does not disrupt the interaction between *NEAT1* and Integrator, these data favor a model in which up-regulation of *NEAT1_2* levels and PS formation in stressed cells (exhibiting elevated transcriptional rates of *NEAT1*) occurs because the amount of functional Integrator available to process *NEAT1* transcripts becomes rate liming. Consistent with this model, overexpression of exogenous INTS11 (fig. S3D) abolished stress-induced up-regulation of *NEAT1_2* (Fig. 6, D and E), decreased PS assembly (Fig. 6F), and phenocopied the decrease in p53 activation and increase in levels of DNA damage observed following *NEAT1_2* KD (Fig. 6, G and H).

Low levels of Integrator components correlate with poorer survival and response to chemotherapy

We previously established a genetic link between PS formation and tumorigenesis and demonstrated that PS can be detected in about 65% of the human carcinomas analyzed, including skin squamous cell carcinoma and ovarian carcinomas. We also showed that expression of *NEAT1_2*, but not *NEAT1_1*, reliably predicts the response of ovarian cancer to platinum-based chemotherapy (*12*). Given that Integrator modulates *NEAT1_2* expression and PS biogenesis, we therefore assessed whether correlations between (altered) expression of Integrator subunits and overall patient survival (OS) may exist. Analysis of patients that underwent chemotherapy in the ovarian cancer cohort (GSE30161) analyzed in our previous study (*12*) confirmed that lower levels of INTS10 and INTS11 significantly correlated with worse OS (Fig. 7A). In this cohort, the differential expression levels of INTS11 and INTS10 exhibit an inverse relationship with that of *NEAT1_2* as shown in Fig. 7B. Analysis was then expanded to publicly available TCGA datasets corresponding to 11 epithelial cancer cohorts. In addition to gene expression levels, the results were adjusted for the effect of other risk factors (covariates) such as age, race, stage, and gender by performing a multivariate analysis using the Cox proportional hazards model (table S5). Moreover, within these studies, only participants who underwent treatment with various chemotherapeutic agents were retained. Consistently, patients with lower levels of INTS6, INTS7, INTS8, INTS10, INTS11, and INTS12 exhibited poorer OS (Fig. 7C and fig. S4). Notably, the most striking effect was observed with the catalytic subunit of Integrator, INTS11. We subsequently performed a similar analysis using multiple Affymetrix gene expression cancer datasets, including two colorectal cancer cohorts (GSE33113 and GSE39582) and two breast cancer cohorts, one of which is split into two Gene Expression Omnibus (GEO) submissions (GSE9195, GSE6532.1, and GSE6532.2). Again, a significant correlation between low expression levels of Integrator subunits and a poorer OS was observed (table S5). Together, these data support a model in which decreased levels/activity of the Integrator complex may affect chemotherapeutic response via modulation of the biogenesis of *NEAT1_2* and PS.

DISCUSSION

Using an unbiased proteomics screen, we have identified known and previously unknown *NEAT1* RNA binding partners, such as the transcription factor TCF7L2 and a member of the F-box protein family, FBXO11, which were both subsequently validated as bona fide *NEAT1* interactors and novel PS proteins. This study therefore provides a new list of factors that may modulate *NEAT1* and PS biology. A large proportion of the identified *NEAT1* interactors belongs to two functionally related protein complexes, namely, the core 3′ end processing and Integrator complexes. The Integrator complex contains two essential Integrator subunits, INTS11 and INTS9, which are homologous of CPSF73 (alias CPSF3) and CPSF100 (alias CPSF1), respectively. Integrator interacts with the CTD of RNA Pol II and processes newly transcribed RNA molecules, mainly nonpolyadenylated transcripts and UsnRNAs. Integrator has also been recently implicated in the

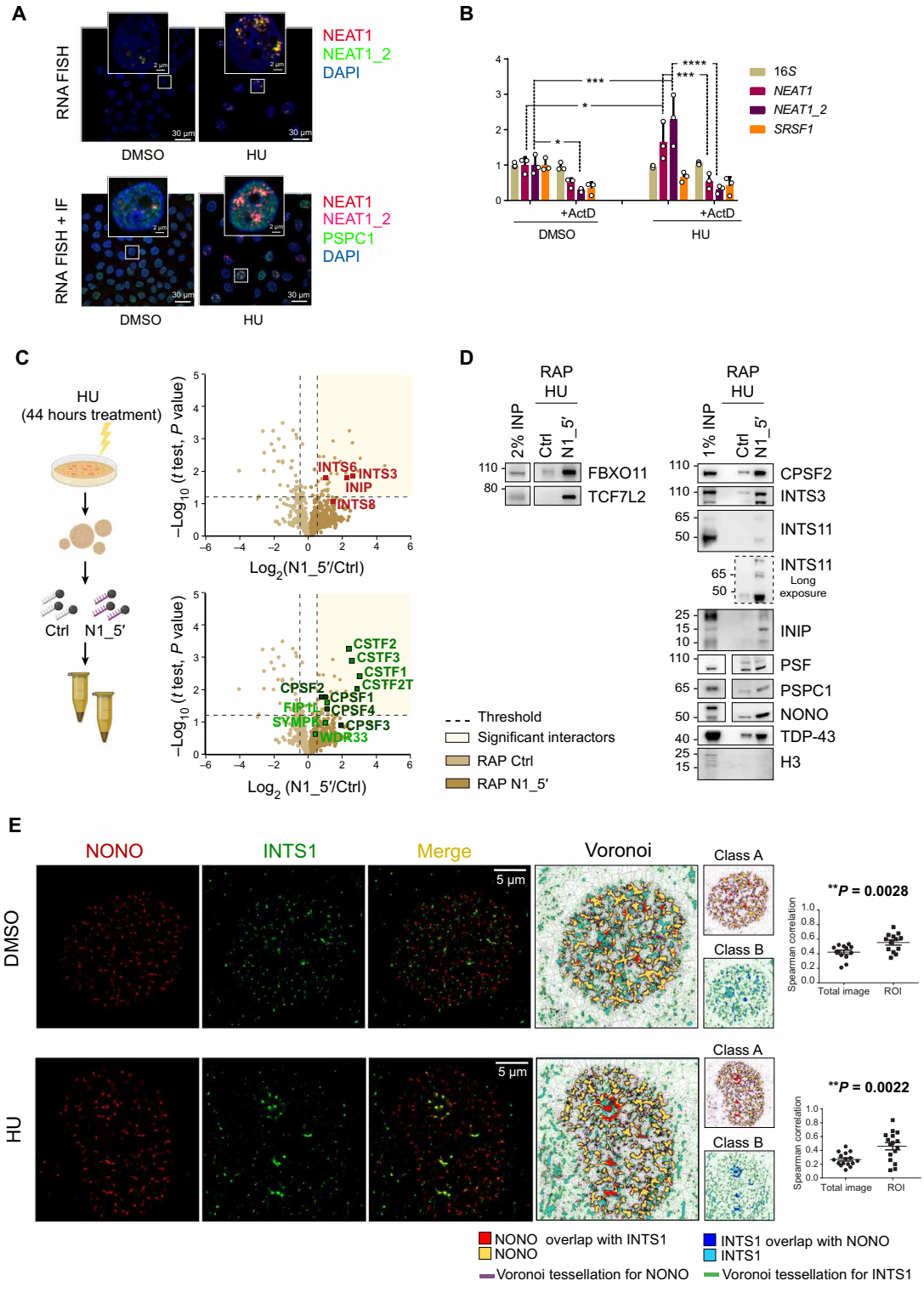

Fig. 4. *NEAT1* interactors do not dissociate upon DNA damage–induced PS formation. (**A**) Top: RNA FISH for total *NEAT1* (red) *NEAT1_2* (green) nucleus (blue). Bottom: RNA FISH combined to IF for total *NEAT1* (red), *NEAT1_2* (magenta), PSPC1 (green), and nucleus (blue) in MCF-7 exposed to HU. (**B**) RT-qPCR of total *NEAT1* transcript (*NEAT1*) and long form (*NEAT1_2*), in MCF-7 exposed to HU after a pulse of actinomycin D (ActD). *SRSF1* is a positive control, and 16*S* is a negative control. Error bars represent means ± SD. *P* values were calculated by two-way ANOVA, using three biological replicates. **P* < 0.05, ****P* < 0.001, and *****P* < 0.0001. (**C**) RAP-MS experimental approach in MCF-7 treated with HU. In the volcano plots, significantly enriched protein candidates (*t* test, *P* < 0.05 and FC > 1.6) are color-coded by protein complexes. (**D**) RAP Western blot in MCF-7 for PS proteins (positive controls) and the newly identified *NEAT1* interactors. H3 is a negative control. Input (INP) loaded is either 1 or 2% of the total nuclear lysate. (**E**) STORM images of the PS protein NONO (red) and INTS1 (green) in MCF-7 treated with HU. Voronoi diagram and quantification of NONO (class A) and INTS1 (class B) colocalization are shown.

