



Elongin A associates with actively transcribed genes and modulates enhancer RNA levels with limited impact on transcription elongation rate *in vivo*

Received for publication, September 2, 2020, and in revised form, December 6, 2020. Published, Papers in Press, December 17, 2020.

<https://doi.org/10.1074/jbc.RA120.015877>

M. Behfar Ardehali^{1,2}, Manashree Damle^{1,2}, Carlos Perea-Resa^{1,2}, Michael D. Blower^{1,2}, and Robert E. Kingston^{1,2,*}

From the ¹Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts, USA; and ²Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA

Edited by Karin Musier-Forsyth

Elongin A (EloA) is an essential transcription factor that stimulates the rate of RNA polymerase II (Pol II) transcription elongation *in vitro*. However, its role as a transcription factor *in vivo* has remained underexplored. Here we show that in mouse embryonic stem cells, EloA localizes to both thousands of Pol II transcribed genes with preference for transcription start site and promoter regions and a large number of active enhancers across the genome. EloA deletion results in accumulation of transcripts from a subset of enhancers and their adjacent genes. Notably, EloA does not substantially enhance the elongation rate of Pol II *in vivo*. We also show that EloA localizes to the nucleoli and associates with RNA polymerase I transcribed ribosomal RNA gene, *Rn45s*. EloA is a highly disordered protein, which we demonstrate forms phase-separated condensates *in vitro*, and truncation mutations in the intrinsically disordered regions (IDR) of EloA interfere with its targeting and localization to the nucleoli. We conclude that EloA broadly associates with transcribed regions, tunes RNA Pol II transcription levels *via* impacts on enhancer RNA synthesis, and interacts with the rRNA producing/processing machinery in the nucleolus. Our work opens new avenues for further investigation of the role of this functionally multifaceted transcription factor in enhancer and ribosomal RNA biology.

Transcription by RNA polymerase II (Pol II) is a multistep process, facilitated by a large number of auxiliary factors that modulate the activity of Pol II (1, 2). In addition to recruitment, pausing of Pol II molecules 20–60 bp downstream of transcription start site (TSS) introduces another regulatory layer, which controls release of Pol II into productive transcription elongation (3–5), and recent findings point to the generality of this mechanism in regulation of transcription in various metazoan organisms (6). Upon release of Pol II into productive elongation, prebound or recruited transcription elongation factors facilitate passage of Pol II through gene bodies by increasing the elongation rate and/or processivity of

Pol II (7–9). The rate of Pol II elongation has been shown to play a critical role in efficient splicing and development of the organism (10, 11). The elongation rate of Pol II has been shown to vary for different genes with measurements ranging from 1 to 4 kb/min (12).

One of these transcription elongation factors, Elongin A (EloA), was discovered as a biochemical activity that enhances Pol II elongation rate by decreasing the frequency of Pol II arrest and transient pausing in run-off transcription assays (13–15). EloA forms a heterotrimeric complex with Elongin B and Elongin C, which enhance the complex stability and specific activity of EloA, respectively (15). Internal deletion analysis of EloA identified the C-terminal region of the factor (residues ~400–773 of rat EloA) as necessary and sufficient for *in vitro* elongation activity (16). *In vivo*, EloA has been shown to play a critical role in optimal gene induction in response to environmental stress and developmental stimuli (17–20). Moreover, in addition to its role as a transcription elongation factor, EloA is also the substrate-recognition subunit of Cullin-RING E3 ubiquitin ligase complexes, with the largest subunit of Pol II, RPB1 being its main client. Assembly of this complex is triggered upon genotoxic stress and DNA damage, which targets stalled Pol II for degradation (21).

While the role of EloA in transcription elongation by RNA polymerase II has been examined in great detail *in vitro* (14–16), there are limited reports on the role of Elongin A in transcription *in vivo*. Recently we showed that Elongin A is methylated by PRC2 at K754, a highly conserved residue within a region predicted to form coiled-coil protein interaction domain (16, 22). Here we characterize the *in vivo* function of EloA and show that it localizes to thousands of transcriptionally active protein coding genes in mESC with a strong presence at the TSS/Pol II pause site and downstream of the polyadenylation signal. Surprisingly, evaluation of the Pol II elongation rate upon reactivation of productive elongation revealed that EloA has minimal effects on transcription elongation by Pol II *in vivo*. We also observed that EloA is present at many transcriptionally active enhancers and superenhancers and that its loss results in accumulation of enhancer RNA (eRNA) at a subset of these, as well as elevated levels of nascent RNA at adjacent genes. Furthermore, we show that

This article contains supporting information.

* For correspondence: Robert E. Kingston, kingston@molbio.mgh.harvard.edu.

EloA enrichment at active genes and enhancers

EloA is associated with RNA polymerase I transcribed ribosomal RNA gene, *Rn45s*, and is present in the nucleoli. EloA is a highly disordered protein, and we demonstrate that purified EloA forms phase-separated condensates *in vitro* and that truncations within the highly disordered region interfere with targeting and localization of EloA to the nucleoli. Our work here uncovers new aspects of EloA function by showing its widespread association with genic and intergenic transcribed regions, its impact on eRNA synthesis, and its presence on rRNA genes and in the nucleolus.

Results

Elongin A shows widespread association with actively transcribed genes

The association of EloA with Pol II transcribed genes has been tested at a small number of housekeeping and stress response genes in cells expressing tagged EloA (20). To gain a comprehensive understanding of the EloA genome-wide distribution and its role in transcription by Pol II, we carried out ChIP-seq, as well as Cleavage Under Targets and Release Using Nuclease (CUT&RUN) (23, 24) on endogenous EloA in wild-type and *EloA*^{-/-} mESCs. We characterized EloA protein levels in wild-type and *EloA* null mES cells by fractionation and by immunoblotting. We found that, as expected, EloA is predominantly a nuclear protein (Fig. 1A). While ChIP-seq and Cut&Run-seq peak calling algorithms both identified many of the robustly expressed genes as being bound by EloA, Cut&Run-seq offered a superior signal/noise ratio for the remainder of genes, allowing for a more comprehensive profiling of EloA binding sites (Fig. S1, A–B).

CUT&RUN-seq of EloA identified 17,753 peaks (SEACR, FC > 1.5 over null) throughout the genome, more than 70% of which localized to different regions of annotated genes, with a notable bias toward the promoter region (Fig. 1B). We established that the EloA CUT&RUN-seq signal is specific by comparing it with that of EloA and IgG CUT&RUN-seq signals in null and wild-type cells, respectively (Fig. 1C). EloA was present at the majority of expressed genes ($n = 6198$, RPKM > 0.5), while 4464 expressed genes were not bound by EloA. Association of EloA at many highly expressed genes exhibited a bimodal distribution pattern, with one strong peak around the TSS/Pol II pausing site and another peak at the 3'-end, downstream of cleavage/polyadenylation signal site (Fig. 1, C–D). We grouped EloA-bound genes into three classes based on enrichment levels (Fig. 1D) and found a strong positive correlation with transcription activity of target genes, as determined by profiling nascent transcription by Bru-seq (Fig. 1E). These observations are consistent with an independent report in human DLD1 cells (Wang *et al.*, submitted). Expressed genes enriched for EloA also had higher levels of histone modification marks associated with active genes in both the promoter regions (H3K27ac, H3K4me3) and gene bodies (H3K36me3, H3K79me2) (Fig. 1F). Molecular pathway analysis of EloA positive genes found that mRNA processing, translation, and pluripotency gene classes were

enriched; all categories showing high expression in stem cells (Fig. 1G). We also examined nascent transcription by Bru-seq in wild-type and null EloA mESCs and found a small effect on gene expression with 679 upregulated and 198 genes downregulated in null cells (FC > 1.5) (Fig. S1C). We conclude that EloA binds to a large number of actively transcribed genes and that loss of EloA has a rather minor effect on gene expression under steady-state conditions, predominantly leading to upregulation of a subset of genes.

In mESCs, roughly half of actively transcribed genes have been shown to be regulated by Pol II pausing early in the transcription process (25). We also noted a positive correlation between EloA enrichment and Pol II pausing, where paused genes displayed elevated levels of EloA near the promoter and around the pause site in comparison with nonpaused genes (Fig. S1, D–E). We also noted a moderate decrease in the levels of Pol II-Ser2P in *EloA*^{-/-} cells (Fig. S1F).

Elongin A is enriched at active enhancers and regulates their transcription

Enhancers are mainly distal regulatory elements that harbor binding sites for transcription factors and which, irrespective of orientation, are capable of activating gene expression over large distances (26, 27). We were intrigued by the fact that a quarter of our identified EloA peaks (4573/17,753) were located outside of annotated genes (Fig. 1B) and asked whether these intergenic EloA binding sites show any overlap with annotated intergenic enhancers in mESC. Interestingly, out of 4573 intergenic EloA peaks, 40% (1841) overlapped with annotated enhancers, with 361 of these sites localizing to clusters of enhancers, commonly referred to as superenhancers (Fig. 2A) (28). We ranked EloA-bound enhancers based on the intensity of EloA enrichment and found a positive correlation between EloA binding and enrichment of Sox2, Nanog, and Oct4, all master transcription factors recruited to enhancers and required for establishment and maintenance of pluripotency in ESCs (Fig. 2B). Interrogation of the chromatin landscape at EloA-bound enhancers also revealed positive correlation between EloA levels and intensity of H3K27ac and H3K4me3 at these sites (Fig. 2, B–C), both modification marks delineating active enhancers (29). Whereas, H3K4me1, an enhancer mark associated with both active and poised enhancers, did not display strong correlation with the intensity of EloA binding (Fig. 2, B–C).

Active enhancers are transcribed and generate unstable, short-lived RNAs termed enhancer RNAs (eRNAs) (29–31). Given that EloA is a Pol II transcription factor and that the intensity of EloA binding at enhancers positively correlated with H3K4me3 levels, a mark of transcriptionally active enhancers, we examined the role of EloA in transcription from these distal regulatory elements and measured nascent transcription at enhancers by Bru-seq (± 1 kb EloA peak).

Interestingly, our analysis of nascent RNA transcription at EloA-bound annotated enhancers revealed misregulation in eRNA levels at 40.8% (751/1841) of these sites in *EloA* null

