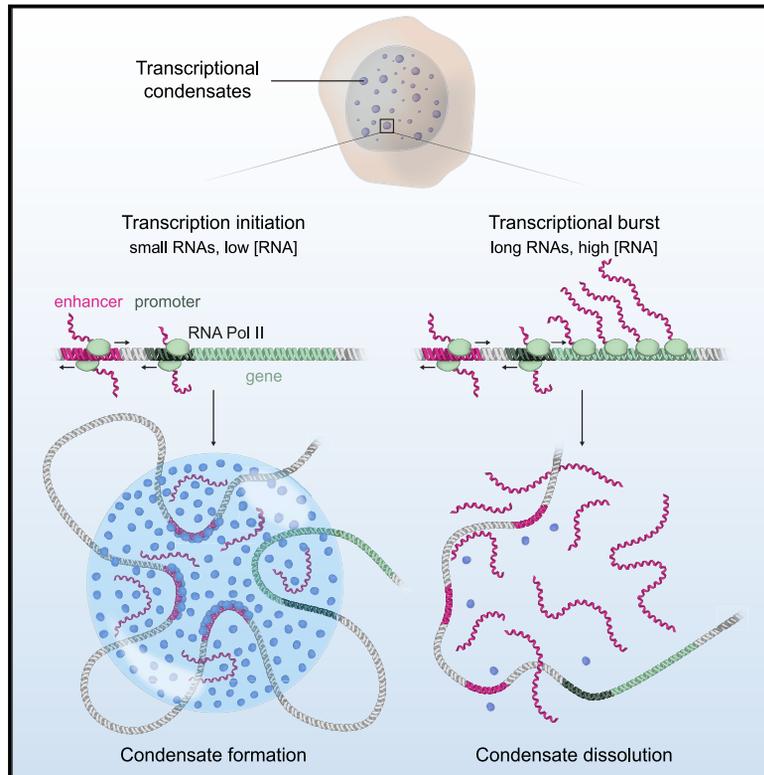


RNA-Mediated Feedback Control of Transcriptional Condensates

Graphical Abstract



Authors

Jonathan E. Henninger, Ozgur Oksuz, Krishna Shrinivas, ..., Ibrahim I. Cissé, Arup K. Chakraborty, Richard A. Young

Correspondence

arupc@mit.edu (A.K.C.),
young@wi.mit.edu (R.A.Y.)

In Brief

During the early steps of transcription initiation, nascent RNAs stimulate transcriptional condensate formation, whereas the burst of RNAs produced during elongation stimulates condensate dissolution.

Highlights

- Regulation of transcription by a non-equilibrium RNA feedback mechanism
- Low levels of RNA at regulatory elements promote condensate formation
- High levels of RNA from gene transcription can dissolve condensates
- Charge balance of electrostatic interactions can account for RNA feedback regulation

Article

RNA-Mediated Feedback Control of Transcriptional Condensates

Jonathan E. Henninger,^{1,13} Ozgur Oksuz,^{1,13} Krishna Shrinivas,^{2,6,8,13} Ido Sagi,¹ Gary LeRoy,^{9,10} Ming M. Zheng,⁴ J. Owen Andrews,⁴ Alicia V. Zamudio,^{1,3} Charalampos Lazaris,^{1,11} Nancy M. Hannett,¹ Tong Ihn Lee,¹ Phillip A. Sharp,^{3,7,14} Ibrahim I. Cissé,^{4,14} Arup K. Chakraborty,^{2,4,5,6,12,14,*} and Richard A. Young^{1,3,14,15,*}

¹Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA

²Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

³Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁴Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁵Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁶Institute of Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁷Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁸NSF-Simons Center for Mathematical & Statistical Analysis of Biology, Harvard University, Cambridge, MA 02138, USA

⁹Howard Hughes Medical Institute, New York University School of Medicine, New York, NY 10016, USA

¹⁰Department of Biochemistry and Molecular Pharmacology, New York University School of Medicine, New York, NY 10016, USA

¹¹Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

¹²Ragon Institute of Massachusetts General Hospital, MIT and Harvard University, Cambridge, MA 02139, USA

¹³These authors contributed equally

¹⁴Senior author

¹⁵Lead Contact

*Correspondence: arupc@mit.edu (A.K.C.), young@wi.mit.edu (R.A.Y.)

<https://doi.org/10.1016/j.cell.2020.11.030>

SUMMARY

Regulation of biological processes typically incorporates mechanisms that initiate and terminate the process and, where understood, these mechanisms often involve feedback control. Regulation of transcription is a fundamental cellular process where the mechanisms involved in initiation have been studied extensively, but those involved in arresting the process are poorly understood. Modeling of the potential roles of RNA in transcriptional control suggested a non-equilibrium feedback control mechanism where low levels of RNA promote condensates formed by electrostatic interactions whereas relatively high levels promote dissolution of these condensates. Evidence from *in vitro* and *in vivo* experiments support a model where RNAs produced during early steps in transcription initiation stimulate condensate formation, whereas the burst of RNAs produced during elongation stimulate condensate dissolution. We propose that transcriptional regulation incorporates a feedback mechanism whereby transcribed RNAs initially stimulate but then ultimately arrest the process.

INTRODUCTION

Diverse biological processes have evolved feedback mechanisms to enable positive and negative regulation. Examples of biological processes that are known to incorporate feedback regulation include signal transduction (Brandman and Meyer, 2008), production of RNA splicing factors (Jangi and Sharp, 2014), circadian rhythms (Dunlap, 1999), red blood cell production (Ebert and Bunn, 1999), and response to DNA damage (Lahav et al., 2004). In transcription, some factors that regulate amino acid biosynthetic pathway genes can be regulated allosterically by intermediates produced by those pathways (Bergot et al., 1992; Bruhat et al., 2000; Sellick and Reece, 2003), but a general feedback mechanism has not been described. Evidence

that feedback control is often mediated by the product of the process (Brandman and Meyer, 2008; Elowitz and Leibler, 2000; Gardner et al., 2000; Monod and Jacob, 1961; Umbarger, 1956) is one of the factors that led us to postulate that RNA may regulate transcription by a feedback mechanism.

Mammalian transcription produces diverse RNA species from regulatory elements and genes (Smith et al., 2019), and transcription of genes occurs in bursts of RNA synthesis (Chubb et al., 2006; Raj and van Oudenaarden, 2008; Raj et al., 2006). Transcription factors and coactivators recruit RNA polymerase II (RNA Pol II) to enhancer and promoter elements, where short (20–400 bp) RNAs are bidirectionally transcribed before RNA Pol II pauses (Adelman and Lis, 2012; Core and Adelman, 2019; Jin et al., 2017; Kim et al., 2010; Seila et al., 2008). These

RNA species are short-lived and are reported to have various regulatory roles, although there is not yet a consensus regarding their functions (Andersson et al., 2014; Catarino and Stark, 2018; Core et al., 2014; Gardini and Shiekhhattar, 2015; Henriques et al., 2018; Lai et al., 2013; Li et al., 2016; Mikhaylichenko et al., 2018; Nair et al., 2019; Pefanis et al., 2015; Rahnamoun et al., 2018; Schaukowitch et al., 2014; Scruggs et al., 2015; Sigova et al., 2015; Smith et al., 2019; Struhl, 2007). RNA Pol II pause release leads to processive elongation, which occurs in periodic bursts (~1–10 min in duration), where multiple molecules of RNA Pol II can be released from promoters within a short time frame and produce multiple molecules of mRNA (~1–100 molecules per burst) (Cisse et al., 2013; Fukaya et al., 2016; Larsson et al., 2019). How and whether the diverse RNA species produced during transcription—which differ in length, half-life, and number—affect or regulate transcription is currently unclear.