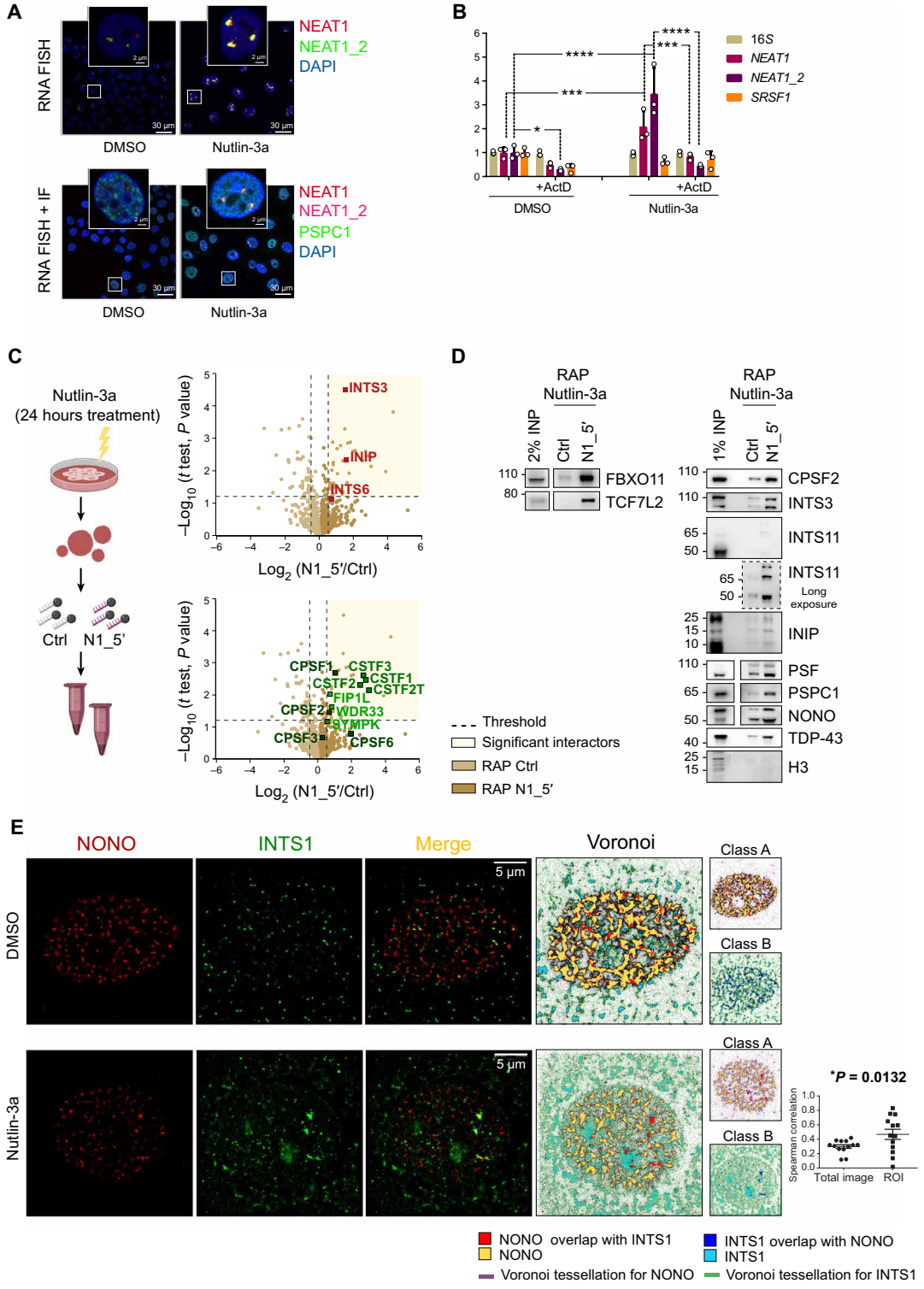

Fig. 5. *NEAT1* interactome does not change upon oncogenic stress–induced PS formation. (**A**) Top: RNA FISH for total *NEAT1* (red) *NEAT1_2* (green) nucleus (blue). Bottom: RNA-FISH combined to IF for total *NEAT1* (red), *NEAT1_2* (magenta), PSPC1 (green), and nucleus (blue) in MCF-7 exposed to Nutlin-3a. (**B**) RT-qPCR of total *NEAT1* transcript (*NEAT1*) and long form (*NEAT1_2*), in MCF-7 exposed to Nutlin-3a after a pulse of actinomycin D. *SRSF1* is a positive control, and 16*S* is a negative control. Error bars represent means ± SD. *P* values were calculated by two-way ANOVA, using three biological replicates. (**C**) Scheme of the RAP-MS approach in MCF-7 treated with Nutlin-3a. In the volcano plots, significantly enriched protein candidates (*t* test, $P < 0.05$ and FC > 1.6) are color-coded by protein complexes. (**D**) RAP Western blot performed in MCF-7 exposed to Nutlin-3a. Western blot was probed with the PS proteins (positive controls) and the newly identified *NEAT1* interactors. H3 is used as negative control. Input (INP) is either 1 or 2% of the total nuclear lysate. (**E**) STORM images of the PS protein NONO (red) and INTS1 (green) in MCF-7 cells treated with Nutlin-3a. Voronoi diagram and quantification of NONO (class A) and INTS1 (class B) colocalization are shown.