EloA enrichment at active genes and enhancers

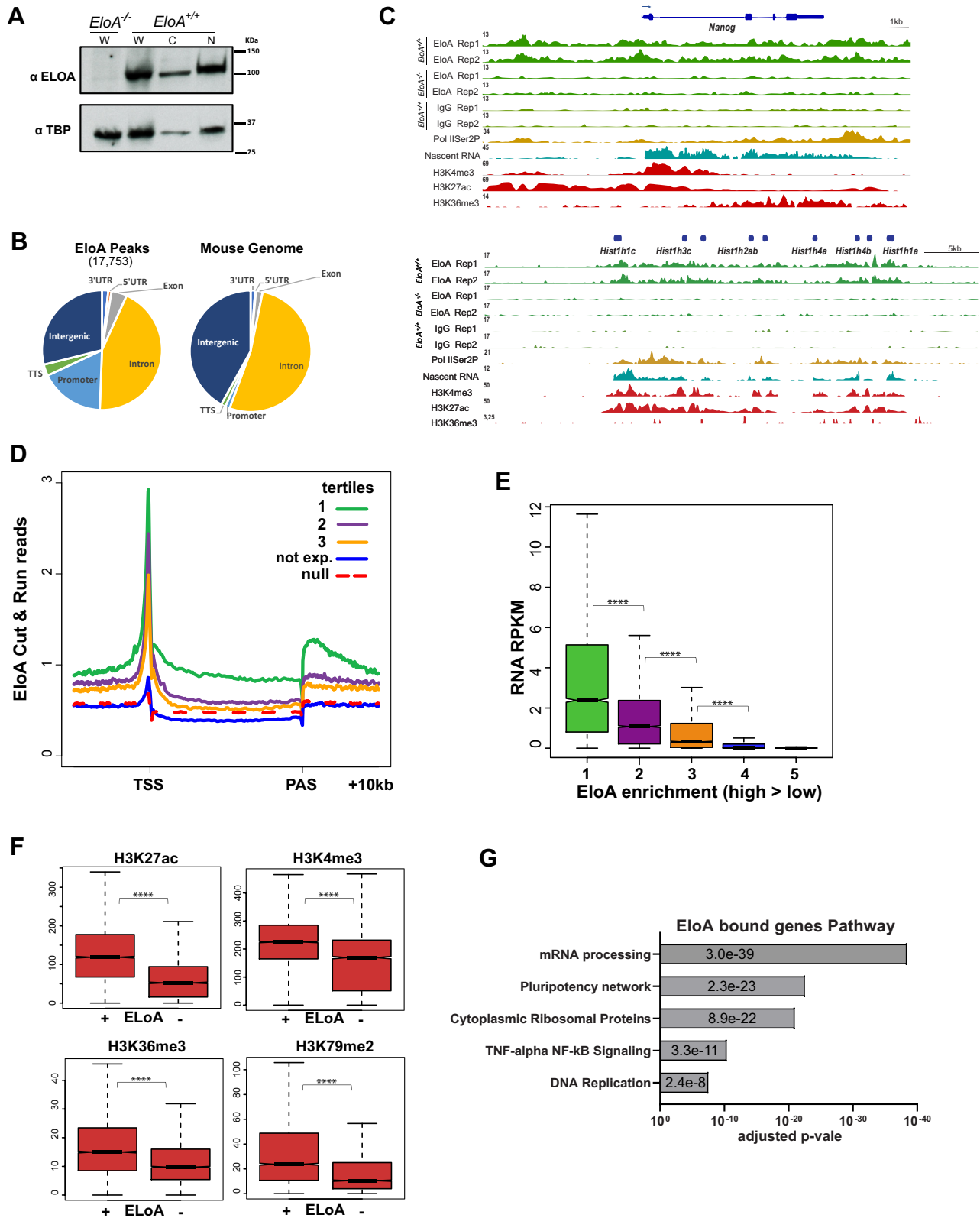


Figure 1. Elongin A is present at a large number of actively transcribed genes. A, immunoblot of whole cell (W), cytoplasmic (C), and nuclear (N) extract of wild-type and *EloA* null mESC, showing nuclear localization of EloA, TBP immunoblot serves as loading and nuclear localization control. B, pie chart depicting distribution pattern of 17,753 EloA C&R-seq peaks (FC > 1.5 over null) over known genomic features and its strong association with 5' ends/promoter region of genes. Pie chart on right illustrates distribution of sequences and genomic features across the mouse genome. C, EloA Cut&Run-seq genome browser profiles of representative genes, *Nanog* and a histone gene cluster, depicting strong association of EloA with the promoter and post PAS of *Nanog* and widespread association across the transcribed histone gene cluster. D, metagene plot of EloA Cut&Run-seq grouped into tertiles based on enrichment level, showing strong TSS and post polyadenylation signal (PAS) enrichment for EloA. E, box plot showing positive correlation between EloA enrichment (x-axis) and transcription levels (y-axis, nascent RNA levels). F, box plots showing strong positive correlation between EloA Cut&Run-seq peaks and histone modifications associated with active transcription. G, molecular pathways showing strongest enrichment as determined by wikipathways. p-values for (E-F) derived from Wilcoxon test, ****p < 0.0001.

EloA enrichment at active genes and enhancers

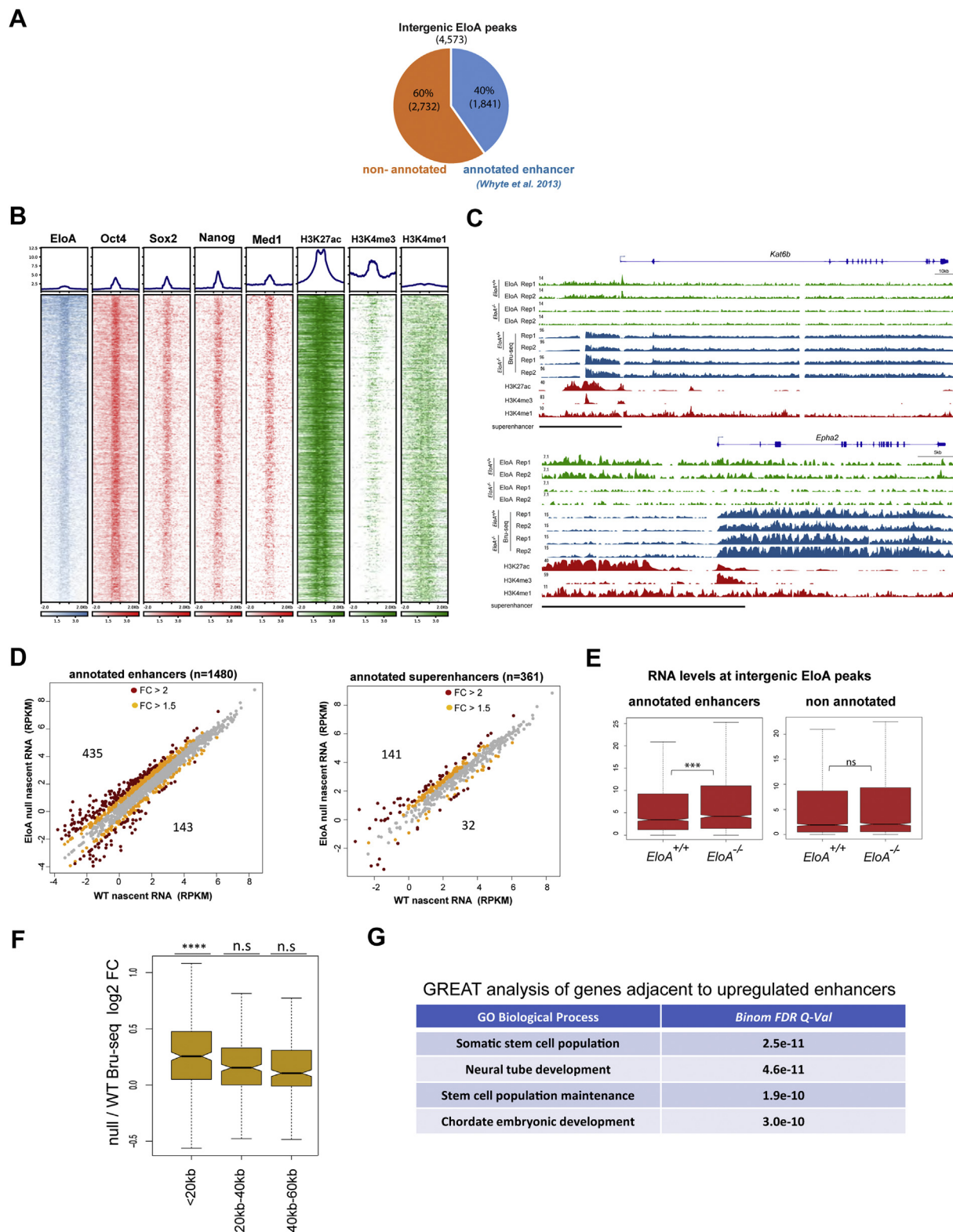


Figure 2. Elongin A is enriched at transcriptionally active enhancers and superenhancers. *A*, pie chart depicting the degree of overlap between 4573 intergenic EloA peaks and annotated enhancers as described by Whyte *et al.* (28). *B*, heat map of EloA, TFs (Oct4, Sox2, Nanog), mediator subunit, Med1, and histone modifications (H3K27ac, H3K4me3, H3K4me1) occupancy rank-ordered by decreasing EloA enrichment displaying positive correlation between intensity of EloA binding and master TFs, histone H3K27ac and H3K4me3 levels ($n = 4573$). *C*, genome browser screenshots depicting enrichment of EloA and eRNA upregulation at two representative superenhancers upstream of *Kat6b* and *Epha2* genes. *D*, scatter plot showing eRNA nascent RNA levels as detected by Bru-seq in WT and null mESC (left panel: enhancers, right panel: superenhancers). Numbers within the scatterplot denotes number of up or downregulated eRNAs ($FC > 1.5$). *E*, box plot showing eRNA levels at EloA-bound annotated enhancers and nonannotated sites. *F*, box plot showing the ratio of null over wild-type nascent RNA signal for genes <20 kb, between 20–40 kb and 40–60 kb away from enhancers that are upregulated in *EloA* null cells. *p*-values for (*E–F*) derived from Wilcoxon test, $***p < 0.001$, $****p < 0.0001$. *G*, table depicting Genomic Regions Enrichment of Annotations Tool (GREAT) analysis results for identification of gene classes adjacent to enhancers bound by EloA and upregulated in *EloA*^{-/-} cells.

cells. Strikingly, the majority of these differentially expressed enhancer elements (76.6%) had elevated levels of eRNA transcription in *EloA*^{-/-} mES cells (576/751, FC > 1.5) (Fig. 2D). In contrast only 3.6% of EloA-bound genic regions show differential expression after loss of EloA (225/6198) (Fig. S1G). The observed upregulation of transcription at EloA-bound intergenic loci was statistically significant for only annotated enhancers and not for other intergenic EloA-bound sites (Fig. 2E). We also noted that genes in the immediate vicinity (<20 kb) of these upregulated enhancers had higher levels of nascent transcription in *EloA* null cells (Fig. 2F). We asked whether these differentially expressed enhancers in *EloA* null cells are adjacent to any specific class of genes and performed Genomic Regions Enrichment of Annotations Tool (GREAT) analysis to address this. Genes adjacent to downregulated enhancers in *EloA*^{-/-} did not show a statistically significant correlation with any GO biological process; however, positive correlation was observed between eRNA misaccumulation in *EloA*^{-/-} mES cells and genes involved in maintenance of pluripotency as well as neural tube and embryonic development (Fig. 2G). In addition to enrichment of EloA at many of these distal intergenic enhancers, the CUT&RUN-seq identified EloA peaks at 469 out of 1912 intragenic enhancers described by Cinghu *et al.* (32) in mESC (Fig. S2, A–B). We conclude that EloA is recruited to both intergenic and intragenic enhancers and that its loss results in increased eRNA nascent transcription at several of these sites as determined by Bru-seq. This increased eRNA synthesis correlated with upregulation of adjacent protein coding genes.

***EloA* has a negligible effect on transcription elongation by pol II in vivo**

EloA was initially purified as an activity that stimulates transcription elongation by Pol II *in vitro* (13, 14). However, the extent to which EloA facilitates Pol II elongation *in vivo* has not been explored. To examine the role of EloA in transcription elongation in mESCs, we shut down productive transcription elongation by treating the cells with flavopiridol (FP), a potent P-TEFb inhibitor blocking Pol II C-terminal domain (CTD) Ser2 phosphorylation (33), and measured movement of Pol II within the body of the gene upon removal of the drug. To minimize secondary effects, FP treatment was carried out for only 60 min. After washing out FP, cells were incubated in presence of 5-bromouridine (Bru) for either 5 or 10-min, and nascent RNA was immunopurified and sequenced (Fig. 3A). We estimated the position where the read density of nascent transcription approaches background levels on the gene body to be 10–12 and 22–24 kb downstream of the TSS for the 5 and 10-min recovery samples, respectively (intersection of the dotted lines with *x*-axis, Fig. 3, B–C). This corresponds to an elongation rate of 2–2.4 kb/min, which is consistent with previously reported Pol II elongation rates for genes in mESC (12).

The density profile of nascent RNA (transcribing Pol II) immediately downstream of the TSS (+1 kb) at EloA-enriched genes was similar or even slightly higher in *EloA* null cells

(Figs. 3, B–C and S2C); however, we noted a discernable, yet modest (~10%) decrease in the density of reads toward the “wave front” of transcription in the 5-min recovery sample. At 5 min the pioneering wave of Pol II appeared to be lagging by about 1 kb past the 10 kb mark in *EloA*^{-/-} cells (Fig. 3B). A decrease in nascent RNA signal in null cells was also observed in the 10-min recovery samples (Fig. 3C); however, this decrease did not appear more prominent near the leading edge of Pol II transcription (~+20 kb) but was more notable between the 5 and 10 kb mark. This latter phenotype is not consistent with a substantive role for EloA in modulation of Pol II elongation rate; a more striking difference toward the “wave front” of transcription would be expected. Rather, this suggests that in addition to a potential role in transcription elongation, EloA might also be acting to increase the efficiency of a step upstream of productive elongation such as release from pause, Pol II loading, or initiation. Further, a defect in Pol II processivity would be predicted to cause a drop in the density of transcribing Pol II toward the end of the gene. To examine this issue, we compared the TSS distal to TSS proximal steady-state nascent RNA levels for the same set of genes and did not detect a larger decrease in the density of nascent transcription toward the middle and end of the gene body in null cells (Fig. 3, D–E). Figure 3F depicts movement of Pol II along the body of two representative genes, *Eed* and *Dhx9*, at various time points after transcription reactivation. From these experiments we conclude that the loss of EloA does not cause a significant global change in RNA polymerase processivity. Furthermore, the effect of EloA on Pol II elongation is limited and that the slight decrease in the density of pioneering Pol II wave upon transcripton reactivation might be mediated at a stage upstream of productive transcription elongation. Our observation regarding the limited effects of EloA on Pol II elongation rate is consistent with an independent study in human DLD1 cells line (Wang *et al.*, submitted).