Recent studies have shown that transcriptional condensates can compartmentalize and concentrate large numbers of transcription factors, cofactors, and RNA Pol II at super-enhancers, clusters of enhancers that regulate genes with prominent roles in cell identity (Boija et al., 2018; Cho et al., 2018; Cramer, 2019; Hnisz et al., 2017; Sabari et al., 2018). The component enhancer elements of such genes promote transcriptional condensate formation by crowding transcription factors and Mediator at densities above sharply defined thresholds for condensate formation (Shrinivas et al., 2019). Transcriptional condensates are highly dynamic and can be observed in live cells to form and dissolve at timescales ranging from seconds to minutes (Cho et al., 2018). The periodic formation and dissolution of dynamic transcriptional condensates, coupled with evidence that different species and levels of RNAs are produced at different stages of transcription, led us to wonder whether transcriptional condensates are regulated by a non-equilibrium feedback mechanism mediated by its RNA product.

RNA molecules are components of and play regulatory roles in diverse biomolecular condensates. These include the nucleolus, nuclear speckles, paraspeckles, and stress granules (Fay and Anderson, 2018; Roden and Gladfelder, 2020; Sabari et al., 2020; Strom and Brangwynne, 2019). RNA has a high negative charge density because of its phosphate backbone, and the effective charge of a given RNA molecule is directly proportional to its length (Boeynaems et al., 2019). Condensates are thought to be formed by an ensemble of low-affinity molecular interactions, including electrostatic interactions, and RNA can be a powerful regulator of condensates that are formed and maintained by electrostatic forces (Banani et al., 2017; Maharana et al., 2018; Peran and Mittag, 2020; Shin and Brangwynne, 2017). Indeed, RNA has been shown to enter and modify the properties of simple condensates formed by polyelectrolyte-rich molecules (Drobot et al., 2018; Frankel et al., 2016; Mountain and Keating, 2020). In a phenomenon called complex coacervation, a type of liquid-liquid phase separation mediated by electrostatic interactions between oppositely charged polyelectrolytes, low levels of RNA can enhance condensate formation, whereas high levels can cause their dissolution (Lin et al., 2019; Overbeek and Voorn, 1957; Sing, 2017; Srivastava and Tirrell, 2016). Condensate formation and subsequent dissolution with increasing RNA concentration is an example of reentrant

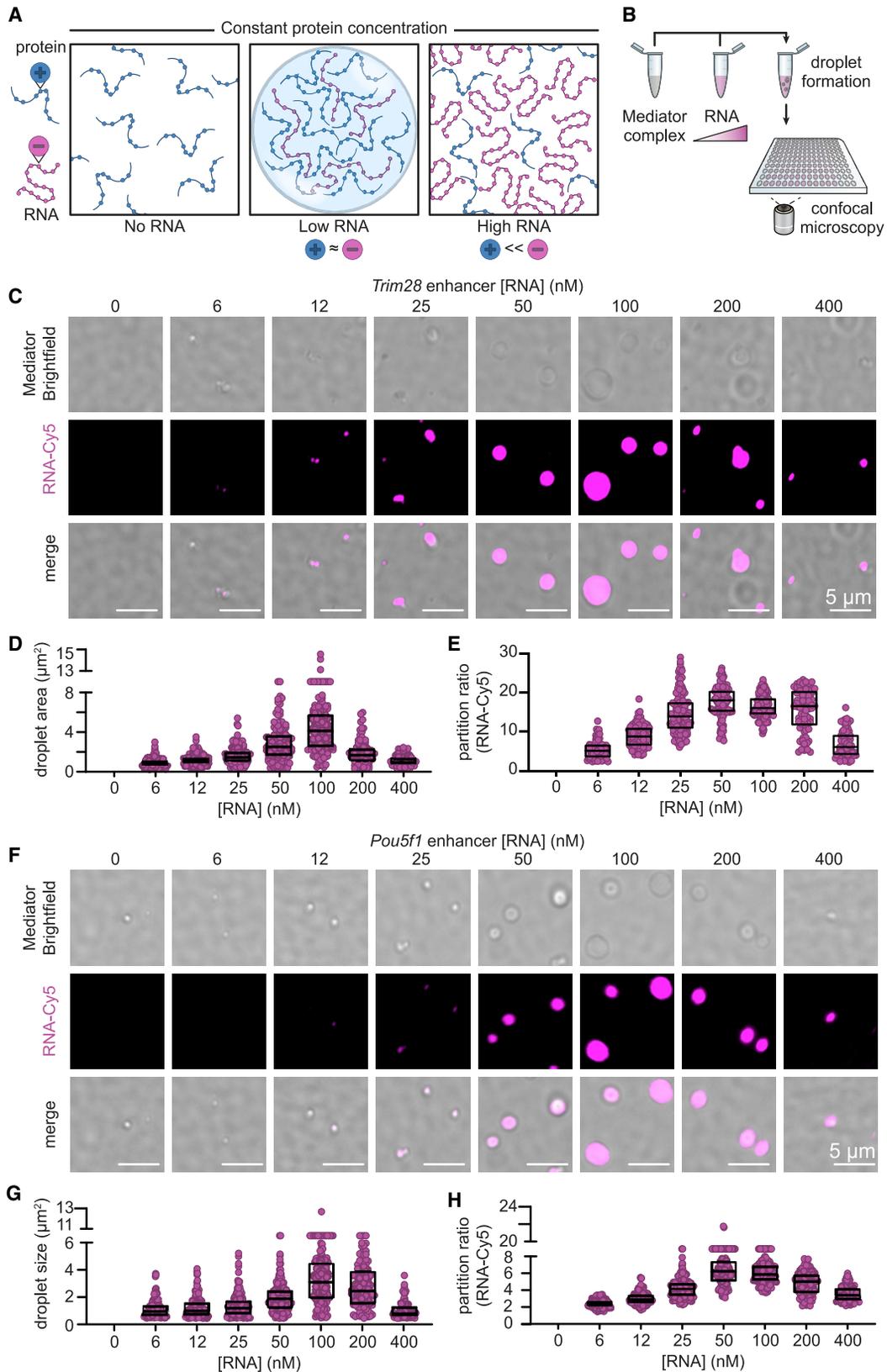
phase behavior, which is driven by favorable opposite-charge interactions at low RNA concentrations (formation) and repulsive like-charge interactions at high RNA concentrations (dissolution) (Banerjee et al., 2017; Milin and Deniz, 2018). We wondered whether such a reentrant equilibrium phase behavior coupled to the non-equilibrium processes that occur during transcription could regulate transcriptional output.

By combining physics-based modeling and experimental analysis, we propose and test a model where the products of transcription initiation stimulate condensate formation and those of a burst of elongation stimulate condensate dissolution. We provide experimental evidence that physiological RNA levels can enhance or dissolve transcriptional condensates. These results show a mechanism by which the products of transcription regulate condensate behaviors and, thus, transcription and suggest that this non-equilibrium process provides negative feedback to dissolve the transcriptional condensates that support initiation and thereby arrest transcription.

RESULTS

Low Levels of RNA Enhance and High Levels Dissolve Mediator Condensates

To explore the potential role of RNA in regulating transcriptional condensates, we sought to estimate the number and effective charge of RNA and protein molecules in a typical transcriptional condensate at different stages of transcription. In early stages of transcription, low levels of small noncoding RNAs are produced by RNA Pol II at enhancers and promoter-proximal regions (Figure S1A; Adelman and Lis, 2012; Core and Adelman, 2019; Kim et al., 2010; Seila et al., 2008). During pause release, RNA Pol II produces longer genic RNAs during bursts of transcription elongation (Figure S1A; Adelman and Lis, 2012; Core and Adelman, 2019). These protein- and RNA-rich states can be thought of as mixtures of poly-electrolytes that may undergo complex coacervation (Figure 1A; Lin et al., 2019; Overbeek and Voorn, 1957; Sing, 2017; Srivastava and Tirrell, 2016). We reasoned that this phenomenon is likely to be relevant to transcriptional condensates because electrostatic interactions contribute to formation of these condensates, even in the absence of RNA (Boija et al., 2018; Sabari et al., 2020). Complex coacervate formation through phase separation is promoted when polyelectrolytes are present at concentrations where their net charges are approximately balanced. When the concentration of a poly-electrolyte, such as RNA, becomes sufficiently high, domination of repulsive like-charge interactions can suppress phase separation (Banerjee et al., 2017; Lin et al., 2019; Milin and Deniz, 2018; Muthukumar, 2016; Overbeek and Voorn, 1957; Zhang et al., 2018). Thus, at constant protein concentration, titrating RNA levels results in reentrant phase behavior, whereby low RNA levels promote and high RNA levels suppress condensate formation (Figure 1A; Banerjee et al., 2017; Milin and Deniz, 2018; Zhang et al., 2018). We wondered whether the reentrant phase behavior might apply to regulation of transcriptional condensates during transcription. Because the quantities of the diverse RNA species and proteins present in transcriptional condensates in populations of cells can be estimated (Figure S1; STAR Methods), it is possible to conduct experimental tests to



(legend on next page)

determine whether reentrant phase behavior occurs under physiologically relevant conditions of these molecules.