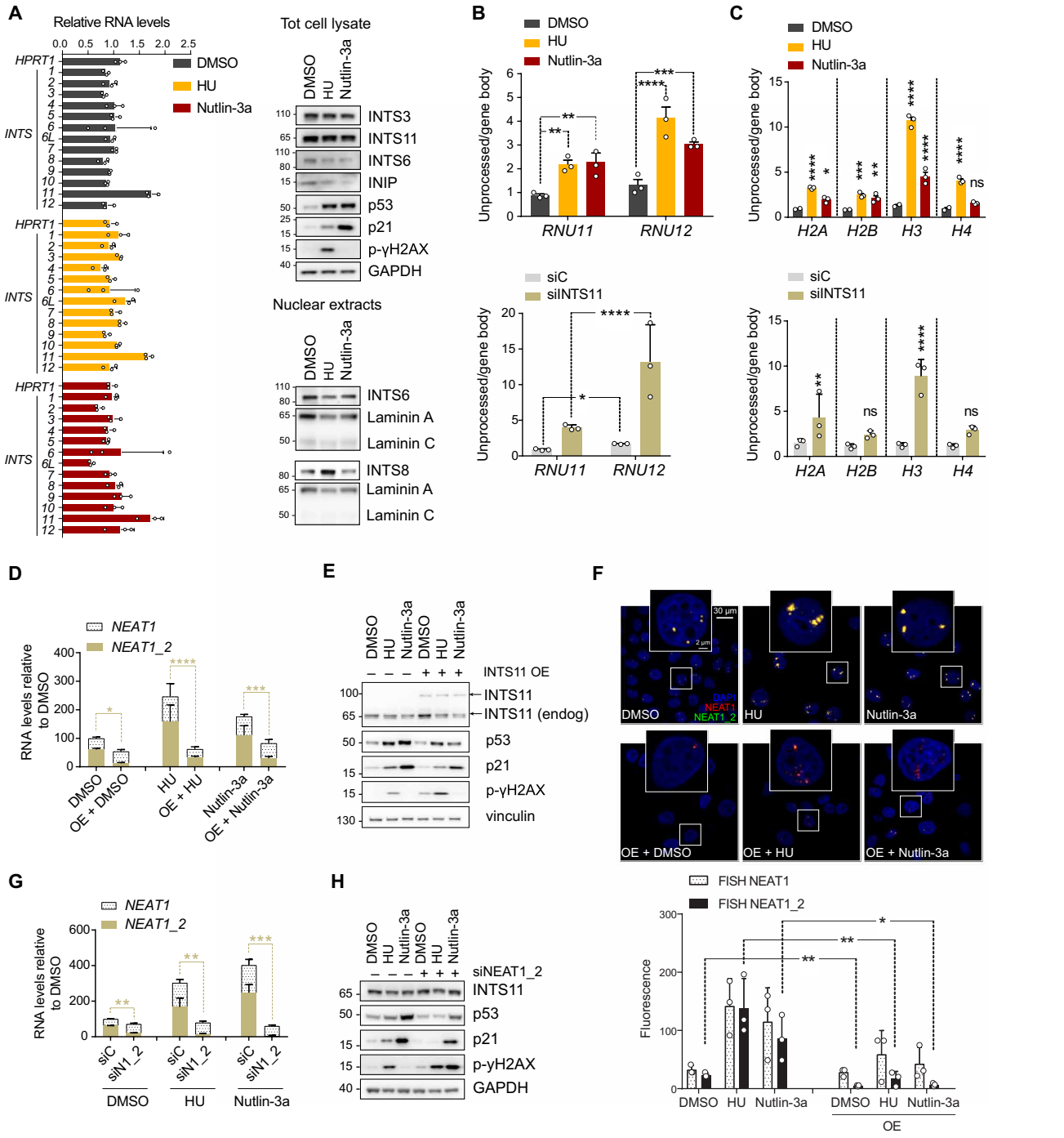

Fig. 6. Stress compromises Integrator activity and determines read-through at target genes. (**A**) RT-qPCR (left) and Western blot (right) for integrator subunits (INTS) in MCF-7 exposed to Nutlin-3a and HU. RT-qPCR: error bars represent means ± SD; *P* values calculated by two-way ANOVA using three biological replicates were not significant (ns). Western blot: GAPDH and Laminin A + C are loading controls; p53, p21, and phospho-γH2AX (p-γH2AX) show efficacy of the treatment. (**B** and **C**) RT-qPCR of read-through at known INTS11 targets in MCF-7 exposed to HU, Nutlin-3a, or siINTS11 (C). Ratios between read-through (unprocessed) and gene body are calculated on normalized expression levels. Error bars represent means ± SD. *P* values were calculated by two-way ANOVA, using three biological replicates. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, and *****P* < 0.0001. (**D**) RT-qPCR for total *NEAT1* and *NEAT1_2* in MCF-7 exposed to stress and overexpressing WT INTS11 (or empty vector). Normalized expression levels are relative to the control (DMSO). Error bars represent means ± SD. *P* values were calculated by two-way ANOVA, using three biological replicates. **P* < 0.05, ****P* < 0.001, and *****P* < 0.0001. (**E**) Western blot confirming INTS11 overexpression (OE). p53, p21, and phospho-γH2AX detect p53 activity and DNA damage; vinculin is a loading control. (**F**) FISH staining and relative quantification (biological triplicate) of *NEAT1* (both transcripts, red) and *NEAT1_2* (long transcript, green) in MCF-7 with WT INTS11 OE and exposure to stress. Error bars represent means ± SD. *P* values were calculated by two-way ANOVA, using three biological replicates. **P* < 0.05 and ***P* < 0.01. (**G**) RT-qPCR for *NEAT1* and *NEAT1_2* in MCF-7 transfected with siN1_2 and exposed to stress. Normalized expression levels are relative to the control (siC). Error bars represent means ± SD. *P* values were calculated by two-way ANOVA, using three biological replicates. ***P* < 0.01 and ****P* < 0.001. (**H**) Western blot analysis for p53, p21, and phospho-γH2AX, confirming the efficacy of the treatments; GAPDH is a loading control.

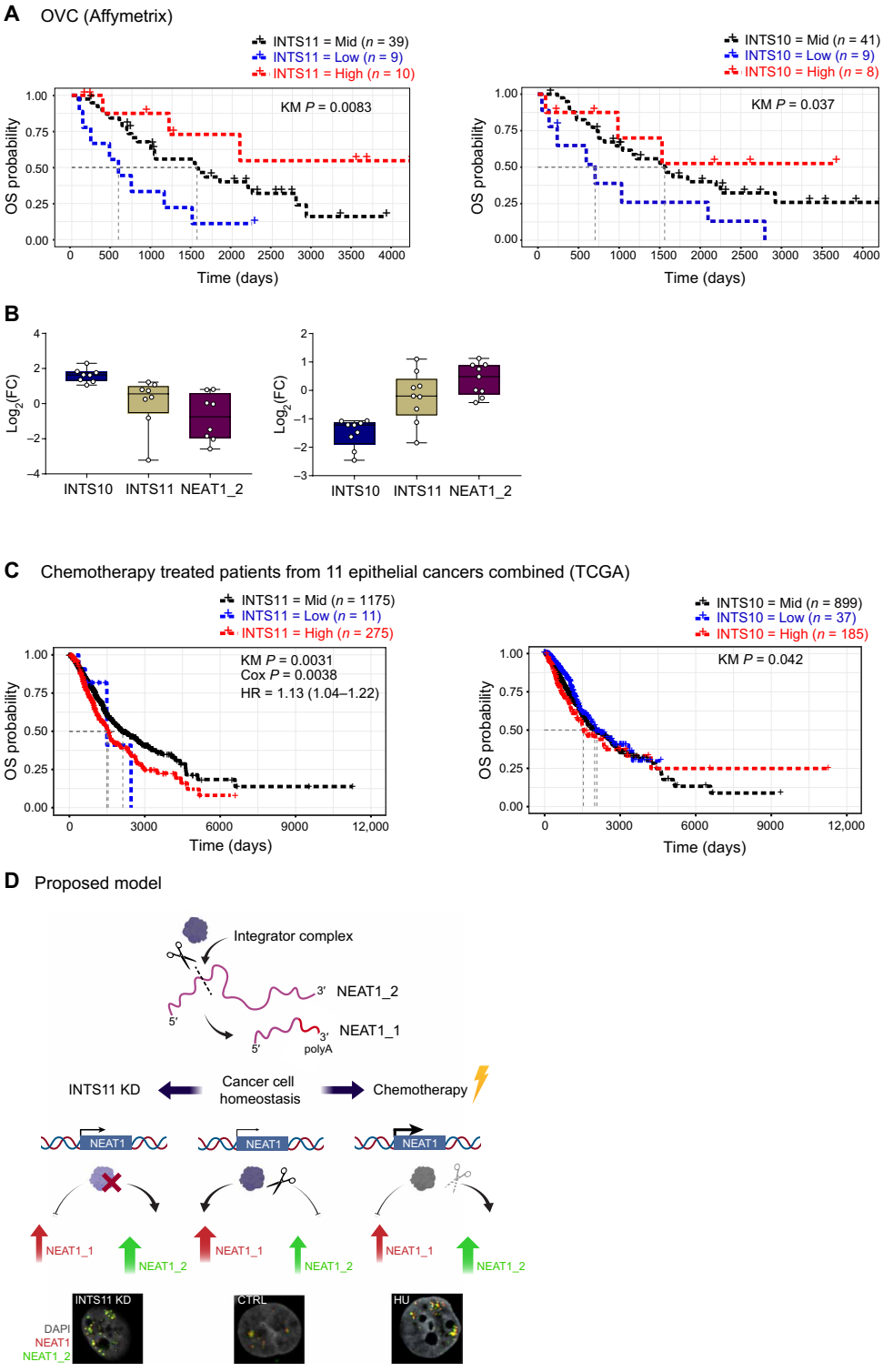

Fig. 7. Low levels of Integrator components correlate with poorer survival and response to chemotherapy. (**A**) Kaplan-Meier (KM) survival plots of patients from the ovarian cancer Affymetrix cohort (OVC; GSE30161). Patients treated with chemotherapy were selected and stratified on the basis of the expression levels of INTS11 and INTS10. The KM log rank test *P* values indicate the difference in overall survival between the patients with significantly low (and high) expression levels of Integrator subunit. (**B**) Bar plots represent the log FCs of INTS10 and INTS11 and the corresponding FCs for *NEAT1_2* among respective samples. FCs were calculated using *z* scores to represent the classic null hypothesis of no overall FC in the mean of all samples. (**C**) KM survival curve of patients treated with chemotherapy from 11 epithelial cancer cohorts combined (TCGA database). Patients were stratified in three groups based on the expression levels of Integrator subunits (blue, low; black, intermediate; red, high). Graphs refer to the subunits INTS11 and INTS10, whose lower expression levels significantly correlate with worse prognosis of chemo-treated patients. We also indicate the *P* value of the Cox analysis and the hazard ratio (HR) calculated for INTS11. (**D**) Proposed model.

modulation of gene expression via regulation of protein-coding gene transcription initiation and premature termination (*7*, *27*), in RNA Pol II pause release (*31*, *32*), and in the biogenesis of enhancer RNAs (*6*). Here, we provide functional evidence for an unexpected role of Integrator in the regulation of the isoform switching of the lncRNA *NEAT1*. Our data are compatible with a model in which Integrator is recruited to the *NEAT1* transcript and participates in the cleavage and subsequent processing of the polyadenylated *NEAT1_1* isoform (Fig. 7D). Although eCLIP data indicate that Integrator is also recruited to the 5′ end of the *NEAT1* transcript, our data highlight a role for Integrator in the processing of the 3′ end of *NEAT1_1* to restrain the expression of the long isoform and, thereby, PS formation. Moreover, our previous observation that *NEAT1_1* is constantly made and degraded by the exosome (*17*) raises the possibility that processing of this 3′ end site by Integrator is a critical step in this degradation process. Previous data have already implicated the 3′ end CFIm complex in the processing of *NEAT1_1*. Whether Integrator and the core 3′ end processing machinery cooperate to process polyadenylated transcripts such as *NEAT1_1* or work independently on different pools of transcripts remains to be addressed.

The interaction between *NEAT1* and Integrator is not disrupted in cells exposed to stress (i.e., HU-induced replication stress) or in cells in which we artificially increased the transcription rate of *NEAT1* (i.e., upon Nutlin-3a exposure). These data therefore favor a model in which bypassing *NEAT1* cleavage may occur because the pool of Integrator available is not sufficient to process the high amounts of *NEAT1* transcripts being produced in cells exposed to stress (or in which the transcriptional rate of *NEAT1* is artificially elevated). The ratio between the rate of *NEAT1* transcription and overall expression levels of the Integrator complex may therefore determine whether *NEAT1_2* remains expressed and whether PS are being assembled (Fig. 7D). This model is further supported by the fact that *NEAT1* is an unusually abundant lncRNA (the "A" in *NEAT1* refers to "abundant"), being expressed at levels that rival highly expressed housekeeping genes, such as *GAPDH*.