EloA also shows strong enrichment downstream of polyadenylation signal (PAS) (Fig. 1, C–D). To examine whether EloA plays role in Pol II termination, we also looked at the levels of nascent RNA up to 10 kb downstream of transcripton end site and noted no difference in the density of Bru-seq signal (Pol II levels) between WT and *EloA* null cells (Fig. S2F).

Characterization of EloA binding partners

To better understand the biological processes that EloA may be participating in, we sought to identify the protein interaction network of EloA. We carried out coimmunoprecipitation mass spectrometry (Co-IP/MS) analysis on endogenous EloA from MEF-depleted mESC nuclear extract (Fig. 1A). As expected, EloA showed the strongest interaction with EloC and EloB, the other subunits of the heterotrimeric Elongin complex (Fig. 4A), as well as the putative RNA exonuclease Rexo1 (EloA-BP1), a previously reported major binding partner of EloA (34). Our results also revealed strong binding between Elongin A and PAF1, LEO1, CDC73, and WDR61(SKI8) (Fig. 4, A–B), all components of the multifunctional polymerase-associated factor 1 complex (Paf1C), which is involved in regulation of various

EloA enrichment at active genes and enhancers

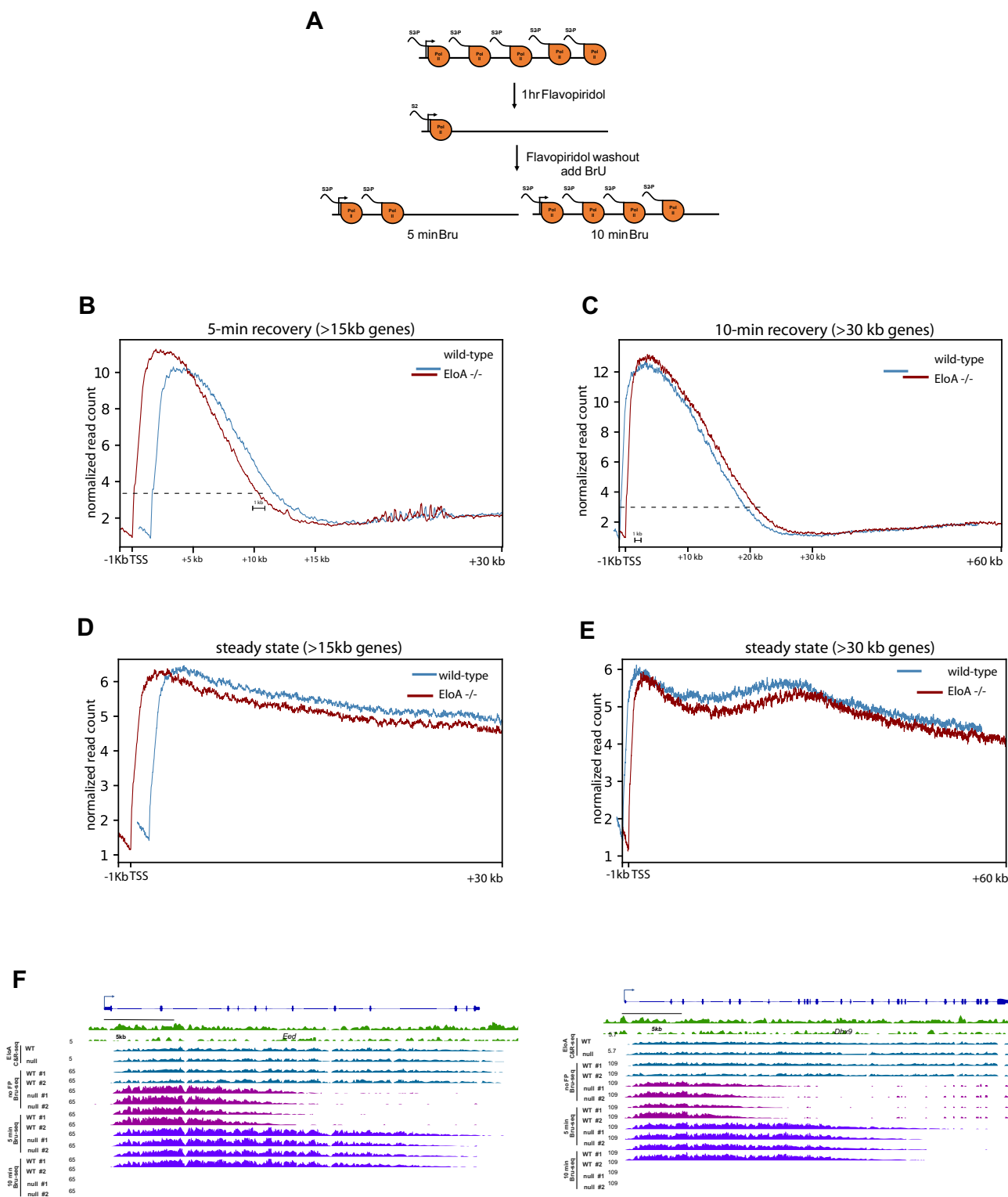


Figure 3. Examining the role of EloA in transcription elongation by Pol II *in vivo*. *A*, experimental outline for measurement of Pol II elongation rate. *B–C*, line plots showing read density of nascent transcription 5-min (*B*) and 10-min (*C*) after release from FP block. Only transcribed EloA-bound genes longer than 15 (5-min) and 30 kb (10-min) were used for measurement of transcription elongation. *Dotted lines* denote the position where nascent RNA signal of Pol II “wave front” approaches background levels for the purpose of estimating the elongation rate (lines are average of two independent experiments). *D–E*, line plots showing nascent transcription density (Pol II levels) during steady-state transcription in the absence of FP. Normalized reads: reads per million/bin. *F*, IGV tracks of *Eed* and *Dhx9*, two representative genes showing distribution of nascent transcripts along the gene body 5 and 10-min after release from FP block. The no FP blue tracks depict steady-state levels of nascent RNA.

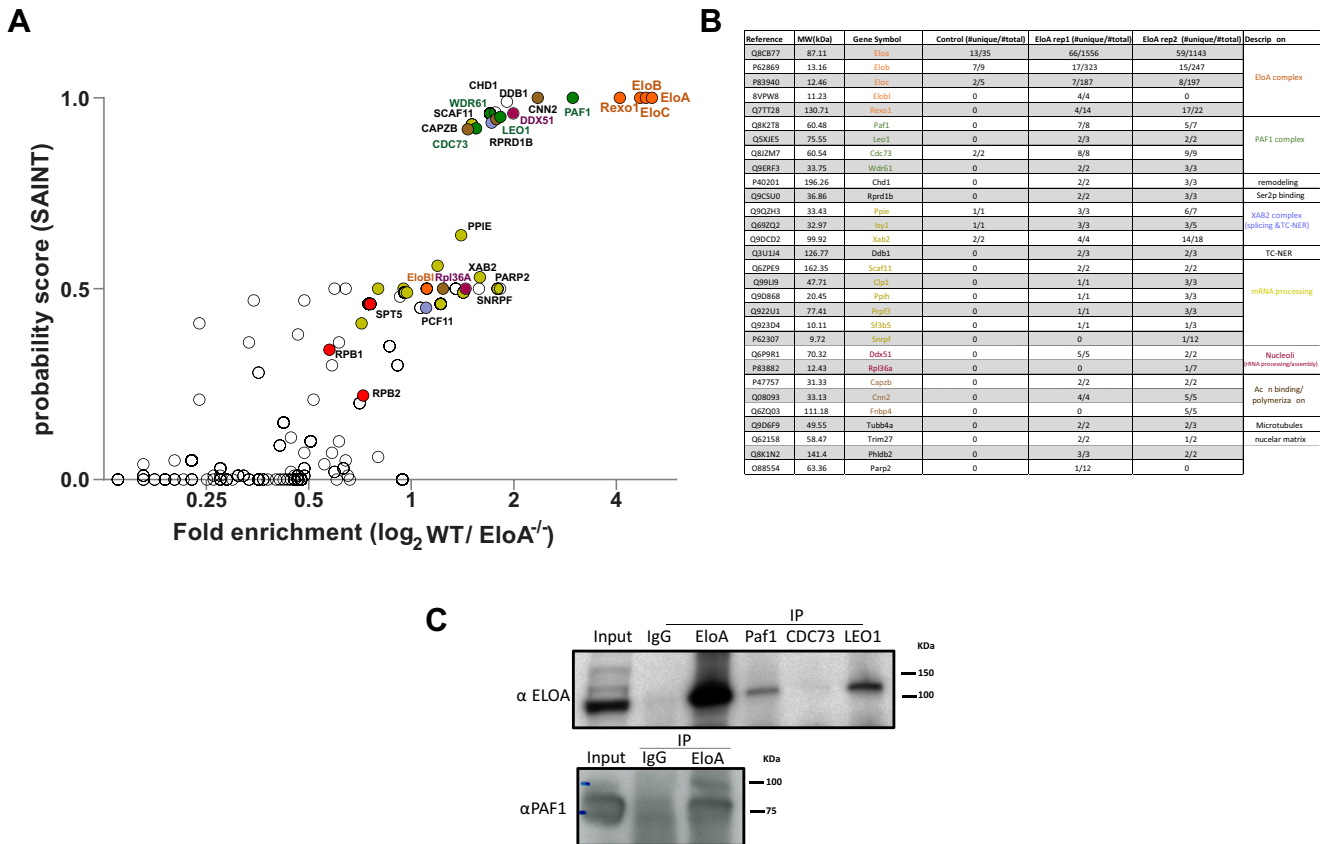


Figure 4. Characterization of Elongin A proteome interaction network. A, affinity purification mass spectrometry (Co-IP/MS) analysis of native Elongin A binding partners from mESC nuclear extract showing major interaction between EloA complex (orange) and Paf1 complex subunits (green) as well as pre-mRNA processing and splicing factors (olive green), and cytoskeletal filament interacting proteins (brown). x-axis denotes \log_2 WT/null EloA enrichment and y-axis displays SAINT (significance Analysis of INteractome) score. B, total peptide count from each pull-down for a selected group of factors. Control IP was performed using EloA antibody on MEF depleted EloA null mESC. Leftover MEF cells have likely contributed to EloA peptide count in control sample. C, IP immunoblot experiment for validation of interaction between EloA and PAF1 and other Paf1 complex components. List of all identified proteins and their corresponding peptide count, and the raw data are accessible under ProteomeXchange ID: PXD022138.

stages of transcription by Pol II (35). We validated these interactions by IP-Immunoblot experiments (Fig. 4C). This observation is concordant with recent reports where pull-down of Paf1C identified EloA as a major interacting partner of the complex in murine myoblast and ESC (7, 36). The protein interactome analysis uncovered evidence for physical interaction between EloA and chromatin remodeling factor CHD1, as well as pre-mRNA processing factors, in particular splicing factors (Fig. 4, A–B). There were also interactions between EloA and nucleolar proteins (Ddx51, Rpl36a), a finding relevant to the studies of nucleoli we describe below. We found multiple components of the Xeroderma pigmentosum group A (XPA)-binding protein 2 (XAB2) complex (XAB2, ISY1, and PPIE), a multifunctional complex involved in splicing and transcription-coupled DNA repair among EloA interacting proteins (37). The IP was performed on nuclear extracts without including the solubilized chromatin fraction, which in part explains lower levels of interaction with Pol II and other transcription factors (Fig. 4A, red dots).