As an initial test of whether low levels of RNA stimulate transcriptional condensate formation while high levels of RNA favor condensate dissolution, we used an *in vitro* droplet assay (Figure 1B). Using components at physiologically relevant conditions, we investigated whether an enhancer RNA transcribed from the *Trim28* super-enhancer, which has been shown previously to form a transcriptional condensate in living cells (Boija et al., 2018; Guo et al., 2019), influences condensate formation by purified Mediator complex. Measurement of enhancer RNA levels in cells indicated that ~ 0.2 molecules of this enhancer RNA exist at steady state in murine embryonic stem cells (mESCs) (Figure S1F). Given that multiple loci in a super-enhancer are transcribed into enhancer RNAs, this roughly corresponds to ~ 100 – $1,000$ nM of RNA in a typical Mediator condensate in cells (STAR Methods). These condensates typically contain Mediator at a concentration of around 1 – 20 μ M (STAR Methods). The results showed that addition of 6 – 400 nM *Trim28* enhancer RNA to 200 nM purified Mediator complex had a dose-dependent effect on the size of Mediator/RNA droplets (Figures 1C–E). Droplet sizes peaked at 100 nM RNA (Figure 1D), and the relative enrichment of RNA in the droplets, as measured by the ratio of average intensity inside versus outside the droplet (partition ratio), followed a similar trend (Figure 1E). Similar results were obtained using an enhancer RNA transcribed from the *Pou5f1* super-enhancer (Figures 1F–H). Thus, within the range of physiological levels observed in cells, low levels of RNA can enhance condensate formation, and high levels of RNA can reduce condensate formation by Mediator *in vitro*.

Charge Balance Mediates Regulation of MED1-IDR Condensates by RNA

We next sought to determine whether the reentrant phase behavior of mixtures of RNA and transcriptional proteins is predominantly regulated by charge balance considerations, with other types of RNA-protein interactions playing a less significant role. We performed *in vitro* droplet assays (Figure 2A) using the MED1 C-terminal intrinsically disordered region (MED1-IDR), which has proven to be a useful surrogate for the multisubunit Mediator complex because it is not possible to purify sufficient

amounts of this complex to test all parameters of interest (Boija et al., 2018; Guo et al., 2019; Klein et al., 2020; Li et al., 2020; Sabari et al., 2018; Shrinivas et al., 2019; Zamudio et al., 2019). Fusion of GFP to MED1-IDR allows quantification by fluorescence of a single species whose effective charge can be calculated to determine the charge ratio between protein and RNA. Addition of increasing levels of RNA to a constant protein concentration should have predictable effects on partitioning of either component according to its charge ratio (Figure 2B). Non-coding and coding RNAs produced from three different super-enhancer loci and their associated genes (*Trim28*, *Pou5f1*, and *Nanog*; Figure S1) were selected for this analysis based on prior studies of nascent RNA sequencing data in mESCs (Boija et al., 2018; Guo et al., 2019; Sabari et al., 2018; Sigova et al., 2015; Whyte et al., 2013). Addition of 6 – 400 nM of each of these RNAs to $1,000$ nM MED1-IDR (protein:RNA ratios = $167:2.5$) stimulated formation of MED1-IDR condensates at low RNA concentrations and dissolved MED1-IDR condensates at higher RNA concentrations (Figures 2C, 2D, S2A, and S2B). BRD4 is another key component of transcriptional condensates, and BRD4-IDR protein exhibits condensate behaviors very similar to those of MED1-IDR (Sabari et al., 2018); the effects of increasing RNA levels on formation and dissolution of BRD4-IDR condensates were very similar to those observed for MED1-IDR (Figures S2C and S2D). RNA did not stimulate formation of droplets with GFP alone or OCT4-GFP, both of which have a net negative charge (Figure S2E). Condensates exhibited internal dynamic reorganization (Figure S2F) with apparent diffusion coefficients 3 – $5 \pm 0.8 \times 10^{-2} \mu\text{m}^2/\text{s}$ (STAR Methods), consistent with liquid-like behavior (Nott et al., 2015; Sabari et al., 2018; Taylor et al., 2019). The incomplete recovery after photobleaching seen here has been observed previously with other condensates (Nott et al., 2015; Sabari et al., 2018; Taylor et al., 2019) and could result from a small portion of the droplet material being relatively immobile, rate-limiting material in the dilute phase or droplet aging (Taylor et al., 2019). These results show that diverse RNAs are capable of stimulating MED1-IDR condensate formation when present at relatively low levels and dissolving MED1-IDR condensates at high levels.

We sought to further test whether charge balance is the predominant phenomenon underlying the RNA-mediated effects on MED1-IDR condensates (STAR Methods). If so, then MED1-

Figure 1. Low Levels of RNA Enhance and High Levels Dissolve Mediator Condensates

(A) Diagram of reentrant phase transition in response to increasing concentrations of RNA over constant protein concentration. The condensed fraction of protein peaks at the RNA concentration at which the charges between protein and RNA are balanced, whereas alteration of this charge balance in either direction decreases the condensed fraction.

(B) Experimental design for the *in vitro* droplet formation assay. Whole Mediator complex is mixed with increasing concentrations of RNA under physiologically relevant buffer conditions, and droplets are imaged using confocal microscopy.

(C) Representative images of droplets formed by the unlabeled whole Mediator complex (200 nM) and Cy5-labeled *Pou5f1* enhancer RNA at increasing concentrations (0 – 400 nM). Bright-field images of the Mediator complex were divided by a median-filtered image (pixels = 15) here and in the subsequent panels (scale bars = $5 \mu\text{m}$).

(D) Droplet sizes in (C).