The observation that the *NEAT1*-Integrator association is not disrupted in stressed cells may have important functional implications. We showed that several components of Integrator colocalize with PS in stressed cells. PS assembly is thought to phase separate its content from the nucleoplasm (*19*, *28*), and thus Integrator recruitment to PS may affect its recruitment and activity at other loci. Paralleling this possibility, a comparable cross-regulation between TDP-43 and *NEAT1*/PS was recently shown to promote pluripotency-differentiation transition (*15*). In addition to repressing the formation of PS by enhancing the maturation of *NEAT1_1*, TDP-43 also regulates alternative 3′ end processing of transcripts encoding pluripotency factors, such as SOX2. PS sequester TDP-43, just like Integrator, and thereby reduce its binding to polyadenylated RNAs to promote exit from pluripotency (*15*). In a similar way, sequestration of Integrator to PS may contribute to an overall decrease in the processing of small and/or enhancer RNAs in stressed cells and thereby cause an overall down-regulation of gene expression and/or rewiring toward a "stress" transcriptome that helps cells cope with (chemotherapy-induced) stress. In support of this hypothesis, our data show that the processing of two well-known Integrator targets *RNU11* and *RNU12* is compromised in cells exposed to HU and Nutlin-3a (Fig. 6B). Similarly, the 3′ end processing of the replication-dependent histones, previously shown to be affected by silencing of INTS3 (*7*), was also compromised under these experimental conditions (Fig. 6C). On the other hand, overexpression of INTS11 under condition of stress abolished the increase in *NEAT1_2* and PS and increased DNA damage, thus phenocopying the effects observed upon *NEAT1_2* KD (Fig. 6, G and H). This model is compatible with the switch from cell cycle arrest/dormancy to apoptotic cell death we observed in cancer cells exposed to chemotherapy following *NEAT1_2* silencing (and PS disruption) and suggests a role for PS as key modulators of 3′ end RNA processing.

Last, our data also establish an important mechanistic link between Integrator and PS biology. Given the recently recognized role of PS as modulators of cancer development and sensitivity to cancer therapy, our work therefore highlights the importance of studying Integrator in a cancer biology context. In keeping, we provide evidence that decreased expression of various components of the Integrator complex, as well as, in particular, its catalytic subunit INTS11, correlates with poorer clinical outcome for patients exposed to chemotherapy. These observations may ultimately bear important therapeutic implications. Agents that may increase either the half-life or the recruitment of Integrator to the *NEAT1* locus, or stimulate INTS11 catalytic activity, would be expected to impair PS formation and thereby increase chemosensitivity.

MATERIALS AND METHODS

Cell culture and cloning

All cell lines were acquired from the American Type Culture Collection Cell Biology collection and kept in culture at 37°C and 5% CO_2 in medium supplemented with 1% penicillin and streptomycin (Invitrogen) and 10% fetal bovine serum (Invitrogen). All cell lines tested negative for mycoplasma contamination. MCF-7 breast cancer cell line was grown in RPMI 1640 GlutaMAX (Gibco, Invitrogen) supplemented with insulin (10 μg/ml; Sigma-Aldrich, I9278).

INTS11- and GFP-inducible KD clones (HeLa cells) were established as previously described in (*6*). HeLa rescue cells were established by cloning the same shINTS11 sequence into Tet-pLKO-neo vector (Addgene), and single clones were selected with G418 (500 μg/ml). shRNA-resistant N-terminal Flag-tagged WT or E203Q mutant INTS11 complementary DNA (cDNA) (*5*) was cloned into Cumate-pLenti-Cloning-2A-GFP vector (ABM Inc.) and transfected into a shINTS11-Tet-pLKO-neo single clone. Stable cell lines were maintained in puromycin (2 μg/ml) and G418 (200 μg/ml) containing Dulbecco's modified Eagle's medium. KDs were induced by adding of doxycycline (1 μg/ml) into the culture medium daily for 3 days. WT INTS11 cDNA was cloned into a VP16 plasmid (Addgene) to transiently overexpress (OE) INTS11 in MCF-7 cells to perform rescue experiments.

Cell transfections

For transient KD experiments, MCF-7 cells were seeded in six-well plates (200,000 cells per well) and transfected with Lipofectamine RNAiMax (Thermo Fisher Scientific) according to the manufacturer's instructions, using 30 nM siNEAT1 or siNEAT1_2 siPOOLs (siTOOLS Biotech) or 35 nM ON TARGETplus siCPSF3L (siINTS11, Dharmacon). Transient transfections with the plasmid of interest were performed in six-well plates (120,000 cells per well) using Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's instructions. We transfected either 10 μg of DNA for the pVP16-INTS11 overexpression construct or 60 μg of DNA for the U7-GFP reporter construct (*30*). Cell medium was refreshed after 8 hours from transfection, and treatments started 24 hours after transfection.

Cell treatments

MCF-7 cells were treated with 5 μM Nutlin-3a (Selleckchem) for 24 hours or with 1 mM HU (Sigma-Aldrich) for 44 hours. For actinomycin D experiments in Figs. 4B and 5B, MCF-7 cells were seeded in six-well plates (180,000 cells per well) and exposed to 1-hour pulse of 3 μM actinomycin D (Sigma-Aldrich) 24 hours after seeding. After two washes in phosphate-buffered saline (PBS), cells were treated with either DMSO (vehicle), 5 μM Nutlin-3a, or 1 mM HU for 24 hours. RNA was extracted with TRIzol lysis reagent (QIAGEN) according to the manufacturer's instructions, and deoxyribonuclease (DNAse) treated to measure the transcript levels of the 16*S* ribosomal RNA, the lncRNA *NEAT1* (*NEAT1*) and its long form specifically (*NEAT1_2*), and *SRSF1* (used here as positive control) by RT-qPCR.

In the rescue experiments, cells were first transfected with either INTS11-overexpressing construct or with siNEAT1_2 siPooLs and then continuously treated with DMSO, Nutlin-3a (5 μM), or HU (1 mM) for 24 hours (rescue with siNEAT1_2) or 72 hours (rescue with INTS11 OE plasmid). For the RNA read-through experiments of Fig. 6 (B and C), MCF-7 cells were either transfected with 35 nM ON TARGETplus against CPSF3L (siINTS11) or exposed to stress for 108 hours (5 μM Nutlin-3a or 1 mM HU).

Cell fractionation

Nuclear and cytoplasmic extracts were prepared from 15-cm plates using the Nuclei EZ prep kit (Sigma-Aldrich) according to the manufacturer's instructions. The quality of the nuclear isolation was verified by RT-qPCR, assessing the cytoplasmic RNA encoding the 40*S* ribosomal protein *S14* and the exclusively nuclear noncoding RNAs *NEAT1* and *MALAT1*.

RAP and quantitative label-free MS

Briefly, for antisense purification of the protein interactors of *NEAT1*, 100 μg of Streptavidin Sepharose High Performance beads (GE Healthcare) were coupled overnight at 4°C to 800 pmol of biotinylated RAP probes against the 5′ portion of the *NEAT1* transcript (N1_5′; Biosearch Technologies) or RAP probes designed against the melanoma-specific *LINC00698* (Ctrl; Biosearch Technologies). MCF-7 breast cancer cells (1.5×10^7 cells per treatment) were washed twice in PBS and UV cross-linked dry at 400 mJ/cm^2 with a CL-1000 Crosslinker (254-nm lamp). After performing cell fractionation as indicated above, nuclei were lysed in pulldown buffer [20 mM tris-HCl (pH 8.0), 200 mM NaCl, 2.5 mM $MgCl_2$, and 0.05% Triton X-100 in diethyl pyrocarbonate (DEPC) water] supplemented with a cocktail of protease inhibitors [Halt Protease and Phosphatase Inhibitor Single-Use Cocktail (100×), Thermo Fisher Scientific], 1 mM dithiothreitol, and SUPERase• In RNase (ribonuclease) Inhibitor (60 U/ml; Life Technologies). Lysates were incubated with the beads coupled to the RAP probes at 4°C for 3 hours. Beads were rinsed three times with pulldown buffer and twice with DEPC-treated water. For MS analysis, proteins were rinsed in trypsin digestion buffer [20 mM tris-HCl (pH 8.0) and 2 mM $CaCl_2$] and eluted by on-beads digestion with 1 μg of trypsin (Promega) overnight at 37°C. Peptides were purified with OMIX Tips (C18 resin) and dried to be stored till MS analysis (see "Liquid chromatography–MS/MS analysis" section). For Western blot, proteins were directly eluted in 30 μl of Laemmli buffer supplemented with tris(2-carboxyethyl)phosphine (TCEP), boiled for 15 min at 95°C, and stored at −80°C. For RNA elution, samples were first decross-linked at 56°C in decross-linking buffer [100 mM tris-HCl (pH 7.5), 50 mM NaCl, 10 mM EDTA, and 0.5% SDS] with proteinase K (Roche) to a final working concentration of 2 mg/ml for 30 to 40 min and then extracted in TRIzol-chloroform and precipitated overnight at −80°C in 1/10th (v/v) NaCl and 100% EtOH. The purified RNA was treated with DNAse, measured with a nanodrop, and stored at −80°C.