EloA is present at sites of RNA polymerase I transcription

The 45S pre-rRNA is transcribed by RNA polymerase I (Pol I) and processed into 18S, 5.8S, and 28S rRNAs. Interestingly, we observed association of EloA with segments that map to

Rn45s, a 45S coding gene on chromosome 17. The EloA CUT&RUN-seq signal was specific for wild-type mES cells and close to IgG background levels in *EloA* null cells, and we did not detect enrichment of Ser2P Pol II at this rRNA gene site (Fig. 5A). We carried out coimmunofluorescence microscopy on 3T3 fibroblast cells and observed colocalization between EloA and NOLC1 (Nopp140) (Fig. 5B), a nucleolar factor, present at the fibrillar center (FC) and dense fibrillar component (DFC), which are sites for rRNA transcription and processing, respectively (38, 39). This finding is in agreement with our observation that nucleolar factors Rpl36a and DDX51 were identified as interacting partners of EloA (Fig. 4, A–B).

Nucleolar proteins have been shown to display rapid exchange dynamics between the nucleolus and the nucleoplasm (40), and a subset of tested nucleolar proteins possess liquid-like *in vivo* dynamics related to the known liquid-liquid phase separating properties of nucleoli (41). We characterized the nucleolar dynamics of EloA using a system in which monomeric red fluorescent protein tagged EloA (mCherry-EloA) was expressed from a doxycycline (dox)-inducible construct (Fig. S3B). Consistent with the immunofluorescence results on endogenous EloA, mCherry-EloA was present in the nucleolus in addition to the nucleoplasm (Fig. 5C, upper panel). We examined the exchange kinetics of mCherry-EloA

EloA enrichment at active genes and enhancers

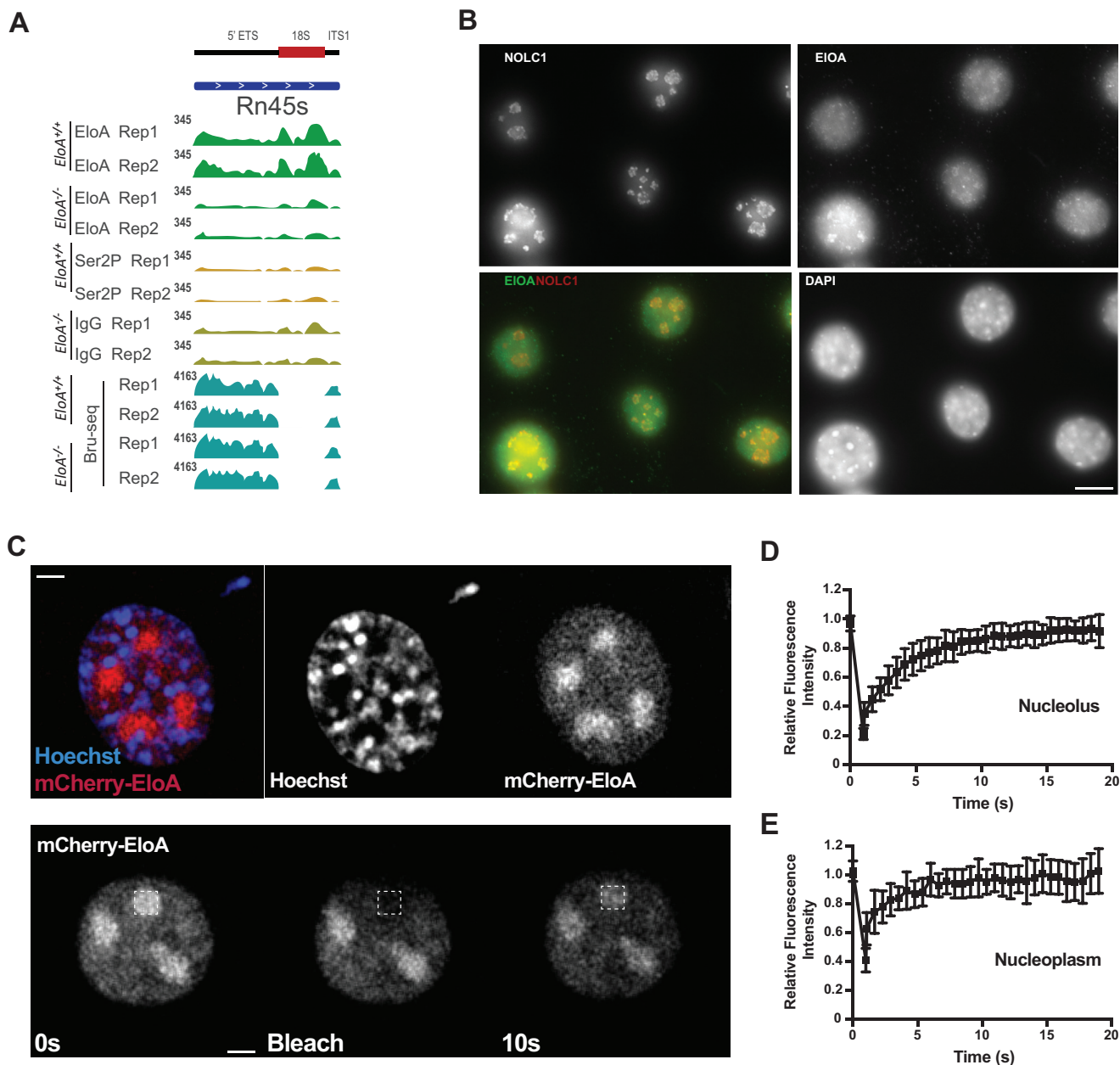


Figure 5. Elongin A associates with Pol I transcribed ribosomal genes. *A*, IGV genome browser screen shot of Pol I transcribed 45S peribosomal RNA, *Rn45s* located on chromosome 17 showing enrichment of EloA as determined by Cut&Run-seq. Background level of EloA in null mESC serves as a control for specificity of EloA enrichment. *B*, coimmunofluorescence micrograph (z-stacks) of formaldehyde fixed 3T3 cells showing colocalization of EloA and nucleolar/Cajal bodies marker protein NOLC1. Scale bar, 11 μ m. *C*, FRAP analysis of nucleolar mCherry-EloA demonstrating the dynamic nature of EloA within this organelle. Top panel: Live cell image of mCherry-EloA (red) in 3T3 cells and DNA (Hoechst, blue). Bottom panel: Bleached cells imaged for fluorescence recovery. Images of prebleach, bleached, and 10 s post recovery are shown from representative cell. Scale bar: 3 μ m. *D* and *E*, relative fluorescence intensity between bleached (squares) and nonbleached reference areas plotted over time. Recovery analysis of nucleolar (*D*) and nucleoplasmic (*E*) signal showing distinct recovery profiles. $n \geq 20$ from two independent experiments. error bars denote SD.

in 3T3 cells to determine whether it showed exchange characteristics similar to that of other nucleolar proteins. Fluorescence recovery after photobleaching (FRAP) analysis of EloA molecules in the nucleoli showed rapid recovery of photobleached EloA (10–15 s) post photobleaching (Fig. 5, *D–E*), demonstrating the dynamic nature of EloA in the nucleolus. Nucleolar EloA recovery was slower compared with the nucleoplasmic mCherry-EloA recovery profile (5 s), which might represent the diffusion dynamics of nucleoplasmic EloA. The slower recovery rate of nucleolar mCherry-EloA might

reflect interaction of EloA with binding partners within the nucleolus. We conclude that the nucleolar fraction of EloA possesses dynamics in cell culture that are consistent with those of previously described nucleolar proteins (40, 42).

Elongin A phase separates and forms condensates in vitro

The liquid–liquid phase separation properties (LLPS) of the nucleolus have been proposed to be driven in part by proteins with large intrinsically disordered regions (IDR) that facilitate

multivalent interactions and formation of phase separated condensates (41). Many of these proteins also interact with RNA, and this interaction is thought to play role in compartmentalization of this highly dynamic nuclear body (41, 43, 44). Analysis of Elongin A protein sequence using predictor of natural disordered regions (PONDR) revealed that 82% of EloA is predicted to be disordered (Fig. 6A). Given the localization of EloA in the nucleolus, we examined the ability of EloA to exhibit properties of phase separation *in vitro* and whether this might be related to nucleolar localization.

To determine whether EloA has any of the properties of phase separating proteins, we purified full-length EloA and demonstrated that EloA forms a turbid solution in a concentration-dependent manner (Fig. 6B). We also expressed and purified monomeric enhanced GFP (mEGFP) tagged EloA (mEGFP-EloA WT) (Fig. S3A) and found that upon centrifugation, mEGFP-EloA separates from solution and forms condensate pellets (Fig. 6C). To examine whether the IDR region of EloA plays a role in phase separation, we purified a truncated EloA construct where three highly disordered stretches of amino acids within the IDR were deleted (mEGFP-EloA IDR-skip). This truncation interfered with pellet formation and the protein predominantly remained in solution (Fig. 6C). Furthermore, using fluorescence microscopy, we demonstrated that, in the presence of a volume excluder, mEGFP-EloA forms protein-rich spherical droplets in a concentration-dependent manner (Fig. 6D), whereas the IDR truncation mutant failed to form these protein-rich foci and had diffuse distribution even at the higher tested concentration (Fig. 6D). We conclude that EloA is able to form phase-separated condensates *in vitro* and that truncations within the highly disordered stretches of EloA IDR appear to interfere with phase separation properties of the protein. We noted the presence of a smaller protein band in our mEGFP-EloA IDR-skip protein preparations that might reflect stability issues of the truncated protein and may have impacted our *in vitro* results (Fig. S3A, right panel).

Next, we introduced doxycycline (Dox)-inducible wild-type and IDR-skip mutant constructs in *EloA* null 3T3 fibroblast cells using lentiviral transduction and determined the doxycycline concentration required for expression of the constructs at the endogenous levels (Fig. S3, C–D). Similar to endogenous EloA (Fig. 5B), Dox-induced wild-type EloA localized to the nucleoli (Fig. 6E, upper panel). Interestingly, EloA IDR-skip appeared to be partly excluded from the nucleoli and displayed a more diffuse nuclear distribution pattern (Fig. 6E, lower panel). We conclude that the highly disordered stretches of amino acid within EloA IDR play a role in localization of EloA to the nucleolus. This might reflect direct effects of these disordered domains on phase separation, as noted above, and also could be influenced by possible impact of this mutation on interactions with other proteins such as EloB and EloC, although the interacting BC box remains intact in the truncated EloA construct.

Discussion

The role of EloA in transcription elongation by Pol II on DNA templates *in vitro* has been examined in great detail

(14–16); however, its role in Pol II transcription *in vivo* has remained underexplored. Here we show that EloA is enriched at thousands of transcriptionally active genes in mESC and that it shows a strong association with TSS and sites of Pol II pausing (Fig. 1). This is consistent with recent findings from an accompanying report (Wang *et al.*, submitted), as well as previous immunofluorescence (IF) analysis showing greater overlap between EloA and Ser5P form of Pol II, as opposed to Ser2P (20). We also observed that loss of EloA resulted in disruption, mostly upregulation, of a few hundred genes in mESC, which is in part in agreement with previous reports (45) and recent findings in DLD1 cells (Wang *et al.*, submitted). While loss of EloA has little effect on steady-state gene expression, it has been shown to play a critical role in transcription induction in response to thermal and genotoxic stress (17, 19, 20), as well as activation of developmental genes in response to developmental cues (18, 46). Beyond its role as a transcription factor that is primarily associated with TSS and sites of Pol II pausing, EloA might also be involved in certain steps of signal transduction upon encountering these external stimuli. Future genome-wide experiments studying the role of EloA in response to external stimuli should shed more light on the function and importance of this factor in these events.

Notably, we observed that EloA is enriched at intragenic and intergenic enhancers, which are regulated by Pol II pausing (29, 32) and that its loss in ES cells results in accumulation of transcripts from many of these annotated enhancers and superenhancers, as well as elevated pol II transcription of adjacent genes (Fig. 2). Recently it was shown that PAF1 negatively regulates enhancer transcription *via* controlling the release of paused Pol II and that its loss leads to upregulation of eRNA transcription (47). Moreover, other studies indicated that the integrator complex plays a role in processing and regulation of eRNA transcription and that its depletion results in accumulation of unprocessed eRNA and increased productive elongation (48, 49). EloA shows strong physical interaction with both of these aforementioned complexes (this work and Wang *et al.*, submitted), and we observed that its loss also leads to accumulation of nascent eRNA. Future experiments that examine the extent of interplay between EloA, PAF1, and the integrator complex would further elucidate the mechanism of enhancer RNA transcription and regulation.