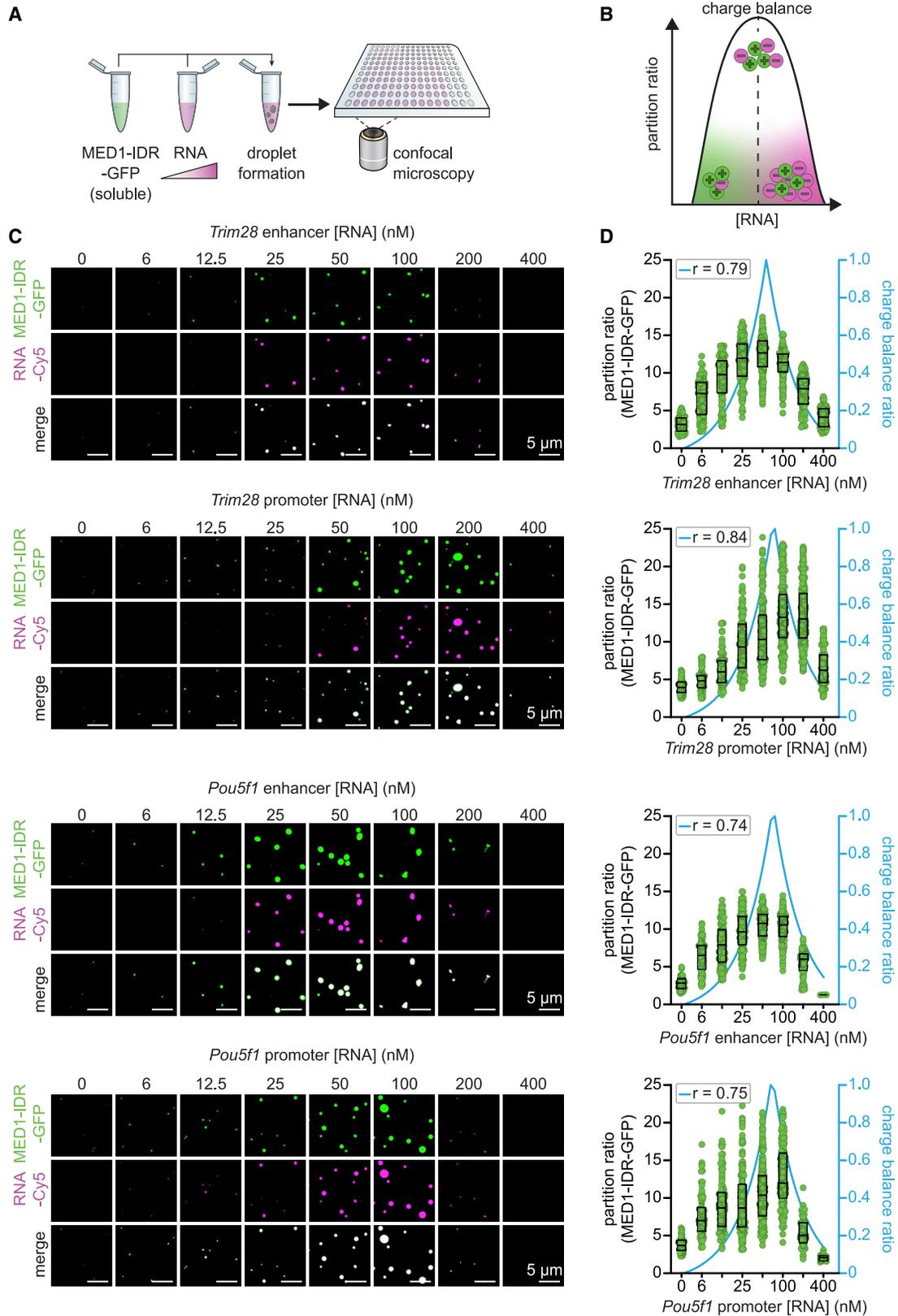
(E) Partition ratios of Cy5-labeled RNA within the droplets in (C).

(F) Representative images of droplets formed by the unlabeled whole Mediator complex (200 nM) and Cy5-labeled *Trim28* enhancer RNA at increasing concentrations (0 – 400 nM) (scale bars = $5 \mu\text{m}$).

(G) Droplet sizes in (F).

(H) Partition ratios of Cy5-labeled RNA within the droplets in (F).

See also Figure S1.



(legend on next page)

IDR/RNA condensate formation should be enhanced when the protein and RNA polymers are balanced in charge, and they should be sensitive to disruption of this balance. We quantified the relative charge of RNA and MED1-IDR and computed the correlation with the partition ratio of MED1-IDR (STAR Methods). RNA-mediated effects on MED1-IDR condensates correlated with their charge balance, as observed by the concordance of higher partition ratios near charge balance and lower partition ratios away from this balance (Figure 2D). We would expect an RNA length-dependent shift in the RNA level required for peak MED1-IDR partitioning when RNAs of different length are introduced into the droplet assay in equal numbers. This expectation, that a higher concentration of shorter RNAs is needed to disrupt condensate formation, was observed (Figures S3A and S3B). Another prediction from charge balance considerations is that these interactions should be largely independent of RNA sequence, so antisense versions of any one of the RNA species should exhibit the same quantitative effects as the sense strand, and this was also observed (Figures S3B and S3C). Consistent with charge balance considerations, MED1-IDR condensates formed with RNA were sensitive to increasing monovalent salt, which screens charged interactions (Figure S3D). The expectations from charge balance considerations also held when MED1-IDR and RNA concentrations were varied (Figures S4A–S4D) and when alternative polyanions (heparin and single-stranded DNA [ssDNA]) were employed (Figures S4E and S4F). RNA did not stimulate condensate formation by a MED1-IDR mutant lacking positively charged residues (MED1-IDR R/H/K > A) (Figures S4G and S4H). Although the experiments described above show a strong correlation between charge balance and partition ratios, the lack of complete correlation suggests that other features of RNA and MED1-IDR, such as RNA secondary structure (Roden and Gladfelter, 2020) or non-electrostatic interactions (Sabari et al., 2018), may influence the observed phase behavior. Nonetheless, these results further support the concept that RNA-mediated effects on equilibrium behavior of MED1-IDR condensates are predominantly regulated by electrostatic effects.

RNA-Mediated Effects on Condensates in Reconstituted *In Vitro* Transcription Assays

We sought to investigate the functional consequence of the RNA-mediated reentrant phase behavior on transcription. RNA Pol II-dependent transcription can be reconstituted *in vitro* with purified components (Roeder, 2019), so we investigated whether droplets containing transcriptional components are formed in these assays and whether conditions that alter droplet levels similarly alter transcriptional output. We used a classic reconstituted mammalian transcription sys-

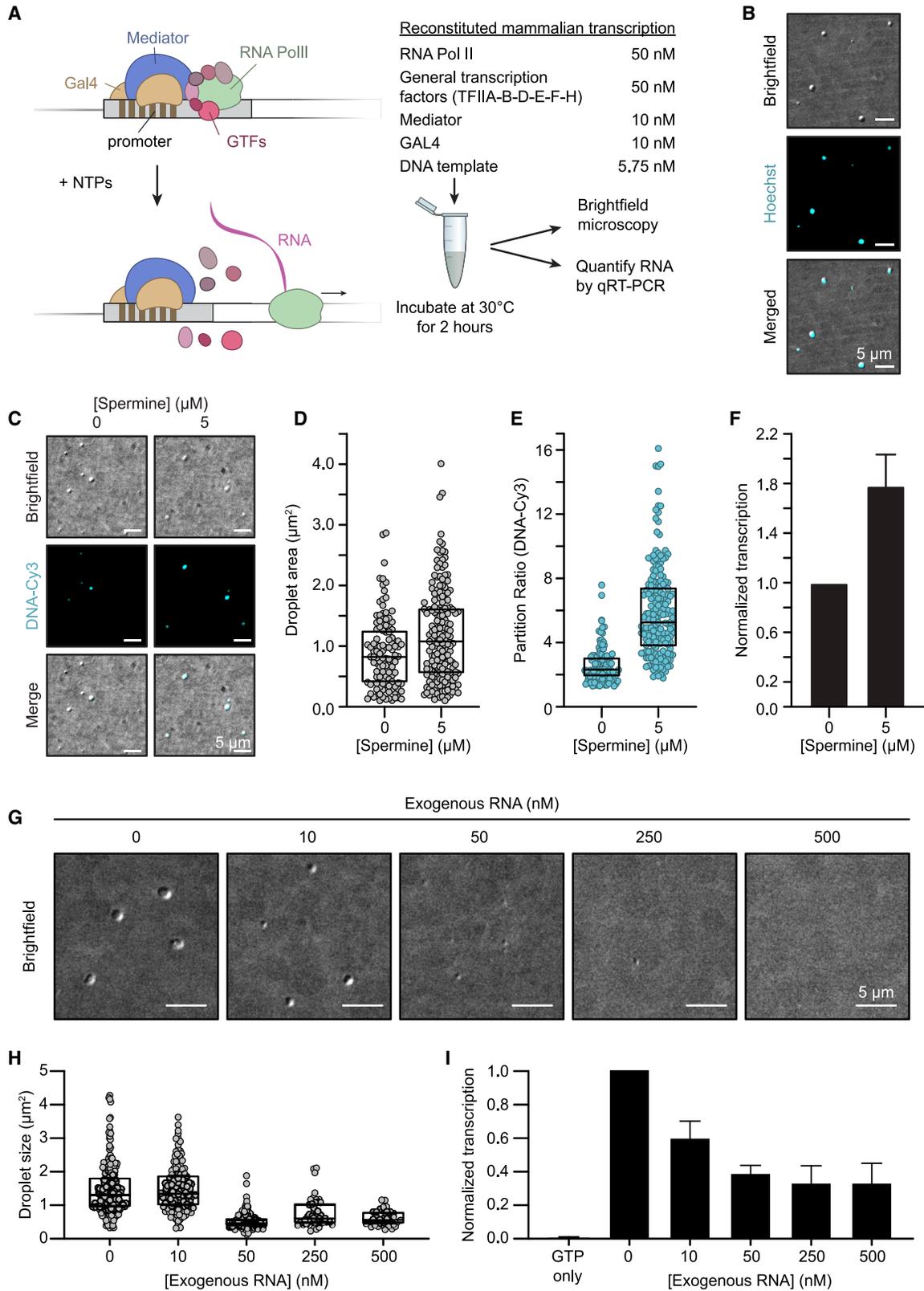
tem with purified components, including RNA Pol II, general transcription factors, Mediator, and a transcriptional activator (Gal4), where addition of nucleotides permits transcription of a linear DNA template (Figure 3A). We observed that component mixtures and buffer conditions that are optimal for transcriptional output (Carey et al., 2009; Flores et al., 1992; LeRoy et al., 2008; Orphanides et al., 1998) produced droplets containing the DNA template (Figure 3B). Quantification of the newly synthesized RNA in this system showed that 3.5 (± 0.5) pM RNA was produced in the transcription reaction (STAR Methods). We were unable to demonstrate that RNA synthesis actually occurs in the droplets because we cannot eliminate the possibility that synthesis occurs in the bulk phase and the product subsequently partitions into the droplet, but the observation that protein and template DNA concentrate in droplets under conditions optimal for transcription (Figure 3B) and evidence that diverse condensate-altering treatments have similar effects on transcription, described below, are consistent with the notion that transcription occurs within condensates in this reconstituted system.