Liquid chromatography–MS/MS analysis

The cleaned peptide mixtures were dried completely and resuspended in 20 μl of loading solvent [0.1% trifluoroacetic acid in water/acetonitrile, 2/98 (v/v)]. Two microliters of the peptide mixtures were analyzed by liquid chromatography (LC)–MS/MS on an UltiMate 3000 RSLCnano LC (Thermo Fisher Scientific, Bremen, Germany) in-line connected to a Q Exactive mass spectrometer (Thermo Fisher Scientific). Peptides were separated with a linear gradient at 300 nl/min from 98% solvent A (0.1% formic acid in water) to 55% solvent B [0.1% formic acid in water/acetonitrile, 20/80 (v/v)] in 120 min before ultimately reaching 99% solvent B. The mass spectrometer was operated in data-dependent, positive ionization mode, automatically switching between MS and MS/MS acquisition for the 10 most abundant peaks in a given MS spectrum.

Proteomics data analysis

Data analysis was performed with MaxQuant (version 1.5.4.1) using the Andromeda search engine with default search settings including a false discovery rate (FDR) set at 1% on both the peptide and protein level. Spectra were searched against human proteins in the UniProt/Swiss-Prot database (database release version of August 2016 containing 20,210 human protein sequences; www.uniprot.org). The mass tolerance for precursor and fragment ions was set to 20 and 4.5 parts per million, respectively, during the main search. Enzyme specificity was set to C terminus to arginine and lysine, also allowing cleavage at arginine/lysine-proline bonds with a maximum of two missed cleavages. Variable modifications were set to oxidation of methionine (to sulfoxides) and acetylation of protein N termini. A minimum of one peptide was required for protein identification. We allowed for matching between runs using a 1-min match time window and a 20-min alignment time window. Proteins were quantified by the MaxLFQ algorithm integrated in the MaxQuant software. A minimum ratio count of two unique or razor peptides was required for quantification.

Further data analysis was performed with the Perseus software (version 1.5.5.3) loading the protein groups file from MaxQuant. First, proteins only identified by site, reverse database hits, and potential contaminants were removed. The label-free quantification (LFQ) intensities were $\log_2$ transformed, the replicate samples were grouped, and protein groups with less than three valid values in at least one group were removed. Missing values were then imputed with values from the lower part of the normal distribution representing the detection limit, leading to a list of 1063 reliably quantified proteins. Moreover, we filtered out proteins identified by less than three peptides ($n = 995$). Then, a *t* test was performed (FDR = 0.05) to compare the RAP N1_5′ with the RAP Ctrl samples and generate the volcano plots depicted in Figs. 1E, 4C, and 5C and figs. S2 (C and E). Of the 995 quantified protein candidates, 698 candidates were enriched by N1_5′ RAP probes. Significantly enriched proteins ($P < 0.05$) with a N1_5′/Ctrl FC of >1.6 (arbitrary cutoff) were considered as highly confident *NEAT1* interaction partners (tables S1, S3, and S4).

Principal component analysis

Ellipses represent 95% confidence intervals, around each cluster's centroid, calculated using Hotelling's T^2 statistics. The axes are the respective first and second principal components with the percent variance captured by each principal component in the parentheses. The figure was generated in R using the "factoextra" package. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the Proteomics Identification Data (PRIDE) partner repository with the dataset identifier PXD015158.

RNA immunoprecipitation

RIP was performed on freshly isolated nuclei from MCF-7 cells (2.5 10^7 cells per sample) after UV cross-linking with UV_{254nm} (0.4 J/cm^2). Nuclei were lysed with polysome buffer [20 mM tris-HCl (pH 8.0), 200 mM NaCl, 2.5 mM $MgCl_2$, and 1% Triton X-100 in DEPC water] supplemented with a cocktail of protease inhibitors [Halt Protease and Phosphatase Inhibitor Single-Use Cocktail (100×), Thermo Fisher Scientific], 1 mM dithiothreitol, and SUPERase• In RNase Inhibitor (60 U/ml; Life Technologies) and precleared with protein A beads for 1 hour at 4°C. RIP was performed overnight at 4°C on a rotating wheel using 5 μg of the specific antibody INTS11 (Sigma-Aldrich, A107128) or normal rabbit IgG (Millipore, 12-370) used as control. On the following day, 50 μl of protein A Dynabeads (Invitrogen) were coupled to the antibody for 3 hours at 4°C. The beads were rinsed five times with polysome buffer and split in two to either elute proteins or RNA (see RAP protocol for elution steps).

Reverse transcription quantitative polymerase chain reaction

Total RNA was extracted with TRIzol lysis reagent (QIAGEN) according to the manufacturer's instructions. To improve the extractability of *NEAT1_2*, we routinely perform the following additional step: TRIzol samples are heated at 56°C for 5 minutes or syringed 20 times with 22-gauge insulin syringes (BD). RNA is DNAse-treated, and reverse-transcribed using the High-Capacity cDNA RT Kit (Thermo Fisher Scientific). RNA expression levels were measured by qPCR on a LightCycler 480 (Roche). Data were analyzed in qbase + 3.0 (Biogazelle) using *HPRT1*, *TBP*, and *GAPDH* as reference genes. For the sequences of the RT-qPCR primers (see table S6). Primers for *HIST* transcripts (Fig. 6C) were taken from (*7*).

RAP (and RIP) analysis

The RAP (and RIP) efficiency was estimated by RT-qPCR starting from 0.2 μg of RNA per sample. The enrichment of the gene of interest for the RAP (RIP) experiment (*NEAT1* and *NEAT1_2* primers) was calculated applying the $\Delta(\Delta C_t)$ method. Briefly, the C_t value of the RAP (RIP) elution was subtracted from the C_t value of the input for every gene, thus obtaining the ΔC_t for each gene in the RAP (RIP) sample. From the RAP (RIP), ΔC_t was subtracted by the ΔC_t of the RAP control (Ctrl probes, targeting the melanoma-specific *LINC00698*) or of the normal IgG (RIP), for every gene, thus obtaining the $\Delta(\Delta C_t)$. The equation "fold enrichment = $2 - \Delta(\Delta C_t)$" was used to calculate the FC for each gene and was plotted as such.

Immunoblotting

Cells were scraped on ice in radioimmunoprecipitation assay buffer (RIPA) containing protease and phosphatase inhibitor cocktails (Thermo Fisher Scientific). The cell lysates were syringe five times with a 22-gauge needle, vortexed, incubated on ice for 10 min, and then centrifuged at 21,000*g* for 15 min at 4°C. Thirty or 20 μg of total protein lysate were loaded on NuPAGE Novex 4 to 12% Bis-Tris Protein Gels (Invitrogen) and probed with primary antibodies at 4°C overnight (see the "Antibodies" section below).

eCLIP assay

eCLIP was performed in HeLa cells in duplicates as previously described in (*26*). Briefly, 2×10^7 cells were cross-linked by UV-C irradiation (254 nm, 400 mJ/cm^2) and lysed on ice, followed by sonication. Antibodies (INTS11: Abcam ab75276 or Sigma Prestige HPA029025) were incubated with Dynabeads M-280 Sheep Anti-Rabbit IgG (Invitrogen, 11204D) for 1 hour. After limiting RNase I (Ambion) digest in presence of DNase, the lysate was subjected to immunoprecipitation at 4°C for 16 hours. In the following, 2% of the lysate was removed for size-matched input control. Immunoprecipitation efficiency and specificity were verified by immunoblot using 20% of the immunoprecipitation material. Coimmunoprecipitated RNA was dephosphorylated, followed by 3′ RNA adapter ligation using T4 RNA Ligase (New England Biolabs). Input and IgG controls and INTS11-RNA complexes were run on a NuPAGE 4 to 12% Bis-Tris Gel, transferred to nitrocellulose, and cut from the membrane between 65 and 145 kDa. Protein-bound RNA was released from the membrane by urea/proteinase K digest, followed by acid phenol/chloroform/isoamyl alcohol RNA extraction and purification using RNA Clean & Concentrator (Zymo Research). After RT (AffinityScript Reverse Transcriptase, Agilent), RNA was treated with exonuclease (ExoSAP-IT, Affymetrix) and removed by combined NaOH/HCl treatment. A 3′ linker was ligated to the cDNA, and the resulting library was PCR-amplified using Q5 Polymerase (New England Biolabs), purified, and size-selected for sequencing. Single-end (SE100) sequencing was performed to an average of 40 million reads per sample using Illumina HiSeq 3000 sequencer. Data were processed according to (*26*), including removal of repetitive sequences before mapping against the human genome version hg19. eCLIP sequencing coverage of noncoding RNAs (*MALAT1*, *RN7SL*, *TUG1*, *CRNDE*, *RNU11*, *NEAT1_1*, and *NEAT1_2* only) was quantified using bigWigAverageOverBed (*33*). Mean eCLIP signal per transcript was normalized to the expression levels of the lncRNA based on total RNA-seq (with the *NEAT1_2* transcript arbitrarily set to 1). *RNU11* was used as positive control, and highly expressed *RN7SL* (3000-fold higher expressed than *NEAT1_2*), moderately expressed *MALAT1* (35-fold higher expressed than *NEAT1_2*), and lowly expressed *TUG1* (0.5-fold) and *CRNDE* (0.1-fold) are also shown. Significant INTS11 binding compared to input was determined using the CLIPper tool with a threshold of $\log_2$ of >3.7 and $P < 10^{-26}$ (*34*).