We observed an increase in the level of nascent transcription immediately downstream of genic TSS (+0.5 kb), as well as upstream of divergent antisense transcription units in null cells (Fig. S2, C–D), but not toward the TES in null cells (Fig. 3, D–E). One explanation for this could be that EloA negatively regulates Pol II pausing and residence time of paused Pol II and that its loss may lead to rapid turnover of paused Pol II and an increase in nascent transcription immediately downstream of the TSS. We also noted strong enrichment of EloA over the promoter and regulatory regions of actively transcribed genes (Figs. 1D and S2E), suggesting that EloA might have preference for the upstream nucleosome depleted region (NDR), which harbors regulatory elements. The presence of EloA at these sites might also reflect its role in transcription from upstream divergent (antisense) sites, as loss of EloA also

EloA enrichment at active genes and enhancers

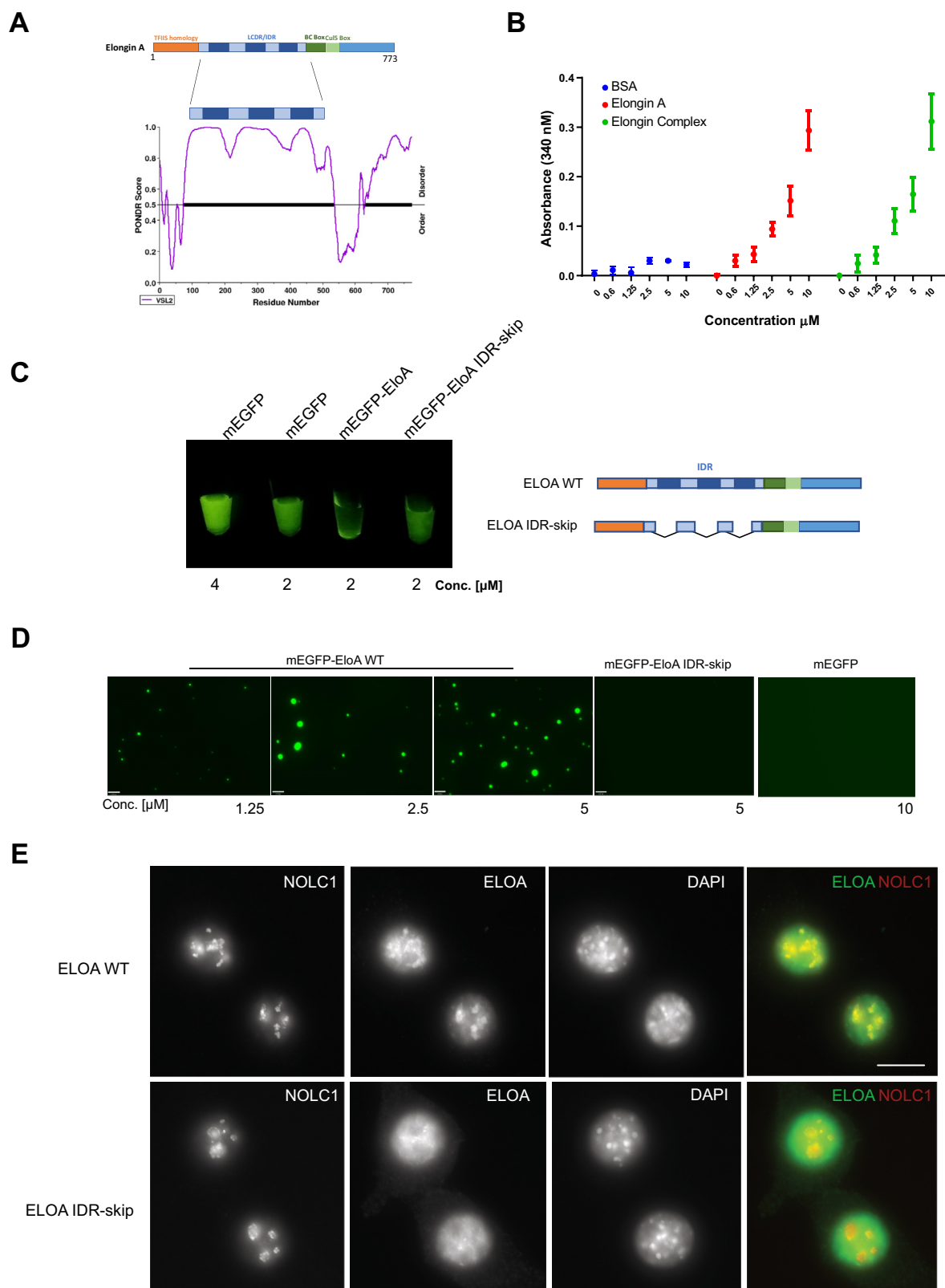


Figure 6. Elongin A forms phase-separated condensates *in vitro*. *A*, Top, schematic representation of Elongin A domains with the highly disordered stretches of amino acids in the low complexity disordered region LCDR (IDR) depicted in *dark blue*. Bottom, predictor of natural disordered regions (PONDNR) graph showing intrinsic disorder regions of EloA using the VSL2 algorithm. *B*, turbidity assay illustrating formation of turbid solution by purified EloA and EloA complex in a concentration dependent manner at 100 mM KCl buffer concentration, BSA used as negative control. $n = 3$, error: SD. *C*, spin-down assay demonstrating formation/separation of protein condensates by mEGFP-EloA, mEGFP-EloA IDR-skip truncation construct remains predominantly in solution. Right: illustration showing truncation within the IDR for the EloA IDR-skip construct. *D*, representative micrographs showing formation of phase-separated condensates by mEGFP-EloA WT at increasing concentrations in the presence of PEG, while mEGFP-EloA IDR-skip fails to form these droplets under the same conditions. Scale bar: 6 μm . *E*, truncation of EloA IDR interferes with formation of EloA nucleolar condensates *in vivo*. Micrograph of 3T3 fibroblast cells upon 50–100 ng/ml Doxycycline induction of EloAWT (top) and EloA IDR-skip (bottom) constructs. Images represent z-stacks from formaldehyde fixed cells immunostained using EloA antibody (green), nucleolar marker, NOLC1 (red) and DNA (DAPI, blue). Scale bar: 11 μm .

leads to upregulation of transcription from the negative strand at upstream divergent transcription sites (Fig. S2D). This observation indicates a possible role for EloA in regulation of antisense transcription.

Under our experimental setting, we estimated the global elongation rate of Pol II to be in the range of 2–2.4 kb/min. This is concordant with previously reported Pol II elongation rates in mESC by studying the “receding wave” of Pol II upon transcription shutdown (12). Elongin A was identified as an activity that significantly stimulates transcription elongation by Pol II *in vitro*. However, we observed a rather modest (~10%) drop in the elongation rate of Pol II in *EloA* null cells upon reactivation of transcription (Fig. 3). Moreover, the decrease in nascent RNA levels in *EloA* null cells in the 10-min sample did not appear accentuated around the “pioneering wave” of transcription and was more prominent upstream. This is not consistent with the phenotype expected for a factor that modulates elongation rate of Pol II (Fig. 3C). Therefore, it is possible that EloA also plays role in steps upstream of productive elongation. There could be several explanations for this discrepancy between *in vitro* and *in vivo* effects of EloA on transcription elongation rate. First, *in vitro* transcription assays are performed at limiting NTP concentrations and monovalent salt concentrations that are below physiological levels (50). The latter could drive condensate formation by highly disordered proteins such as EloA and induce macromolecular crowding, which on top of the elongation activity of EloA might further increase rates of transcription by Pol II *in vitro*. Second, our experiments were carried out in *EloA* null cells, which might over time adjust and adapt to loss of EloA. Moreover, redundancy and compensation by other transcription factors in the highly complex chromatin environment within living cells could also account for this modest effect on transcription elongation. Third, it is plausible that the role of EloA in stimulating transcription elongation is not global and only limited to a subset of genes with unique regulatory features (e.g., stress response genes).

Interestingly, we also provide evidence for the potential of EloA to be involved in the production of rRNA. We show that EloA localizes to nucleoli and that it is present at the ribosomal RNA gene, *Rn45s* (Fig. 5). Physical interaction with rRNA processing factors provided additional support for involvement of EloA in rRNA transcription/processing (Fig. 4B). Our results also show that EloA is capable of forming phase-separated condensates *in vitro*, a feature shared among tested nucleolar proteins (41, 43). Lastly, we show that deletion of three highly disordered stretches within the IDR of EloA interferes with phase separation of EloA *in vitro* and its localization to the nucleolus (Fig. 6E). This is consistent with a study showing truncations within the IDR of EloA interfere with localization of EloA to sites of DNA damage (51). These data are consistent with an expanded range of activities for EloA in the nucleus that go beyond RNA Pol II transcription. In summary, we observe broad association of EloA with genes, and a more limited impact on steady-state transcription, especially transcription elongation, than expected based upon *in vitro* activity of this protein. The behavior we characterize is

consistent with a broad-based function that tunes nuclear processes at a large set of genes

Experimental procedures

Cell culture

Elongin A null (*EloA*^{-/-}) and wild-type mES cells were kind gift of Dr Teijiro Aso (Kochi Medical School) (46). Cells were grown as described elsewhere (22). In brief, gelatinized plates (0.2%) were seeded with mouse embryonic fibroblast (MEF) as feeder layer and cells were grown in ES cell media containing DMEM knockOut medium, 15% hyclone FBS, 1× GlutaMAX, 1× NEAA, 1× pen/strep, 2 × 10³ Units/ml of mLIF (Millipore), 55 μM 2BME, 1 μM and 3 μM of PD0325901 and CHIR99021, respectively. NIH/3T3 (ATCC CRL-1658) mouse embryonic fibroblast cells were grown in DMEM supplemented with 10% FCS.

Cut&Run-seq

Cut&Run-seq was carried out as described by the Henikoff lab with the following modifications: Mouse ES cells were cross-linked with 1% formaldehyde in PBS for 10 min and quenched with 0.125 M glycine for 5 min, cells were pelleted and washed once with PBS, spun down, resuspended in freezing media (DMEM, 10% DMSO), placed in a styrofoam container, and frozen O/N (this step was included to minimize cell membrane rupture and DNA breakage). On the day of the experiment, cross-linked cells were thawed and washed once with PBS and twice in cut&run wash buffer (20 mM HEPES pH7.5, 150 mM NaCl, 0.1% Tween20, 0.1% BSA, 1 Roche complete EDTA-free protease inhibitor tablet/50 ml), Tween20 was included in the wash buffer to minimize concanavalin A magnetic bead aggregation in downstream steps. Next, 5 × 10⁶ cells were resuspended in 950 μl of wash buffer. Fifty microliters of BioMag-Plus Concanavalin A beads (Polysciences Inc) was also washed gently twice with binding buffer (20 mM HEPES pH 7.9, 1 mM CaCl₂, 1 mM MnCl₂) and resuspended in binding buffer at the initial volume (50 μl). Concanavalin A beads (50 μl) were added to 5 × 10⁶ cells in 950 μl of wash buffer and incubated rotating for 15 min at RT. The cells were divided into five 200 μl aliquots for each antibody used (1 × 10⁶ cells/antibody) and after removal of wash buffer, 50 μl of ice cold antibody buffer (wash buffer containing 0.05% Digitonin and 2 mM EDTA) containing 0.5 μg of antibody was added to 1 × 10⁶ bead-bound cell pellet, and the mixture was gently resuspended by pipetting up and down 25 times. Resuspended beads were placed on a vortex at the lowest setting and incubated at room temperature for 15 min. For the EloA cut&run experiment using goat anti EloA (SCBT, R-19), a bridging rabbit anti-goat antibody was used for efficient binding to protein A. Cells were washed once after primary antibody incubation with 1 ml of Digitonin buffer (Wash buffer containing 0.05% Digitonin) resuspended in 50 μl of digitonin buffer with 0.5 μg of rabbit anti-goat IgG (rabbit anti-goat IgG [ab6697]). Left on the lowest setting of the vortex and incubated at RT for 15 min. Next, permeabilized, antibody-bound cells were washed once in digitonin buffer containing

EloA enrichment at active genes and enhancers

1 mM EDTA, and 50 μ l of digitonin buffer containing Protein A-MNase (pA-MN) at a final concentration of 700 ng/ml was added to the pellet, resuspended, and incubated at room temperature for 15 min. Cells were washed twice with digitonin buffer containing 1 mM EDTA. Next, 150 μ l of ice cold digitonin buffer was added to the pA-MN bound cells, resuspended, and digested on ice (0 °C) for 30 min after addition of CaCl₂ to a final concentration of 4 mM. Hundred microliter of 2 \times stop buffer was added as described in (24) and incubated on thermomixer for 15 min at 37 °C to release the cleaved DNA fragments into solution. The mixture was spun down in a cold tabletop centrifuge at max speed for 5 min, and 2 μ l of proteinase K was added to the supernatant eluate. The mixture was incubated at 65 °C for 4 h, and DNA was extracted by using ChIP DNA Clean & Concentrator (Zymo research). Purified DNA was used for library preparation as described elsewhere (52).