We reasoned that if transcription and droplet formation are mutually dependent in the reconstituted system, then treatments that alter transcription should similarly affect condensate formation and vice versa. Addition to the reaction of various chemicals that are known to inhibit transcription (elevated concentrations of nucleoside triphosphates (NTPs), NaCl, or heparin; Carey et al., 2009; Reinberg and Roeder, 1987) caused reductions in droplet area, DNA partitioning, and transcription (Figure S5). Spermine, a positively charged polyamine, enhances droplet formation when it contributes to charge balance in coacervate models (Aumiller et al., 2016). Addition of spermine at concentrations predicted to balance charge in the *in vitro* reactions simultaneously increased droplet area, partitioning of template DNA, and levels of RNA synthesis (Figures 3C–3F; Table S2; Blair, 1985; Moruzzi et al., 1975). These correlations suggest that optimal droplet formation and transcription are co-dependent.

An expectation of the RNA feedback model is that droplets in the reconstituted system might ultimately produce enough RNA to cause a reduction in droplet size and transcriptional output. However, the low concentrations of RNA produced in these systems (3.5 ± 0.5 pM; STAR Methods) are insufficient to dissolve the droplets. For this reason, we tested whether purified RNA, added to the reaction, would similarly affect droplets and transcription. Indeed, addition of exogenous RNA reduced the number and size of the droplets (Figures 3G and 3H) and reduced template-derived RNA synthesis, as measured by qRT-PCR (Figure 3I). Although these results do not rule out additional ways in which RNA may

Figure 2. Charge Balance Mediates Regulation of MED1-IDR Condensates by RNA

(A) Experimental design for the *in vitro* droplet formation assay. Soluble MED1-IDR-GFP is mixed with increasing concentrations of RNA under physiologically relevant buffer conditions, and droplets are imaged using confocal microscopy. (B) Scheme of the charge balance ratio between constant protein concentration and increasing RNA concentrations. (C) Representative images of droplets formed by increasing concentrations (0–400 nM) of the indicated RNAs mixed with 1 μ M of MED1-IDR-GFP (scale bars = 5 μ m). (D) Partition ratios of MED1-IDR-GFP within the droplets in (C) (left y axis). Charge balance ratios between MED1-IDR-GFP and increasing concentrations of the indicated RNAs are shown as blue lines (right y axis). Correlation between partition ratio and charge balance is determined by Pearson correlation (r). See also Figures S2–S4.



(legend on next page)

affect transcription (Pai et al., 2014), they are consistent with the expected behavior of transcriptional condensates if RNA contributes to negative feedback control.

A Model of RNA-Mediated Non-equilibrium Feedback Control of Transcriptional Condensates

The *in vitro* experiments, which provide evidence that key transcriptional proteins and RNA exhibit electrostatics-driven, RNA-protein ratio-dependent reentrant phase transition, were performed under equilibrium conditions (Figures 1 and 2). However, *in vivo*, RNA is synthesized and degraded at specific genomic loci by dynamic, ATP-dependent, non-equilibrium processes (Azofeifa et al., 2018; Li et al., 2016; Pefanis et al., 2015). To investigate how non-equilibrium processes underlying transcription may regulate transcriptional condensates, we built a physics-based model. The model consists of two inter-linked parts: (1) a free-energy function (Figure 4A) that depends on the concentrations of transcriptional proteins and RNA and recapitulates the equilibrium reentrant phase behavior of RNA-protein mixtures (Figures 1 and 2), and (2) a mathematical framework to study spatiotemporal evolution of condensates subject to non-equilibrium dynamical processes of RNA synthesis, degradation, and diffusion (Figure 4B).

The goal of the model described below is to explore non-equilibrium regulation of transcription by RNA output and obtain insights into the pertinent mechanistic principles. Quantitative descriptions of RNA-protein phase behavior even *in vitro* (Adhikari et al., 2018; Delaney and Fredrickson, 2017) and direct measurements of the dynamic parameters underlying transcription are largely unavailable (Rodriguez and Larson, 2020). Therefore, we sought to develop a phenomenological model of non-equilibrium regulation of transcription by RNA output and use it to predict the qualitative effects of perturbing model parameters. Our experimental approaches allow such perturbations to be realized and can test whether the predicted effects are accurate. By coupling the predictions of effects of perturbing model parameters with experimental tests, we aimed to obtain mechanistic insights into how RNA synthesis may dynamically regulate transcription itself.

We first developed a free-energy function to recapitulate the experimentally observed reentrant phase behavior of RNA-pro-

tein mixtures (Figure 4A). The free-energy function depends on the concentrations of transcriptional proteins and RNA ($\phi_p(\vec{r}, t)$) and ($\phi_r(\vec{r}, t)$), which vary in space and time. For simplicity, all transcriptional proteins are combined into one pseudo-species. As noted above, our goal is not quantitative recapitulation of known experimental data but to obtain mechanistic insights into RNA-mediated non-equilibrium regulation of transcription that could be tested experimentally. Therefore, we first sought to develop a free-energy function that qualitatively recapitulates the observed reentrant phase behavior of RNA/protein mixtures. Following a long tradition in the physics of phase transitions, we employed a general Landau approach (Kardar, 2007; Landau, 1937) and expanded the free energy as a function of RNA and protein concentrations. We include terms to describe repulsive RNA-RNA interactions, favorable interactions among the transcriptional proteins that drive condensate formation of transcriptional proteins in the absence of RNA (Figure 4C, equation 1, green) as well as a surface tension term important for describing condensate formation (Figure 4C, equation 1, blue; STAR Methods). The free-energy function also includes protein-RNA interactions that are described by a concentration-dependent interaction term, which is expanded in the standard Landau fashion (Figure 4C, equation 1, red; Kardar, 2007). Magnitudes of the coefficients of the various terms in the expansion ($\chi_{a,b,c}$) account for the effective strength of RNA-protein interactions (STAR Methods), which implicitly include solvent effects. Although symmetry arguments do not preclude any specific terms in this expansion, analysis of the pertinent Jacobian matrix shows that the choice of ($\chi > 0, c > 0, a, b \ll 1$) ensures a reentrant phase transition (schematic in Figures 4B and S6A; STAR Methods) with a minimal number of higher-order terms. Results using the Landau model (Figure 4C, equation 1) are recapitulated using a different method for obtaining the free energy (Flory-Huggins) to highlight the generality of our Landau approach (Figures S6A and S6B; STAR Methods). Given the universality of its application, easily characterizable phase behavior, and numerical ease of investigation (e.g., ~ 50 times faster than the Flory-Huggins to study coupled dynamics), we employed the Landau free energy in the rest of this work to study how the dynamics of transcriptional condensates are regulated by transcription.