RNA sequencing

A total of ~3×10^7 cells were used for total RNA extraction using TRIzol reagent (Thermo Fisher Scientific, #15596026) according to the manufacturer's instructions. Genomic DNA was removed by Turbo DNAse treatment (Invitrogen, #AM1907). Total RNA-seq libraries were produced using TruSeq Stranded Total RNA library prep kit (Illumina, #20020596) with 500 ng of DNAse-treated input RNA. Genome-wide experiments were performed as two independent biological replicates. To avoid a batch effect in library preparation and sequencing flow cell, these replicates were processed together. Raw fastq RNA-seq data were processed with Trimmomatic v0.32 (*35*) and aligned to the human genome (hg19 version) using STAR aligner v2.5.3a (*36*) with default parameters. For visualization on the

University of California, Santa Cruz (UCSC) Genome Browser, all tracks were CPM (counts per million) normalized against the total number of usable reads in that data set using deepTools2 (*37*).

Small RNA analysis

A total of ~3 × 10^8 cells were used for nuclear fractionation, and RNA was extracted using TRIzol reagent (Thermo Fisher Scientific, #15596026) according to the manufacturer's instructions. Genomic DNA was removed by Turbo DNAse treatment (Invitrogen, #AM1907). Small RNA libraries were prepared using the SMARTer smRNA-seq Kit (Takara, #635030) with 750 ng of nuclear-enriched total RNA, and the experiments were performed as two independent biological replicates. Raw fastq reads were then adapter-trimmed (AAAAAAA) as recommend by SMARTer smRNA-seq kit (Takara, #635030) protocol using Cutadapt (v1.14), and reads less than 17 base pairs (bp) were discarded. First, we aligned the reads against human elements in Repbase (v23.08) with STAR (v2.5.3a) (*36*), repeat-mapping reads were removed, all others were then mapped against the full human genome (hg19 version), and we keep all unique aligned reads. For visualization on the UCSC Genome Browser, all tracks were CPM normalized against the total number of usable reads in that data set using deepTools2 (*37*).

3′ end RNA- seq (3′ quant-seq) and data analysis

Total RNA was extracted and treated with TURBO DNase for 60 min at 37°C. We used QuantSeq 3′ mRNA-Seq Library Prep Kit REV (Lexogen) to prepare 3′ end libraries. 3′ Quant-seq was performed on NEXTSeq 500 machine with single-end 75-bp sequencing. For the data analysis, we followed the Lexogen protocol. Briefly, raw fastq data were processed with BBMap (https://sourceforge.net/projects/bbmap/) to remove the adapter contamination, poly(A) read-through and low-quality tails, and aligned to the human genome (hg19 version) using STAR aligner v2.5.3a (*36*) with the following parameters (– outFilterType BySJout – outFilterMultimapNmax 20 –alignSJoverhangMin 8 – alignSJDBoverhangMin 1 – outFilterMismatchNmax 999 –outFilterMismatchNoverLmax 0.1 – alignIntronMin 20 – alignIntronMax 1000000 –alignMatesGapMax 1000000 – outSAMattributes NH HI NM MD). For visualization on the UCSC Genome Browser, all tracks were CPM normalized against the total number of usable reads in that dataset using deepTools2 (*37*).

RNA FISH

RNA FISH was performed using Stellaris FISH probes (Biosearch Technologies) for human *NEAT1*: SMF-2036-1 for NEAT1_5 and VSMF-2251-5 for NEAT1_m. FISH was performed according to the manufacturer's protocol. Briefly, cells were grown on slides (round cover glasses; VWR), fixed in 4% paraformaldehyde (PFA), and permeabilized in 70% EtOH over night at 4°C. Cells can be stained within the following 2 weeks maximum. Cells were washed twice in PBS and incubated for 5 min in FISH washing buffer [2× standard saline citrate (SSC) and 10% formamide]. Hybridization of FISH probes was carried out overnight at 37°C in 2× SSC, 10% formamide, and 10% dextran, in a dark humid chamber. After three washes with FISH washing buffer, slides were mounted in ProLong Gold Antifade containing 4′,6-diamidino-2-phenylindole (DAPI) (Thermo Fisher Scientific) and images acquired on a confocal microscope Nikon C2. Imaging panels were prepared using Imaris 7.2.3 and ImageJ [plugins such as Interactive three-dimensional (3D) surface plot and JACoP (Just Another Colocalization Plugin) were used, respectively, to produce the 3D plots to show colocalization and to quantify fluoresce and signal colocalization]. In Fig. 2I, 25 MCF-7 cells randomly selected cells from three biological replicates were used for the quantification.

Immunofluorescence

Cells were grown on slides, fixed in 4% formaldehyde, and permeabilized in 70% EtOH overnight at 4°C. Cells were washed twice in PBS and blocked for 1 hour in 3% bovine serum albumin (BSA) (Sigma-Aldrich), 10% goat serum (DAKO), and 0.2% Triton X-100 (Sigma-Aldrich). Slides were incubated with primary antibodies at room temperature for 1 hour, washed three times in PBS, and incubated with secondary antibodies, either anti-rabbit or anti-mouse Alexa Fluor 488 or Alexa Fluor 555 (Life Technologies) at room temperature for 45 min. After three washes in PBS, slides were mounted in ProLong Gold Antifade Mountant with DAPI (Thermo Fisher Scientific). Images were acquired on a confocal microscope Nikon C2. Imaging panels were prepared using Imaris 7.2.3 and ImageJ.

IF combined to RNA FISH

Cells were grown on slides and fixed in 4% paraformaldehyde (PFA), permeabilized in 70% EtOH overnight at 4°C, and stained within the following 2 weeks maximum. The protocol for RNA FISH was performed first by incubation overnight at 37°C with FISH probes (NEAT1_5 Quasar 560 and NEAT1_m Quasar 670). The following day cells were incubated 30 min in FISH wash buffer at 37°C, washed twice in PBS, and fixed again at room temperature for 15 min in 2% PFA. After two washes in PBS, cells were blocked for 1 hour in IF buffer: 3% BSA (Sigma-Aldrich), 10% goat serum (DAKO), 0.2% Triton X-100 (Sigma-Aldrich), and SUPERase• In RNase Inhibitor (60 U/ml; Life Technologies), used also for further washes and antibody incubations. Slides were incubated with the primary antibody for 1 hour at room temperature in the dark. After three washes, cells were incubated for 45 min in secondary antibody anti-mouse or anti-rabbit Alexa Fluor 488 (Life Technologies) and washed again in PBS prior of mounting the slides with ProLong Gold Antifade Mountant with DAPI (Thermo Fisher Scientific). Images were acquired on a confocal microscope Nikon C2. Imaging panels were prepared using Imaris software 7.2.3 and ImageJ (the plugin Interactive 3D surface plot was used to produce the 3D plots to show colocalization; JACoP was used to quantify fluoresce and signal colocalization).

Antibodies

Western blotting experiments and/or IF and IF combined to FISH were performed using the following primary antibodies: PSPC1 (Sigma-Aldrich, SAB4200503), PSF (Sigma-Aldrich, P2860), FBXO11 (Novus Biologicals, NB100-59826), TCF7L2 (Cell Signaling Technology, 2565), CPSF1 (Santa Cruz Biotechnology, sc-166281), CPSF2 (Santa Cruz Biotechnology, sc-165983), GAPDH (Abcam, ab9485), NONO (Bethyl Laboratories, A300-587A), TDP-43 (ProteinTech, 12892-1 ap), CPSF3L or INTS11 (Sigma-Aldrich, A107128), INTS1 (Millipore), INIP [C9orf80 (E-12), Santa Cruz Biotechnology, SC-137357], INTS3 (Bethyl Laboratories, A300-427A), INTS6 (Bethyl Laboratories, A301-658A), INTS8 (Bethyl Laboratories, A300-269A), p53-DOI (Santa Cruz Biotechnology, sc-126), p21 (Santa Cruz Biotechnology, sc-6246), phospho-γH2AX (Cell Signaling Technology, 2577), Laminin A + C (Abcam, ab108922), vinculin (Sigma-Aldrich, V9131), GFP (Clontech, 632375), and H3 (Abcam, ab1791).

Stochastic optical reconstruction microscopy

Cells were seeded in ibidi μ-Slide 4 wells and, after the indicated treatments, fixed in 4% formaldehyde for 10 min at room temperature. Cells were permeabilized in 0.5% Triton X-100 for 15 min and incubated overnight at −20°C with the primary antibody (recombinant anti-nmt55/p54nrb antibody, Abcam, ab133574) at a dilution of 1:1500 and with INTS1 antibody (MilliporeSigma, MABS1984) at a dilution of 1:100. Last, samples were incubated for 30 min at room temperature with the secondary antibodies diluted 1:500, JF646 anti-rabbit (Novus Biologicals, NB7156JF646) and Alexa Fluor 568 anti-mouse (Thermo Fisher Scientific, A-11004).