Data processing and bioinformatic analysis of cut-and-run-seq

All experiments were performed in duplicate. Sequencing reads were aligned to the mm10 genome using bowtie2 (53) and filtered using samtools (54) to keep uniquely aligned reads. Genome browser tracks were generated using Homer v4.10.3 (55) and visualized in IGV (56). Peaks were called in each replicate using SEACR (57) using IgG and “relaxed” parameters. Peaks in wild-type cells that were at least 1.5-fold over null cells were used further for analysis. Homer’s annotatePeaks function was used to find the distribution of peaks in the genome (Fig. 1B). The same annotations were used to calculate the proportions of features in the mouse genome for comparison. Deeptools v3.3.0 (58) was used to make average profile line plots. HOMER’s annotatePeaks function was used to calculate signals for boxplot and R was used to make boxplots. EloA-bound genes were defined as having at least one EloA peak from TSS-0.5 kb to TES. Unique transcripts for each gene were chosen based on the start of the Bru-seq signal. Intergenic peaks were defined as peaks not overlapping genes (TSS-0.5 kb-TES).

Nascent RNA sequencing (Bru-seq)

Bru-seq was used as a readout for the position and levels of Pol II on gene bodies to determine the elongation rate of Pol II. Moreover, eRNAs are highly unstable and are rapidly degraded upon transcription, which makes them hard to detect using conventional RNA sequencing methodologies. Bru-seq also allowed for better enrichment and detection of eRNA species that would have been hard to detect using conventional RNA sequencing. For nascent RNA analysis experiment, ES cells were depleted of MEF and plated on gelatin-coated 100 mm plates for 3–4 h. For the steady-state nascent RNA level study, 5-Bromouridine (BrU) (2 mM final conc) was added to cells for 10 min. For the elongation rate analysis using flavopiridol (FP), FP (Cayman Chemical, Inc) was added at a final concentration of 1 μ M for 60 min. Plates were washed with PBS and incubated in media containing 2 mM BrU for 5 or 10 min. Plates

were washed once with PBS and total RNA was isolated by addition of 4 ml of Trizol. Bru-seq was performed as described previously (22, 59) on 100 μ g of total RNA from specified mESC cells. Five-hundred microliters of mouse anti-BrdU antibody (200 μ g/ml, SCBT, sc-32323) was coupled to 1 ml of PBS-DEPC washed Dynabeads M-280 sheep anti-mouse IgG (11201D). Coupling was done in Bru-seq buffer (PBS-DEPC, 0.1% BSA, 0.05% Tween20), and yeast tRNA was added to a final concentration of 500 ng/ μ l. Coupling was carried out for 3 h at 4 °C. Coupled antibody was washed three times and stored in 1 ml of Bru-seq buffer. Prior to enrichment of nascent RNA, two consecutive rounds of mature mRNA depletion were carried out on 100 μ g of total RNA using Oligo d(T)₂₅ Magnetic beads (NEB Inc) according to manufacturer’s protocols. We also generated a pool of BrU-incorporated *in vitro* transcribed *Drosophila* spike in RNAs using 5-Bromouridine 5'-triphosphate (B7166-5MG) and MEGA-script T7 Kit (AM1333) according to manufacturer’s protocol. Names of *Drosophila* genes and the primer sequences used for synthesis of the templates are provided in Table S1. The *in vitro* transcribed spike in RNAs were pooled at varying concentrations and added to the total RNA from each sample. RNA immunoprecipitation and washing steps were carried out in the Bru-seq buffer as described above and previously (22). In brief, 75 μ l of antibody coupled beads was incubated with preheated (95 °C for 1 min followed by rapid cooling on ice) mature mRNA depleted total RNA for 90 min at RT in the dark. After the final washing step (four times total, 5–10 min each), the enriched RNA bound to beads was isolated using Trizol extraction. Ribosomal RNA was depleted using NEB-Next rRNA Depletion Kit (NEB), and first strand and directional second strand DNA synthesis was carried out using NEBNext Ultra II first, and directional second strand synthesis modules (NEB Inc), respectively. Library preparation was carried out as described previously (52).

Bru-seq data processing

All experiments were performed in duplicate. Paired end sequencing reads were aligned to the mm10 and dm3 genomes using bowtie2 and filtered using samtools to keep uniquely aligned reads. Reads from the positive and negative strand were separated using samtools. Browser tracks were generated using Homer v4.10.3 and visualized in IGV. Reads in whole gene of Refseq annotated genes were counted using featureCounts v1.6.1 (60). edgeR (61) and R v3.3.2 were used to normalize reads and calculate RPKMS and fold changes. All further calculations and figures were made using R and Deeptools.

ChIP-seq

ChIP was carried out on 2.5-5 \times 10⁶ cross-linked mESC cells as described previously (22).

ChIP-seq data processing

All experiments were performed in duplicate. Sequencing reads were aligned to the mm10 genome using bowtie2 and filtered

using samtools to keep uniquely aligned reads. Peaks were called in each replicate separately using HOMER's findPeaks function for broad peaks. Peaks that were 1.5FC over null in both replicates were used for the overlap analysis with cut-and-run peaks.

Turbidity assay

EloA and EloA complex that were purified and described previously (22) were used for the turbidity assay. Proteins were diluted to the described concentrations in the following buffer: 20 mM HEPES pH 7.9, 10% glycerol, 1.5 mM MgCl₂, 100 mM KCl and placed in clear-bottomed 384-well plates. Absorbance was measured at 405 nm using a Spectramax M3 plate reader. Values are average of at least three samples and errors denote standard deviation.

Centrifugation assay of mEGFP tagged proteins

mEGFP or mEGFP-ELOA proteins were purified as described above and brought to the specified concentration in 0.2 ml microcentrifuge tubes with the buffer used for turbidity assay. Samples were spun down at 2000g for 5 min and visualized under Safe Imager 2.0 Blue Light Transilluminator.

Fluorescence microscopy of protein condensates

Recombinant mEGFP and mEGFP-EloA were diluted to the specified concentration in 20 mM HEPES pH 7.9, 10% glycerol, 1.5 mM MgCl₂, 100 mM KCl, and 10% PEG-6000 (crowding agent) and placed on a glass slide and covered with a coverslip. Images were acquired using the 100× oil objective on Nikon 90i Eclipse epifluorescence microscope equipped with an Orca ER camera and velocity imaging software (Perkin Elmer).

Immunofluorescence

Cells were grown on coverslips and fixed with 1% formaldehyde for 10 min, washed with PBS, and permeabilized with 0.2% Triton X-100 in PBS for 15 min. Cells were blocked for 30–60 min in Blocking buffer (PBS, 0.05% Triton X-100, 3% BSA) and incubated with primary antibodies (goat anti EloA [1:300 SCBT, R-19], anti Nopp140(NOLC1) [1:250 SCBT, (E-7): sc-374033]) in blocking buffer O/N at 4 °C. Coverslips were washed three times with wash buffer (PBS, 1% BSA, 0.05% Triton X-100) and incubated with secondary antibody conjugated to Alexa Fluor 568 or 488 for 60 min at room temperature, coverslips were washed three times with wash buffer, once with PBS, rinsed with water, and mounted on slides using DAPI-containing mounting medium. Images were acquired using the 60× oil objective of Nikon 90i Eclipse microscope equipped with Orca ER camera (Hamamatsu) and velocity imaging software (Perkin Elmer). Z-stacks were collected using the 0.2 μm spacing and analyzed using maximum intensity of each section using ImageJ Fiji software.

Fluorescence recovery after photobleaching

Cells were grown on glass-bottomed fluorodish (WPI) in phenol-red free DMEM medium. Expression of mCherry-EloA

was induced using doxycycline at the indicated concentration (50–100 ng/ml for 6 h). All images were acquired using a Ti-2 Eclipse microscope (Nikon). Imaging was performed at 37 °C using a temperature-controlled Tokai-Hit culture dish system. Incubation with Hoechst 33,342 (Molecular Probes, 1:5000) for 2 h was carried out for live-cell DNA staining. A square of 1.5 × 1.5 microns was imaged for 2 s at 1 frame/s before stimulation for photobleaching using a 561-nm laser at 10% power for 2 s. Two distinct time-lapse imaging protocols were used: a fast acquisition at 600 ms/frame for 20 s or a slow acquisition at 2 s/frame phase for 1 min. Images were background subtracted and fluorescence intensity evaluated at bleached areas using Fiji software. Measurements were normalized to values obtained before bleaching (t = 0) and compared with an unbleached area used as a reference. All recovery graphs show relative fluorescence intensity (RFI) as a function of time (s). For nucleoplasm FRAP experiment, squares of 1.5 × 1.5 microns were stimulated and imaged as described above.

Generation of cell lines using lentiviral transduction

FLAG-EloA WT or FLAG-EloA skip IDR was cloned into a modified pTRIPZ vector (Dharmacon), described elsewhere (62). The skip IDR truncated construct contained three deletions in the IDR spanning amino acids (aa): 95–185, 236–335, 413–442, which was synthesized as a gBlocks Gene Fragments (IDT) and ligated to the sequence coding the C-terminal half of EloA. Constructs were transfected into HEK293T cells using TransIT lentiviral transfection reagent (Mirus) along with pCMV-dR8.91, which contains gag, pol, and rev genes and VSV-G envelope protein encoding pMD2.G plasmids. After 48 h, medium was collected and passed through 0.45 μm filter, and the medium was used to transduce 3T3 cells at low multiplicity of infection. After 2 days, transduced cells were selected using 1–2 μg/ml of puromycin. Expression of constructs at levels comparable with the endogenous levels of EloA was achieved through testing different concentration of doxycycline. Imaging was carried out on fixed cells as described in the immunofluorescence section above.

Baculoviral expression and purification of proteins

For characterization of phase separation properties of EloA, FLAG-tagged EloA and FLAG-EloA in complex with EloB and EloC constructs were cloned in pFastBac1 baculovirus expression plasmid and purified and described previously (22). The construct for expression of FLAG-mEGFP was also described previously (62). For FLAG-mEGFP-EloA construct, EloA cDNA sequence was cloned downstream of FLAG-mEGFP in pFastBac1. Baculovirus was generated and used to infect Sf9 cells as described in detail elsewhere (22).