Figure 3. RNA-Mediated Effects on Condensates in Reconstituted *In Vitro* Transcription Assays

- (A) Cartoon representation of the reconstituted *in vitro* mammalian transcription assay with purified components (left) and the design of the assay (right) (STAR Methods).
- (B) Bright-field images of droplets formed within the *in vitro* transcription reaction. Droplets are stained with DNA dye (Hoechst). Bright-field images were white tophat filtered and smoothed here and in the subsequent panels (STAR Methods) (scale bars = 5 μm).
- (C) Bright-field images of droplets formed within the *in vitro* transcription reaction performed in the presence of the indicated spermine concentrations. Template DNA is labeled with Cy3 (scale bars = 5 μm).
- (D) Droplet sizes in (C) ($p = 0.0011$, Student's t test).
- (E) Partition ratio of Cy3-labeled template DNA in the droplets in (C) ($p < 0.0001$, Student's t test).
- (F) qRT-PCR of transcriptional output upon addition of spermine. The values are normalized to the no-spermine condition. The mean of 2 replicates is shown, and error bars depict SD ($p = 0.0477$, Student's t test).
- (G) Representative images of droplets in the *in vitro* transcription reaction in the presence of the indicated amounts of exogenous RNA (scale bars = 5 μm).
- (H) Droplet sizes in (G) ($p = 0.9309$ 0 versus 10; $p < 0.001$ for 0 versus 50, 250, and 500; one-way ANOVA).
- (I) qRT-PCR of transcriptional output upon addition of increasing concentrations of exogenous RNA. The values are normalized to the no-RNA condition. The mean of 2 replicates is shown, and error bars depict SD ($p = 0.0001$, GTP only versus 0; $p = 0.0111$, 0 versus 10; $p = 0.0013$, 0 versus 50; $p = 0.0008$, 0 versus 250; $p = 0.008$, 0 versus 500; one-way ANOVA).
- See also Figure S5.

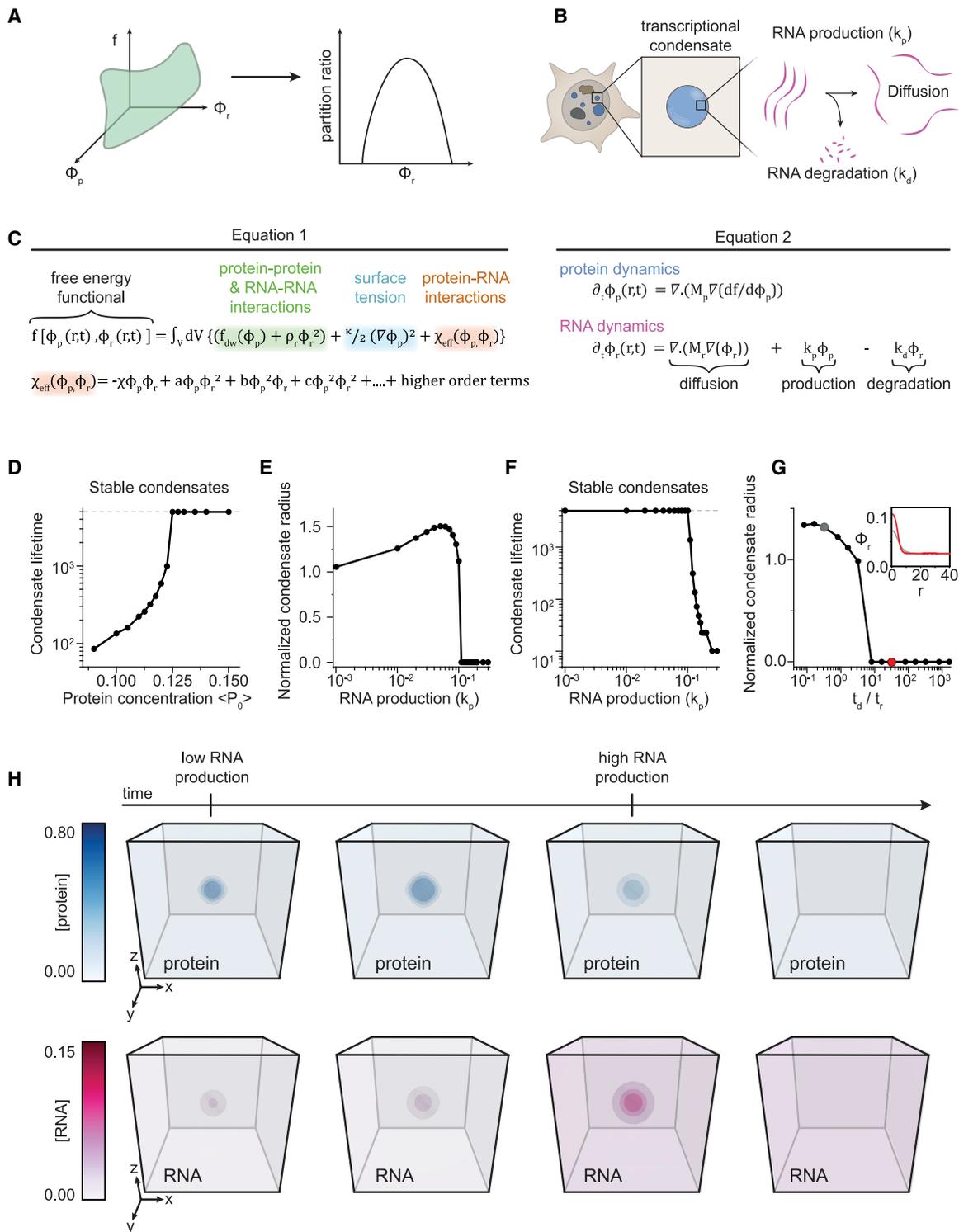


Figure 4. A Model of RNA-Mediated Non-equilibrium Feedback Control of Transcriptional Condensates

(A) Schematic of coarse-grained free energy (f , green surface), which depends on the transcriptional protein (ϕ_p) and RNA (ϕ_r) concentrations. This free energy recapitulates *in vitro* observations of an equilibrium reentrant transition.

(B) Schematic of the non-equilibrium model coupling transcriptional activity with transcriptional condensate dynamics. In the model framework, we focus on a local micro-environment near a single transcriptional condensate (blue). RNA (magenta) is synthesized and degraded and can diffuse.

(C) Equations underlying construction of the free-energy function (equation 1) and dynamics of protein and RNA (equation 2) (STAR Methods).