Imaging experiments were carried out with a Nikon eclipse Ti2 microscope equipped with Nikon Instruments (N-STORM). For two color dSTORM imaging, Janelia 646, and Alexa Fluor 568 secondary antibodies were used with MEA STORM imaging buffer and were imaged continuously with 5000 frames collected per filter range at a frequency of 20 ms. Images were acquired using a 100×, 1.49–numerical aperture objective, and imaged onto a Hamamatsu C11440 ORCA-flash 4.0 camera. Storm localization analysis was carried out with ImageJ, thunderstorm plugin (1.3-2014-11-08). Molecule list files were then exported from ImageJ to be further analyzed using Coloc-Tesseler. Cluster analysis, specifically Voronoi function, was carried out after manually selecting regions of interest (ROIs). For quantification, the whole image was compared versus the selected ROI (1.5 μm by 1.5 μm) area of PS (high NONO signal) for all the treatment groups (at least 13 samples per group, we used 16 for HU). A two-tailed *t* test was performed using GraphPad Prism 5. More details on the analysis method have been published previously (*38*, *39*).

In silico survival analysis

Gene expression data and the corresponding clinical information from 11 cancers were downloaded from TCGA repository using the *GDCquery* function of the TCGAbiolinks R package (*40*). These data were breast cancer, lung adenocarcinoma, colon adenocarcinoma, glioblastoma, low-grade glioma, liver hepatocellular carcinoma, kidney renal cell carcinoma, adrenocortical carcinoma, skin cutaneous melanoma, and ovarian cancer. All the expression and clinical data were then merged and analyzed together excluding genes whose expression level was zero (0) in 50% of the samples. The data were subsequently normalized using the voom normalization (*41*) and partitioned the data into "normal" and "tumor" samples based on the information provided with the clinical data.

Differential expression was calculated on the basis of *z* scores, calculated on the basis of the difference between the means of the normal samples and those from tumor samples correcting for sample heterogeneity using the SD across all genes. With this approach, we assigned differential expression to correspond to genes whose *z* scores were greater or less than ±1.96, respectively, in line with classical *z* statistical theory. We then used a univariate Kaplan-Meier survival analysis conducted to test the effect of the levels of expression of each of the genes of interest (*INTS6*, *INTS7*, *INTS8*, *INTS10*, *INTS11*, and *INTS12*) on the overall survival (OS) of the patients that underwent chemotherapy treatment. In addition, we performed a multivariate Cox proportional hazards analysis to model the effects of age, stage, race, and gender along with the expression levels of the genes on the overall survival.

DNA microarray data (Affymetrix dataset) were downloaded from GEO database repository. The data downloaded were GSE33113, GSE39582, GSE9195, GSE6532, and GSE30161. They were preprocessed using standard tools for microarray data normalization available through the Affy Package in R (*42*). Given the lack of normal samples in these data, differential expression was calculated, again, on the basis of the *z* score, assuming the global mean to represent the expression levels of normal samples. Kaplan-Meier survival analysis was conducted in a similar way to that described for TCGA data.

The bar plots shown in Fig. 7B represent the $\log_2$ FCs of INTS10 and INTS11, and the corresponding FCs for *NEAT1_2* among respective samples. INTS10 and INTS11 "high" represents the samples for which the log FCs of the Integrator subunit is greater than 1.96 (left panel), while INTS10 and INTS11 "low" represents log FC less than 1.96 (right panel). FCs were calculated using *z* scores (as described above) to represent the classic null hypothesis of no overall FC in the mean of all samples. A weighted average of all the five *NEAT1_2* Affymetrix probes was taken to represent the expression values for *NEAT1_2*.

Statistical analysis

The significance between means was determined by two-tailed paired Student's *t* test or with a two-way analysis of variance (ANOVA) test. All *P* values are represented as follows: ns (not significant), $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, and $****P < 0.0001$. All statistical analyses were performed with GraphPad Prism v7.0a.

Graphical output

Figure panels have been generated using Adobe Illustrator 22.1; scientific illustrations were created with the online web-based software BioRender (https://biorender.com/) and iStock (www.istockphoto.com/).

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/6/27/eaaz9072/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. B. Tian, J. Hu, H. Zhang, C. S. Lutz, A large-scale analysis of mRNA polyadenylation of human and mouse genes. *Nucleic Acids Res.* **33**, 201–212 (2005).
2. R. Elkon, A. P. Ugalde, R. Agami, Alternative cleavage and polyadenylation: Extent, regulation and function. *Nat. Rev. Genet.* **14**, 496–506 (2013).
3. A. J. Gruber, R. Schmidt, A. R. Gruber, G. Martin, S. Ghosh, M. Belmadani, W. Keller, M. Zavolan, A comprehensive analysis of 3′ end sequencing data sets reveals novel polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation. *Genome Res.* **26**, 1145–1159 (2016).
4. A. J. Gruber, M. Zavolan, Alternative cleavage and polyadenylation in health and disease. *Nat. Rev. Genet.* **20**, 599–614 (2019).
5. D. Baillat, M.-A. Hakimi, A. M. Näär, A. Shilatifard, N. Cooch, R. Shiekhattar, Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**, 265–276 (2005).
6. F. Lai, A. Gardini, A. Zhang, R. Shiekhattar, Integrator mediates the biogenesis of enhancer RNAs. *Nature* **525**, 399–403 (2015).
7. J. R. Skaar, A. L. Ferris, X. Wu, A. Saraf, K. K. Khanna, L. Florens, M. P. Washburn, S. H. Hughes, M. Pagano, The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Res.* **25**, 288–305 (2015).
8. R. Sandberg, J. R. Neilson, A. Sarma, P. A. Sharp, C. B. Burge, Proliferating cells express mRNAs with shortened 3′ untranslated regions and fewer microRNA target sites. *Science* **320**, 1643–1647 (2008).
9. Z. Xia, L. A. Donehower, T. A. Cooper, J. R. Neilson, D. A. Wheeler, E. J. Wagner, W. Li, Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3′-UTR landscape across seven tumour types. *Nat. Commun.* **5**, 5274 (2014).
10. C. P. Masamha, Z. Xia, J. Yang, T. R. Albrecht, M. Li, A.-B. Shyu, W. Li, E. J. Wagner, CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. *Nature* **510**, 412–416 (2014).
11. A. J. Gruber, R. Schmidt, S. Ghosh, G. Martin, A. R. Gruber, E. van Nimwegen, M. Zavolan, Discovery of physiological and cancer-related regulators of 3′ UTR processing with KAPAC. *Genome Biol.* **19**, 44 (2018).