Coimmunoprecipitation mass spectrometry (Co-IP/MS)

Wild-type and *EloA* null mESC cells growing on 100 mm gelatinized feeder plates were de-MEFed and 2 × 10⁸ cells were washed once in PBS, resuspended in lysis buffer (50 mM Tris

EloA enrichment at active genes and enhancers

pH 8.0, 0.25% NP-40, 50 mM NaCl, 1 mM EDTA, 5 mM MgCl₂, 0.2 mM PMSF, 0.2 mM DTT, Roche protease inhibitor tablet) and left on ice for 10 min, dounced ten times, and spun down at 2850g for 5 min to separate nuclei from cytoplasmic extract. Nuclear extract was prepared as described elsewhere (63), by resuspending the nuclear pellet in low salt buffer (20 mM HEPES, pH 7.9, 25% glycerol, 1.5 mM MgCl₂, 0.2 mM EDTA, 20 mM KCl, 0.2 mM DTT, 0.2 mM DTT, protease inhibitor) and while gently vortexing, equal volume of high salt buffer (20 mM HEPES, pH 7.9, 25% glycerol, 1.5 mM MgCl₂, 0.2 mM EDTA, 1.2 M KCl, 0.2 mM DTT, 0.2 mM DTT, protease inhibitor) was added to the resuspended nuclei. Nuclear extract was obtained by gentle continuous mixing in cold room for 30 min, followed by centrifugation at 20,000g. The concentration of the isolated nuclear extract was brought to 300 mM KCl using the low salt buffer and NP-40 was added to the extract to a final concentration of 0.05% before immunoprecipitation (IP). IP was carried out using 15 µg of goat anti EloA (SCBT, R-19) conjugated to magnetic protein G for 3 h at 4 °C. Beads were washed three times using wash buffer (20 mM HEPES, pH 7.9, 25% glycerol, 1.5 mM MgCl₂, 0.2 mM EDTA, 300 mM KCl, 0.2 mM DTT, 0.2 mM DTT, protease inhibitor). Following the final washing step, the beads were eluted using 1× SDS gel-loading buffer and separated on a 10% SDS-PAGE gel 1/4 of the way down and stained using coomassie blue. Each immunoprecipitation was excised as 5 gel sections, minced, and digested in-gel using trypsin. Digested gel sections were dehydrated by treatment with acetonitrile and speed-Vac. Rehydration was carried out in 50 mM ammonium bicarbonate supplemented with 12.5 ng/µl of sequencing-grade trypsin (Promega) at 4 °C. Samples were incubated at 37 °C overnight. For peptide extraction, ammonium bicarbonate was removed and washed with a 50% acetonitrile, 1% formic solution and dried in a speed-Vac. Dried samples were resuspended in HPLC solvent A (2.5% acetonitrile, 0.1% formic acid) and loaded into a reverse-phase HPLC column (containing 2.6 µm C18 spherical silica beads in fused silica capillary). Following formation of gradient, elution of peptides was carried out using increasing concentration of HPLC solvent B (97.5% acetonitrile, 0.1% formic acid). Electroionization was performed on the peptides before entering an LTQ Orbitrap Velos Pro ion trap mass spectrometer (Thermo Fisher Scientific). Detection, isolation, and fragmentation of the peptides were carried out to produce a tandem mass spectrum of fragment ions for each peptide. Sequest (Thermo Finnigan) was used for identification of peptide sequences. The false discovery rate (FDR) filter for peptide identification was set at 1% or less. Processing, preparation, and mass spectrometry analysis of the samples were carried out at Taplin Mass Spectrometry Facility (Harvard Medical School). Verification of EloA PAF1 interaction by immunoprecipitation Western Blot was carried out on 1 mg of nuclear extract using 1–3 µg of the following antibodies: Paf1 (Bethyl, A300-172A), CTR9 (Bethyl A301-395A), Leo1 (Bethyl, A300-175A), Parafibromin/CDC73, (Bethyl, A300-170A), EloA (R-19, Santa Cruz Biotechnology, sc-1557). Immunoblots were performed at a primary antibody concentration of 1:2000.

Proteomic data analysis presented in Figure 4A was performed using the CRAPome (crapome.org) analysis pipeline (64).

CRISPR/Cas9 generation of Elongin A knock out in NIH/3T3 cells

Guide RNA (gRNA) oligos targeting the 5' untranslated region and translation initiation site of murine *EloA* gene were identified using crispr.mit.edu guide RNA design tool and cloned into pSpCas9(BB)-2A-Puro (PX459)V2.0 (addgene, Plasmid #62988) (65). Primer sequences for cloning both constructs are provided in Table S2. The day before transfection, 1.5×10^5 3T3 cells were plated in each well of a 6-well plate. On the next day, 1.25 µg of each plasmid (2.5 µg total) was transfected using 5 µl of jetPRIME (Polyplus) transfection reagent. After 24 h, transfected cells were selected using puromycin at a final concentration of 2 µg/ml and 48 h later, cells were washed once with PBS, trypsinized, counted, and diluted down to 5 cells/ml (0.5 cells/100 µl), and 100 µl was plated in each well of a 96-well plate (three plates total). One week later, wells with cells growing in them were expanded in a 12-well plate and harvested after reaching confluency. Harvested cells were frozen down and also used for verification of *EloA* KO by western blot using EloA (R-19, Santa Cruz Biotechnology, sc-1557) antibody.

GREAT and wikipathways analysis

GREAT analysis (66) was carried out by uploading the differentially regulated enhancer BED files to the GREAT website (great.stanford.edu) and downloading the GO biological process list. Wikipathways (67) analysis was carried out by uploading the EloA enriched gene list to Enrichr (<http://amp.pharm.mssm.edu/Enrichr>) (68).

Published data sources and processing

Published data were downloaded from GEO as follows: H3K4me1 (GSM747542) (69), H3K4me3 (GSM1385046) (70), H3K27Ac (GSM2586543) (71), H3K36me3 (GSM1873392) (72), H3K79me2 (GSM307151) (73), Oct4 (GSM1082340), Sox2 (GSM1082341), Nanog (GSM1082342), Med1 (GSM1038259) (28), GRO-seq (GSE48895) (12).

FASTQ files were downloaded from GEO using sratools v2.9.1. All reads were processed as described before.

Data availability

All sequencing data have been deposited to GEO under accession number: GSE151582. Mass spectrometry raw data are accessible under ProteomeXchange: PXD022138 and jPOSTrepo: JPST000951.

Acknowledgments—We are grateful to Dr Teijiro Aso for providing *EloA*^{-/-} and *EloA*^{+/+} mES cells. We thank Brian Dynlacht and Yating Wang for discussion of their unpublished work. We also thank Jesse Cochrane and Mattia Lion for critical reading of the article, MGH NextGen Sequencing Core and Ross Tomaino at Taplin Mass Spectrometry Facility.

Author contributions—M. B. A., R. E. K. designed the project. M. B. A. performed all the experiments with the exception of FRAP analysis. M. D. carried out the bioinformatics analysis and generated plots. C. P. R. and M. D. B. performed and analyzed FRAP experiment. M. B. A. and R. E. K. wrote the manuscript.

Funding and additional information—This work was supported by National Institutes of Health grants R01GM043901, R37GM048405, and R35GM131743 to R. E. K. and grant GM122893 from National Institutes of Health/National Institute of General Medical Sciences to M. D. B. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: CTD, C-terminal domain; DFC, dense fibrillar component; EloA, Elongin A; eRNA, enhancer RNA; FC, fibrillar center; FDR, false discovery rate; GREAT, Genomic Regions Enrichment of Annotations Tool; IDR, intrinsically disordered regions; LLPS, liquid–liquid phase separation; mEGFP, monomeric enhanced GFP; NDR, nucleosome depleted region; Paf1C, polymerase-associated factor 1 complex; PONDR, predictor of natural disordered regions; TSS, transcription start site.

References

- Harlen, K. M., and Churchman, L. S. (2017) The code and beyond: transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nat. Rev. Mol. Cell Biol.* **18**, 263–273
- Fuda, N. J., Ardehali, M. B., and Lis, J. T. (2009) Defining mechanisms that regulate RNA polymerase II transcription *in vivo*. *Nature* **461**, 186–192
- Rougvie, A. E., and Lis, J. T. (1988) The RNA polymerase II molecule at the 5' end of the uninduced hsp70 gene of *D. melanogaster* is transcriptionally engaged. *Cell* **54**, 795–804
- Core, L. J., Waterfall, J. J., and Lis, J. T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* **322**, 1845–1848
- Core, L., and Adelman, K. (2019) Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev.* **33**, 960–982
- Adelman, K., and Lis, J. T. (2012) Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* **13**, 720–731
- Hou, L., Wang, Y., Liu, Y., Zhang, N., Shamovsky, I., Nudler, E., Tian, B., and Dynlacht, B. D. (2019) Paf1C regulates RNA polymerase II progression by modulating elongation rate. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 14583–14592
- Conaway, R. C., and Conaway, J. W. (2019) The hunt for RNA polymerase II elongation factors: a historical perspective. *Nat. Struct. Mol. Biol.* **26**, 771–776
- Baluapuri, A., Hofstetter, J., Dudvarski Stankovic, N., Endres, T., Bhandare, P., Vos, S. M., Adhikari, B., Schwarz, J. D., Narain, A., Vogt, M., Wang, S.-Y., Düster, R., Jung, L. A., Vanselow, J. T., Wiegering, A., *et al.* (2019) MYC recruits SPT5 to RNA polymerase II to promote processive transcription elongation. *Mol. Cell* **74**, 674–687.e11
- Herzel, L., Otz, D. S. M., Alpert, T., and Neugebauer, K. M. (2017) Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat. Rev. Mol. Cell Biol.* **18**, 637–650
- Maslon, M. M., Braunschweig, U., Aitken, S., Mann, A. R., Kilanowski, F., Hunter, C. J., Blencowe, B. J., Kornblihtt, A. R., Adams, I. R., and Cáceres, J. F. (2019) A slow transcription rate causes embryonic lethality and perturbs kinetic coupling of neuronal genes. *EMBO J.* **38**, e101244
- Jonkers, I., Kwak, H., and Lis, J. T. (2014) Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* **3**, e02407
- Bradsher, J. N., Jackson, K. W., Conaway, R. C., and Conaway, J. W. (1993) RNA polymerase II transcription factor SIII. I. Identification, purification, and properties. *J. Biol. Chem.* **268**, 25587–25593
- Bradsher, J. N., Tan, S., McLaury, H. J., Conaway, J. W., and Conaway, R. C. (1993) RNA polymerase II transcription factor SIII. II. Functional properties and role in RNA chain elongation. *J. Biol. Chem.* **268**, 25594–25603
- Aso, T., Lane, W. S., Conaway, J. W., and Conaway, R. C. (1995) Elongin (SIII): a multisubunit regulator of elongation by RNA polymerase II. *Science* **269**, 1439–1443
- Aso, T., Haque, D., Barstead, R. J., Conaway, R. C., and Conaway, J. W. (1996) The inducible elongin A elongation activation domain: structure, function and interaction with the elongin BC complex. *EMBO J.* **15**, 5557
- Gerber, M., Tenney, K., Conaway, J. W., Conaway, R. C., Eisenberg, J. C., and Shilatifard, A. (2005) Regulation of Heat shock gene expression by RNA polymerase II elongation factor, Elongin A. *J. Biol. Chem.* **280**, 4017–4020
- Chopra, V. S., Hong, J.-W., and Levine, M. (2009) Regulation of hox gene activity by transcriptional elongation in *Drosophila*. *Curr. Biol.* **19**, 688–693
- Ardehali, M. B., Yao, J., Adelman, K., Fuda, N. J., Petesch, S. J., Webb, W. W., and Lis, J. T. (2009) Spt6 enhances the elongation rate of RNA polymerase II *in vivo*. *EMBO J.* **28**, 1067–1077
- Kawauchi, J., Inoue, M., Fukuda, M., Uchida, Y., Yasukawa, T., Conaway, R. C., Conaway, J. W., Aso, T., and Kitajima, S. (2013) Transcriptional properties of mammalian elongin A and its role in stress response. *J. Biol. Chem.* **288**, 24302–24315
- Weems, J. C., Slaughter, B. D., Unruh, J. R., Hall, S. M., McLaird, M. B., Gilmore, J. M., Washburn, M. P., Florens, L., Yasukawa, T., Aso, T., Conaway, J. W., and Conaway, R. C. (2015) Assembly of the elongin A ubiquitin ligase is regulated by genotoxic and other stresses. *J. Biol. Chem.* **290**, 15030–15041
- Ardehali, M. B., Anselmo, A., Cochrane, J. C., Kundu, S., Sadreyev, R. I., and Kingston, R. E. (2017) Polycomb repressive complex 2 methylates Elongin A to regulate transcription. *Mol. Cell* **68**, 872–884.e6
- Skene, P. J., and Henikoff, S. (2017) An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *eLife* **6**, e21856
- Skene, P. J., Henikoff, J. G., and Henikoff, S. (2018) Targeted *in situ* genome-wide profiling with high efficiency for low cell numbers. *Nat. Protoc.* **13**, 1006–1019
- Min, I. M., Waterfall, J. J., Core, L. J., Munroe, R. J., Schimenti, J., and Lis, J. T. (2011) Regulating RNA polymerase pausing and transcription elongation in embryonic stem cells. *Genes Dev.* **25**, 742–754
- Furlong, E. E. M., and Levine, M. (2018) Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345
- Field, A., and Adelman, K. (2020) Evaluating enhancer function and transcription. *Annu. Rev. Biochem.* **89**, 213–234
- Whyte, W. A., Orlando, D. A., Hnisz, D., Abraham, B. J., Lin, C. Y., Kagey, M. H., Rahl, P. B., Lee, T. I., and Young, R. A. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319
- Henriques, T., Scruggs, B. S., Inoué, M. O., Muse, G. W., Williams, L. H., Burkholder, A. B., Lavender, C. A., Fargo, D. C., and Adelman, K. (2018) Widespread transcriptional pausing and elongation control at enhancers. *Genes Dev.* **32**, 26–41
- De Santa, F., Barozzi, I., Mietton, F., Ghisletti, S., Polletti, S., Tusi, B. K., Müller, H., Ragoussis, J., Wei, C.-L., and Natoli, G. (2010) A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol.* **8**, e1000384
- Kim, T.-K., Hemberg, M., Gray, J. M., Costa, A. M., Bear, D. M., Wu, J., Harmin, D. A., Laptewicz, M., Barbara-Haley, K., Kuersten, S., Markenscoff-Papadimitriou, E., Kuhl, D., Bito, H., Worley, P. F., Kreiman, G., *et al.* (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* **465**, 182–187
- Cinghu, S., Yang, P., Kosak, J. P., Conway, A. E., Kumar, D., Oldfield, A. J., Adelman, K., and Jothi, R. (2017) Intragenic enhancers attenuate host gene expression. *Mol. Cell* **68**, 104–117.e6