(legend continued on next page)

We next developed a mathematical framework to study the temporal evolution of transcriptional condensates as transcription ensues. Most transcriptional proteins turn over with a half-life of several hours (Cambridge et al., 2011; Chen et al., 2016), which is longer than timescales of transcription-associated events, which range from seconds to minutes (Chen and Larson, 2016; Fukaya et al., 2016; Rodriguez and Larson, 2020). Hence, the overall amount of protein is conserved in the timescales of interest. Thus, the dynamics of the protein concentration (ϕ_p) are represented by standard model B dynamics (Figure 4C, equation 2; Hohenberg and Halperin, 1977). Under model B dynamics, gradients in the protein chemical potential, which depend on the spatial distribution of protein and RNA concentrations, drive diffusive protein fluxes, which, in turn, drive the spatio-temporal evolution of ϕ_p . Because RNA concentrations vary over transcription-associated timescales, the dynamics of ϕ_r are explicitly governed by a reaction-diffusion equation. The key features (schematic in Figure 4B) are that RNA diffuses with mobility M_{ra} and is synthesized and degraded with specific reaction rates: k_p and k_d , respectively. Because the RNA dynamics are far from equilibrium and the free-energy function noted above depends on protein and RNA concentrations, the coupled temporal evolution of transcriptional proteins and RNA (Figure 4C, equations 1 and 2) cannot be obtained from near-equilibrium considerations of simply going downhill in free energy with time. We employ this mathematical framework to study non-equilibrium regulation of transcriptional condensates.

We first sought to determine whether this model is consistent with previous studies (Cho et al., 2016, 2018). These studies have shown that transcriptional condensates at different genomic loci recruit a varying number of transcriptional proteins, which, in turn, correlates with condensate lifetimes. To explore this phenomenon, we numerically simulated equation 2 (Figure 4C) on 2D and 3D grids (STAR Methods). Locus-dependent recruitment of the transcriptional machinery can be mimicked in our model by varying the total transcriptional protein amount ($\langle P_0 \rangle$) with all other parameters fixed because our simulation volume represents a local micro-environment (Figure 4A). Our simulations predict that loci that can recruit more transcriptional proteins (higher $\langle P_0 \rangle$) form relatively stable condensates, whereas condensates that recruit fewer proteins dissolve after a characteristic lifetime (Figure 4D). The model predictions for transcriptional condensate dynamics are qualitatively consistent with published data (Cho et al., 2016) and suggest that features encoded at genomic loci contribute to transcriptional condensate dynamics.

We next investigated how the sizes and lifetimes of transcriptional condensates change as a function of the effective rate of

RNA synthesis (k_p) while keeping all other parameters fixed. In these simulations, the size of condensates initially increases and subsequently decreases with increasing effective rates of RNA synthesis (Figure 4E). Above a threshold rate of RNA synthesis, condensates dissolve (Figure 4E). The underlying reason for this result is the reentrant phase behavior of mixtures of transcriptional molecules and RNA (Figures 1 and 2). We also find that condensates with higher transcriptional activity dissolve faster, as measured by condensate lifetimes (Figure 4F). Condensate lifetimes do not vary over a range of RNA transcription rates that reflect RNA-transcriptional protein ratios that roughly correspond to the charge balance conditions (Figure 4F). The same qualitative results are recapitulated in 3D simulations (Figure S6C) as well as simulations employing the Flory-Huggins free energy (Figure S6D) and further reinforced by partition ratios computed from simulations (Figure S6E). Further, we carried out simulations that accounted for phase-dependent changes in diffusion of RNA; i.e., RNA diffusion was hindered in the dense phase because of crowding (Figure S6F). Predictions of the condensate size and lifetimes exhibited qualitative trends similar to simulations without this phase-dependent diffusion (Figure S6F; Figures 4E and 4F). Overall, our results suggest a model where low effective rates of RNA synthesis (or low transcription activity) stabilize transcriptional condensates, whereas higher rates promote condensate dissolution.

We then investigated the extent to which non-equilibrium effects underlying transcription regulate transcriptional condensate dynamics. RNA synthesis, degradation, and diffusion influence the spatial distribution of RNA, which, in turn, may feed back on transcriptional condensates. To explore this, we varied the diffusivity of RNA and the effective rates of RNA synthesis and degradation while holding the ratio of synthesis and degradation rates constant. The latter constraint ensures that the overall RNA concentration is constant in the condensate as other parameters are varied; thus, any effect on condensate dynamics arises from purely non-equilibrium effects. Varying the parameters that control RNA synthesis/degradation rates and diffusion changes the relative timescales of these processes (t_r and t_d , respectively) (STAR Methods), which in turn, influences the spatial distribution of RNA in the condensate. If diffusion is slower than synthesis/degradation ($t_r < t_d$), then RNA will accumulate near transcription sites, leading to a higher local RNA concentration in the condensate. Conversely, if diffusion is faster than synthesis/degradation ($t_r > t_d$), then RNA will diffuse away from transcription sites, leading to a lower uniform RNA concentration in the condensate. The spatial distribution of RNA will affect condensates according to local charge balance. To study

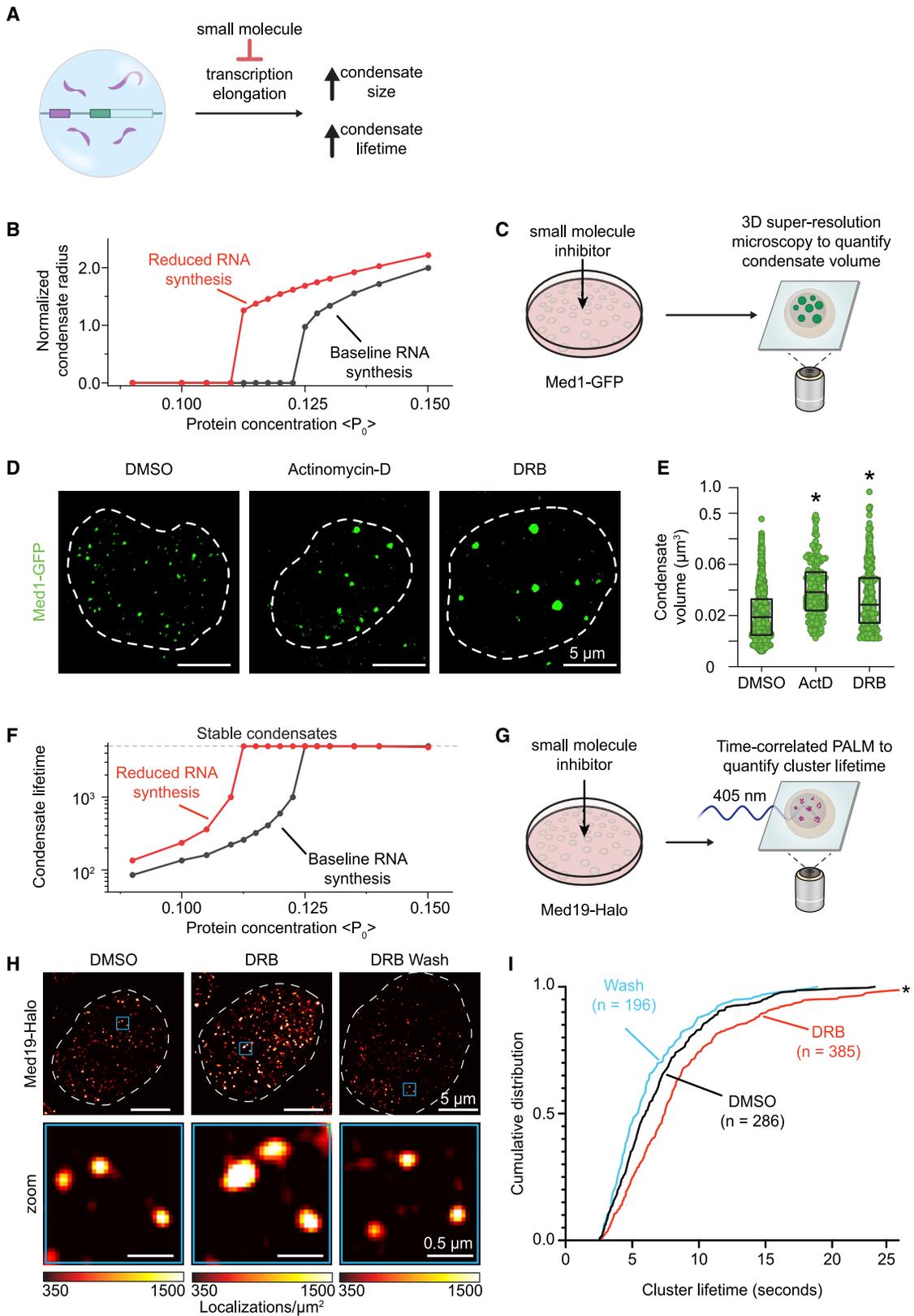
(D) Simulation predictions of transcriptional condensate lifetime with varying total protein concentrations (2D simulation grid). The dashed line represents the lifetime of condensates (in units of simulation time) that do not dissolve at steady state.