12. C. Adriaens, L. Standaert, J. Barra, M. Latil, A. Verfaillie, P. Kalev, B. Boeckx, P. W. G. Wijnhoven, E. Radaelli, W. Vermi, E. Leucci, G. Lapouge, B. Beck, J. van den Oord, S. Nakagawa, T. Hirose, A. A. Sablina, D. Lambrechts, S. Aerts, C. Blanpain, J.-C. Marine, P53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. *Nat. Med.* **22**, 861–868 (2016).
13. Y. T. F. Sasaki, T. Ideue, M. Sano, T. Mituyama, T. Hirose, MENε/β noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 2525–2530 (2009).
14. T. Naganuma, S. Nakagawa, A. Tanigawa, Y. F. Sasaki, N. Goshima, T. Hirose, Alternative 3′-end processing of long noncoding RNA initiates construction of nuclear paraspeckles. *EMBO J.* **31**, 4020–4034 (2012).
15. M. Modic, M. Grosch, G. Rot, S. Schirge, T. Lepko, T. Yamazaki, F. C. Y. Lee, E. Rusha, D. Shaposhnikov, M. Palo, J. Merl-Pham, D. Cacchiarelli, B. Rogelj, S. M. Hauck, C. von Mering, A. Meissner, H. Lickert, T. Hirose, J. Ule, M. Drukker, Cross-regulation between TDP-43 and paraspeckles promotes pluripotency-differentiation transition. *Mol. Cell* **74**, 951–956.e13 (2019).
16. M. Isobe, H. Toya, M. Mito, T. Chiba, H. Asahara, T. Hirose, S. Nakagawa, Forced isoform switching of Neat1_1 to Neat1_2 leads to the loss of Neat1_1 and the hyperformation of paraspeckles but does not affect the development and growth of mice. *RNA* **26**, 251–264 (2020).
17. C. Adriaens, F. Rambow, G. Bervoets, T. Silla, M. Mito, T. Chiba, H. Asahara, T. Hirose, S. Nakagawa, T. H. Jensen, J.-C. Marine, The long noncoding RNA NEAT1_1 is seemingly dispensable for normal tissue homeostasis and cancer cell growth. *RNA* **25**, 1681–1695 (2019).
18. C. M. Clemson, J. N. Hutchinson, S. A. Sara, A. W. Ensminger, A. H. Fox, A. Chess, J. B. Lawrence, An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol. Cell* **33**, 717–726 (2009).
19. T. Yamazaki, S. Souquere, T. Chujo, S. Kobelke, Y. S. Chong, A. H. Fox, C. S. Bond, S. Nakagawa, G. Pierron, T. Hirose, Functional domains of NEAT1 architectural lncRNA induce paraspeckle assembly through phase separation. *Mol. Cell* **70**, 1038–1053.e7 (2018).
20. L. Standaert, C. Adriaens, E. Radaelli, A. Van Keymeulen, C. Blanpain, T. Hirose, S. Nakagawa, J.-C. Marine, The long noncoding RNA Neat1 is required for mammary gland development and lactation. *RNA* **20**, 1844–1849 (2014).
21. S. Nakagawa, M. Shimada, K. Yanaka, M. Mito, T. Arai, E. Takahashi, Y. Fujita, T. Fujimori, L. Standaert, J.-C. Marine, T. Hirose, The lncRNA Neat1 is required for corpus luteum formation and the establishment of pregnancy in a subpopulation of mice. *Development* **141**, 4618–4627 (2014).
22. S. S. Mello, C. Sinow, N. Raj, P. K. Mazur, K. Bieging-Rolett, D. K. Broz, J. F. C. Imam, H. Vogel, L. D. Wood, J. Sage, T. Hirose, S. Nakagawa, J. Rinn, L. D. Attardi, Neat1 is a p53-inducible lincRNA essential for transformation suppression. *Genes Dev.* **31**, 1095–1108 (2017).
23. S. Li, J. Li, C. Chen, R. Zhang, K. Wang, Pan-cancer analysis of long non-coding RNA NEAT1 in various cancers. *Genes Dis.* **5**, 27–35 (2018).
24. S. Nakagawa, T. Naganuma, G. Shioi, T. Hirose, Paraspeckles are subpopulation-specific nuclear bodies that are not essential in mice. *J. Cell Biol.* **193**, 31–39 (2011).
25. E. Leucci, R. Vendramin, M. Spinazzi, P. Laurette, M. Fiers, J. Wouters, E. Radaelli, S. Eyckerman, C. Leonelli, K. Vanderheyden, A. Rogiers, E. Hermans, P. Baatsen, S. Aerts, F. Amant, S. Van Aelst, J. van den Oord, B. de Strooper, I. Davidson, D. L. J. Lafontaine, K. Gevaert, J. Vandesompele, P. Mestdagh, J.-C. Marine, Melanoma addiction to the long non-coding RNA SAMMSON. *Nature* **531**, 518–522 (2016).
26. E. L. Van Nostrand, G. A. Pratt, A. A. Shishkin, C. Gelboin-Burkhart, M. Y. Fang, B. Sundararaman, S. M. Blue, T. B. Nguyen, C. Surka, K. Elkins, R. Stanton, F. Rigo, M. Guttman, G. W. Yeo, Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016).
27. D. C. Tatomer, N. D. Elrod, D. Liang, M.-S. Xiao, J. Z. Jiang, M. Jonathan, K.-L. Huang, E. J. Wagner, S. Cherry, J. E. Wilusz, The Integrator complex cleaves nascent mRNAs to attenuate transcription. *Genes Dev.* **33**, 1525–1538 (2019).
28. T. Hirose, G. Virnicchi, A. Tanigawa, T. Naganuma, R. Li, H. Kimura, T. Yokoi, S. Nakagawa, M. Bénard, A. H. Fox, G. Pierron, NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies. *Mol. Biol. Cell* **25**, 169–183 (2014).
29. Y. Wang, S.-B. Hu, M.-R. Wang, R.-W. Yao, D. Wu, L. Yang, L.-L. Chen, Genome-wide screening of NEAT1 regulators reveals cross-regulation between paraspeckles and mitochondria. *Nat. Cell Biol.* **20**, 1145–1158 (2018).
30. T. R. Albrecht, S. P. Shevtsov, Y. Wu, L. G. Mascibroda, N. J. Peart, K.-L. Huang, I. A. Sawyer, L. Tong, M. Dundr, E. J. Wagner, Integrator subunit 4 is a 'Symplekin-like' scaffold that associates with INTS9/11 to form the Integrator cleavage module. *Nucleic Acids Res.* **46**, 4241–4255 (2018).
31. A. Gardini, D. Baillat, M. Cesaroni, D. Hu, J. M. Marinis, E. J. Wagner, M. A. Lazar, A. Shilatifard, R. Shiekhattar, Integrator regulates transcriptional initiation and pause release following activation. *Mol. Cell* **56**, 128–139 (2014).
32. N. D. Elrod, T. Henriques, K.-L. Huang, D. C. Tatomer, J. E. Wilusz, E. J. Wagner, K. Adelman, The Integrator complex terminates promoter-proximal transcription at protein-coding genes. bioRxiv 725507 [**Preprint**]. 5 August 2019. https://doi.org/10.1101/725507.
33. W. J. Kent, A. S. Zweig, G. Barber, A. S. Hinrichs, D. Karolchik, BigWig and BigBed: Enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204–2207 (2010).
34. M. T. Lovci, D. Ghanem, H. Marr, J. Arnold, S. Gee, M. Parra, T. Y. Liang, T. J. Stark, L. T. Gehman, S. Hoon, K. B. Massirer, G. A. Pratt, D. L. Black, J. W. Gray, J. G. Conboy, G. W. Yeo, Rbfox proteins regulate alternative mRNA splicing through evolutionarily conserved RNA bridges. *Nat. Struct. Mol. Biol.* **20**, 1434–1442 (2013).
35. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
36. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
37. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
38. M. Ovesný, P. Křížek, J. Borkovec, Z. Švindrych, G. M. Hagen, ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics* **30**, 2389–2390 (2014).
39. F. Levet, G. Julien, R. Galland, C. Butler, A. Beghin, A. Chazeau, P. Hoess, J. Ries, G. Giannone, J.-B. Sibarita, A tessellation-based colocalization analysis approach for single-molecule localization microscopy. *Nat. Commun.* **10**, 2379 (2019).
40. A. Colaprico, T. C. Silva, C. Olsen, L. Garofano, C. Cava, D. Garolini, T. S. Sabedot, T. M. Malta, S. M. Pagnotta, I. Castiglioni, M. Ceccarelli, G. Bontempi, H. Noushmehr, TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **44**, e71 (2016).
41. C. W. Law, Y. Chen, W. Shi, G. K. Smyth, Voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).
42. L. Gautier, L. Cope, B. M. Bolstad, R. A. Irizarry, Affy—Analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
43. T. Yamazaki, T. Hirose, The building process of the functional paraspeckle with long non-coding RNAs. *Front. Biosci.* **7**, 1–41 (2015).

Acknowledgments: We thank the Proteomics Core Facility at VIB/UGent for performing label-free MS, the VIB BioImaging Core for support and assistance in this work, and the Oncogenomics Core Facility at Sylvester Comprehensive Cancer Center for performing high-throughput sequencing. We are grateful to E. J. Wagner for sharing with us the U7-GFP reporter construct. We acknowledge G. Bervoets and O. V. Goethem for technical support. **Funding:** J.B. received a PhD research fellowships from the International VIB PhD program (4 years) and by the Emmanuel van der Schueren fellowship number ZKD5324 (1 year). E.L. was supported by the King Baudouin Foundation Fund Emile Carpentier—André Vander Stricht—Van Damme 2017 J1810830-207301 and by the Belgian Federation against Cancer. T.K.K. is supported by the Developmental Origins of Chronic Diseases in Children Network (DEVOTION) of Manitoba. T.H.J. was supported by the Novo Nordisk Foundation. Research reported in this publication was supported by funding from VIB, grants R01 GM078455 and GM105754, and DP1 CA228041 from the NIH to R.S.; and the National Cancer Institute of the NIH under award number P30CA240139. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. **Author contributions:** J.B. designed and conducted experiments, acquired, analyzed, and interpreted the data. G.S.G. analyzed STORM images. E.B. performed the RNA-seq experiments and generated the STORM images. F.B. performed the 3′ mRNA-seq experiments and total RNA-seq, smRNA-seq, and 3′ mRNA-seq analyses. M.M.T. performed eCLIP experiments. N.K. processed eCLIP data. T.K.K. reanalyzed publicly available transcriptomics data and performed the PCA. E.L., R.S., and J.-C.M. conceptualized and designed research studies. J.-C.M. and E.L. wrote the manuscript. All authors read and edited the manuscript. **Competing interests:** J.B. and J.-C.M. are inventors on a patent related to this work filed by VIB (no. PCT/EP2015/052663, 4 November 2019). J.-C.M. is scientific founder and scientific adviser at Flamingo Therapeutics. The other authors declare that they have no competing interests. **Data and materials availability:** High-throughput data are deposited at the GEO under accession number GSE125534. 3′ mRNA-seq is deposited at the GEO under accession number GSE125535. The eCLIP is deposited at the GEO under accession number GSE148755. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors.

Submitted 18 October 2019
Accepted 14 April 2020
Published 3 July 2020
10.1126/sciadv.aaz9072

Citation: J. Barra, G. S. Gaidosh, E. Blumenthal, F. Beckedorff, M. M. Tayari, N. Kirstein, T. K. Karakach, T. H. Jensen, F. Impens, K. Gevaert, E. Leucci, R. Shiekhattar, J.-C. Marine, Integrator restrains paraspeckles assembly by promoting isoform switching of the lncRNA *NEAT1*. *Sci. Adv.* **6**, eaaz9072 (2020).