EloA enrichment at active genes and enhancers

33. Chao, S. H., and Price, D. H. (2001) Flavopiridol inactivates P-TEFb and blocks most RNA polymerase II transcription *in vivo*. *J. Biol. Chem.* **276**, 31793–31799
34. Tamura, K., Miyata, K., Sugahara, K., Onishi, S., Shuin, T., and Aso, T. (2003) Identification of EloA-BP1, a novel Elongin A binding protein with an exonuclease homology domain. *Biochem. Biophys. Res. Commun.* **309**, 189–195
35. Van Oss, S. B., Cucinotta, C. E., and Arndt, K. M. (2017) Emerging insights into the roles of the Paf1 complex in gene regulation. *Trends Biochem. Sci.* **42**, 788–798
36. Ding, L., Paszkowski-Rogacz, M., Nitzsche, A., Slabicki, M. M., Heninger, A.-K., de Vries, I., Kittler, R., Junqueira, M., Shevchenko, A., Schulz, H., Hubner, N., Doss, M. X., Sachinidis, A., Hescheler, J., Iacone, R., *et al.* (2009) A genome-scale RNAi screen for Oct4 modulators defines a role of the Paf1 complex for embryonic stem cell identity. *Cell Stem Cell* **4**, 403–415
37. Kuraoka, I., Ito, S., Wada, T., Hayashida, M., Lee, L., Saijo, M., Nakatsu, Y., Matsumoto, M., Matsunaga, T., Handa, H., Qin, J., Nakatani, Y., and Tanaka, K. (2008) Isolation of XAB2 complex involved in pre-mRNA splicing, transcription, and transcription-coupled repair. *J. Biol. Chem.* **283**, 940–950
38. Chen, H. K., Pai, C. Y., Huang, J. Y., and Yeh, N. H. (1999) Human Nopp140, which interacts with RNA polymerase I: implications for rRNA gene transcription and nucleolar structural organization. *Mol. Cell. Biol.* **19**, 8536–8546
39. Tsai, Y.-T., Lin, C.-I., Chen, H.-K., Lee, K.-M., Hsu, C.-Y., Yang, S.-J., and Yeh, N.-H. (2008) Chromatin tethering effects of hNopp140 are involved in the spatial organization of nucleolus and the rRNA gene transcription. *J. Biomed. Sci.* **15**, 471–486
40. Chen, D., and Huang, S. (2001) Nucleolar components involved in ribosome biogenesis cycle between the nucleolus and nucleoplasm in interphase cells. *J. Cell Biol.* **153**, 169–176
41. Feric, M., Vaidya, N., Harmon, T. S., Mitrea, D. M., Zhu, L., Richardson, T. M., Kriwacki, R. W., Pappu, R. V., and Brangwynne, C. P. (2016) Coexisting liquid phases underlie nucleolar subcompartments. *Cell* **165**, 1686–1697
42. Jacob, M. D., Audas, T. E., Uniacke, J., Trinkle-Mulcahy, L., and Lee, S. (2013) Environmental cues induce a long noncoding RNA-dependent remodeling of the nucleolus. *Mol. Biol. Cell* **24**, 2943–2953
43. Brangwynne, C. P., Mitchison, T. J., and Hyman, A. A. (2011) Active liquid-like behavior of nuclei determines their size and shape in *Xenopus laevis* oocytes. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4334–4339
44. Banani, S. F., Lee, H. O., Hyman, A. A., and Rosen, M. K. (2017) Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298
45. Yamazaki, K., Aso, T., Ohnishi, Y., Ohno, M., Tamura, K., Shuin, T., Kitajima, S., and Nakabeppu, Y. (2003) Mammalian Elongin A is not essential for cell viability but is required for proper cell cycle progression with limited alteration of gene expression. *J. Biol. Chem.* **278**, 13585–13589
46. Yasukawa, T., Bhatt, S., Takeuchi, T., Kawachi, J., Takahashi, H., Tsutsui, A., Muraoka, T., Inoue, M., Tsuda, M., Kitajima, S., Conaway, R. C., Conaway, J. W., Trainor, P. A., and Aso, T. (2012) Transcriptional elongation factor Elongin A regulates retinoic acid-induced gene expression during neuronal differentiation. *Cell Rep.* **2**, 1129–1136
47. Chen, F. X., Xie, P., Collings, C. K., Cao, K., Aoi, Y., Marshall, S. A., Rendleman, E. J., Ugarenko, M., Ozark, P. A., Zhang, A., Shiekhhattar, R., Smith, E. R., Zhang, M. Q., and Shilatifard, A. (2017) PAF1 regulation of promoter-proximal pause release via enhancer activation. *Science* **357**, 1294–1298
48. Lai, F., Gardini, A., Zhang, A., and Shiekhhattar, R. (2015) Integrator mediates the biogenesis of enhancer RNAs. *Nature* **525**, 399–403
49. Elrod, N. D., Henriques, T., Huang, K.-L., Tatomer, D. C., Wilusz, J. E., Wagner, E. J., and Adelman, K. (2019) The integrator complex attenuates promoter-proximal transcription at protein-coding genes. *Mol. Cell.* **76**, 738–752.e7
50. Kong, S. E., Shilatifard, A., Conaway, R. C., and Conaway, J. W. (2003) Preparation and assay of RNA polymerase II elongation factors elongin and ELL. *Methods Enzymol.* **371**, 276–283
51. Weems, J. C., Slaughter, B. D., Unruh, J. R., Boeing, S., Hall, S. M., McLaird, M. B., Yasukawa, T., Aso, T., Svejstrup, J. Q., Conaway, J. W., and Conaway, R. C. (2017) Cockayne syndrome B protein regulates recruitment of the Elongin A ubiquitin ligase to sites of DNA damage. *J. Biol. Chem.* **292**, 6431–6437
52. Bowman, S. K., Simon, M. D., Deaton, A. M., Tolstorukov, M., Borowsky, M. L., and Kingston, R. E. (2013) Multiplexed Illumina sequencing libraries from picogram quantities of DNA. *BMC Genomics* **14**, 1
53. Langmead, B., and Salzberg, S. L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359
54. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome project data processing subgroup (2009) The sequence alignment/map format and SAMtools. *Bioinform. Oxf. Engl.* **25**, 2078–2079
55. Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010) Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589
56. Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., and Mesirov, J. P. (2011) Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26
57. Meers, M. P., Tenenbaum, D., and Henikoff, S. (2019) Peak calling by sparse enrichment analysis for CUT&RUN chromatin profiling. *Epigenetics Chromatin* **12**, 42
58. Ramirez, F., Dundar, F., Diehl, S., Gruning, B. A., and Manke, T. (2014) deepTools: a flexible platform for exploring deep-sequencing data. *Nucleic Acids Res.* **42**, W187–W191
59. Tani, H., Mizutani, R., Salam, K. A., Tano, K., Ijiri, K., Wakamatsu, A., Isogai, T., Suzuki, Y., and Akimitsu, N. (2012) Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals. *Genome Res.* **22**, 947–956
60. Liao, Y., Smyth, G. K., and Shi, W. (2014) featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinform. Oxf. Engl.* **30**, 923–930
61. Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinform. Oxf. Engl.* **26**, 139–140
62. Plys, A. J., Davis, C. P., Kim, J., Rizki, G., Keenen, M. M., Marr, S. K., and Kingston, R. E. (2019) Phase separation of Polycomb-repressive complex 1 is governed by a charged disordered region of CBX2. *Genes Dev.* **33**, 799–813
63. Abmayr, S. M., Carrozza, M. J., and Workman, J. L. (2006) Preparation of nuclear and cytoplasmic extracts from mammalian cells. *Curr. Protoc. Pharmacol.* Chapter 12:Unit12.3
64. Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J.-P., St-Denis, N. A., Li, T., Miteva, Y. V., Hauri, S., Sardi, M. E., Low, T. Y., Halim, V. A., Bagshaw, R. D., Hubner, N. C., Al-Hakim, A., Bouchard, A., *et al.* (2013) The CRAPome: a contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* **10**, 730–736
65. Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308
66. McLean, C. Y., Bristor, D., Hiller, M., Clarke, S. L., Schaar, B. T., Lowe, C. B., Wenger, A. M., and Bejerano, G. (2010) GREAT improves functional interpretation of cis-regulatory regions. *Nat. Biotechnol.* **28**, 495–501
67. Slenter, D. N., Kutmon, M., Hanspers, K., Riutta, A., Windsor, J., Nunes, N., Méliou, J., Cirillo, E., Coort, S. L., Digles, D., Ehrhart, F., Giesbertz, P., Kalafati, M., Martens, M., Miller, R., *et al.* (2018) WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res.* **46**, D661–D667
68. Kuleshov, M. V., Jones, M. R., Rouillard, A. D., Fernandez, N. F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S. L., Jagodnik, K. M., Lachmann, A., McDermott, M. G., Monteiro, C. D., Gundersen, G. W., and Ma'ayan, A. (2016) Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97
69. Stadler, M. B., Murr, R., Burger, L., Ivanek, R., Lienert, F., Schöler, A., van Nimwegen, E., Wirbelauer, C., Oakeley, E. J., Gaidatzis, D., Tiwari, V. K.,

- and Schübeler, D. (2011) DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495
70. Noh, K.-M., Wang, H., Kim, H. R., Wenderski, W., Fang, F., Li, C. H., Dewell, S., Hughes, S. H., Melnick, A. M., Patel, D. J., Li, H., and Allis, C. D. (2015) Engineering of a histone-recognition domain in Dnmt3a alters the epigenetic landscape and phenotypic features of mouse ESCs. *Mol. Cell* **59**, 89–103
71. Dorigi, K. M., Swigut, T., Henriques, T., Bhanu, N. V., Scruggs, B. S., Nady, N., Still, C. D., Garcia, B. A., Adelman, K., and Wysocka, J. (2017) Mll3 and Mll4 facilitate enhancer RNA synthesis and transcription from promoters independently of H3K4 monomethylation. *Mol. Cell* **66**, 568–576.e4
72. Neri, F., Rapelli, S., Krepelova, A., Incarnato, D., Parlato, C., Basile, G., Maldotti, M., Anselmi, F., and Oliviero, S. (2017) Intragenic DNA methylation prevents spurious transcription initiation. *Nature* **543**, 72–77
73. Marson, A., Levine, S. S., Cole, M. F., Frampton, G. M., Brambrink, T., Johnstone, S., Guenther, M. G., Johnston, W. K., Wernig, M., Newman, J., Calabrese, J. M., Dennis, L. M., Volkert, T. L., Gupta, S., Love, J., *et al.* (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. *Cell* **134**, 521–533