(E and F) Simulation predictions of transcriptional condensate radius (E) and lifetime (F) at varying effective rates of RNA synthesis (2D simulation grid). The radius values are normalized to $r = 6.0$ mesh units. The dashed line in (F) represents the lifetime of stable condensates in units of simulation time (STAR Methods).

(G) Variation of normalized condensate radius (ordinate, normalized to $r = 6.0$ mesh units) with changing relative timescales of reaction and diffusion (abscissa, t_d/t_r) (2D simulation grid). In these simulations, the total effective concentration of RNA produced is held constant (see text). The inset graphs the distribution of RNA concentrations at early simulation times ($t_{step} = 100$) for two different values of t_d/t_r (highlighted in the main panel with corresponding colors).

(H) Visualization of protein (blue) and RNA (magenta) concentration fields over simulation time for 3D simulations. The condensate is initialized (first panel) and then grows under low transcriptional activity (second panel). After a finite time ($t_{sim} = 1000$), the effective rate of RNA synthesis (k_p) is increased by 2.5-fold, which, in turn, drives condensate shrinkage (third panel) and, ultimately, dissolution (fourth panel) (STAR Methods).

See also Figures S6 and S7.



(legend on next page)

how varying spatial distributions of RNA affect transcriptional condensates, we simulated conditions where the overall RNA concentration was fixed close to the charge balance condition, promoting condensate formation at equilibrium. In these simulations, condensates that are stable when synthesis/degradation is slower than diffusion ($t_r > t_d$) dissolve when RNA synthesis/degradation is faster than diffusion ($t_r < t_d$) (Figure 4G). When ($t_r > t_d$), RNA concentration is relatively uniform and low throughout the condensate, equilibrium effects dominate. Conversely, when ($t_r < t_d$), RNA is distributed non-uniformly with high local concentrations in the condensate, non-equilibrium effects dominate to result in condensate dissolution (Figure 4G). In the latter case, the localized high RNA concentrations exceed the charge balance condition because of non-equilibrium effects. Approximate estimates for the rates of RNA synthesis, degradation, and diffusion under physiological conditions ($\frac{t_d}{t_r} \approx 2 - 100$, STAR Methods) suggest that transcriptional condensate dynamics are likely driven off equilibrium.

We sought to synthesize our results so far to explore the effect of non-equilibrium dynamics on regulating transcriptional condensates across transcription initiation and productive elongation. Simulations were started at a relatively low effective rate of RNA synthesis, mimicking initiation, followed by an increase to a relatively high effective rate of RNA synthesis, mimicking productive elongation. The simulations predict that low effective rates of RNA synthesis enhance condensate formation, and these condensates subsequently dissolve upon ensuing higher effective rates of RNA synthesis (Figure 4H). Consistent with these simulations, Mediator condensates tend to be depleted in areas of high, RNA Pol II-driven nascent transcription (Figure S7A–S7C). These results suggest that non-equilibrium processes underlying RNA synthesis can potentially regulate formation and dissolution of transcriptional condensates.

Inhibition of RNA Elongation Leads to Enhanced Condensate Size and Lifetime in Cells

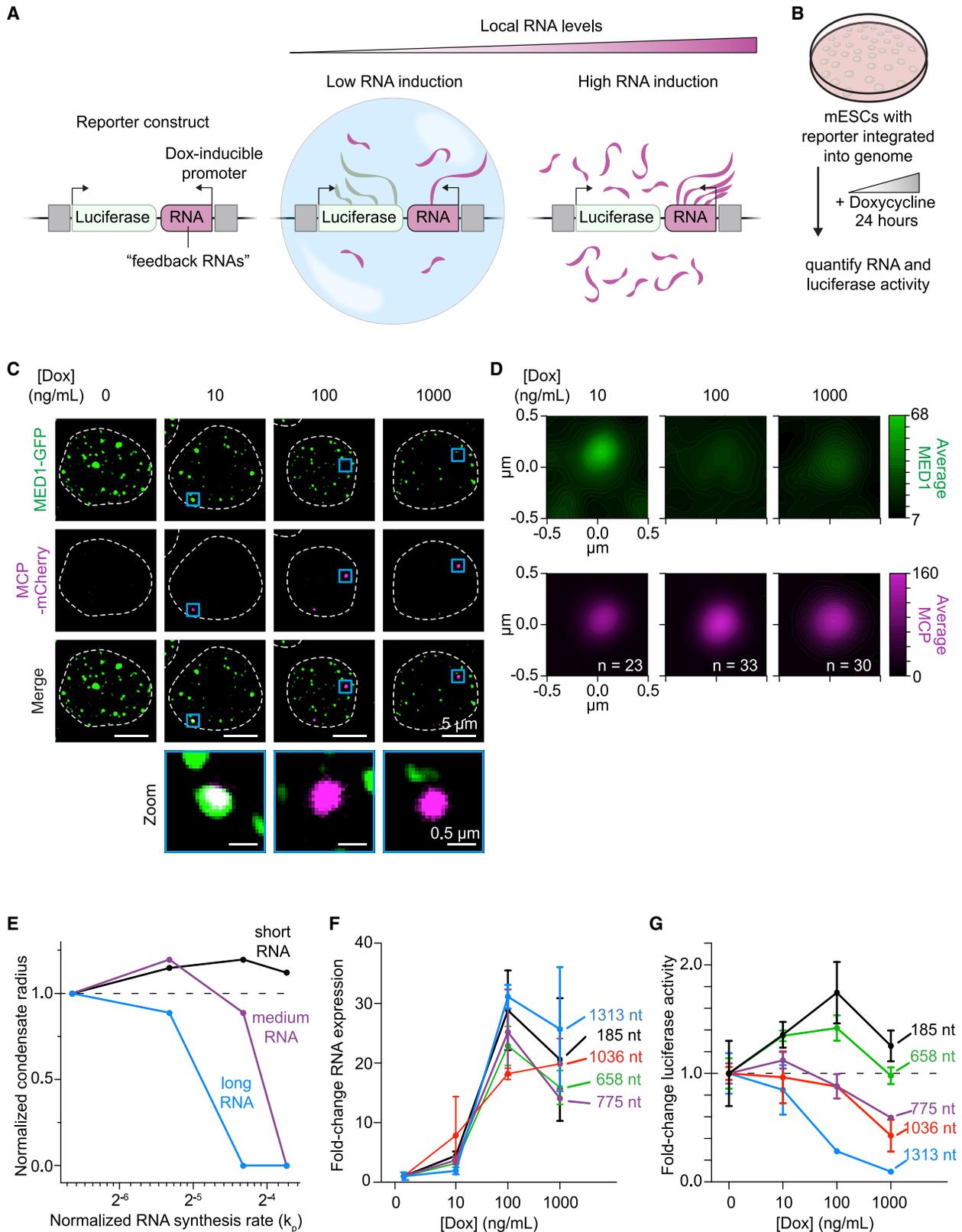
Transcriptional condensates in cells are highly dynamic, forming and dissolving at timescales ranging from seconds to minutes (Cho et al., 2018). We previously showed that condensate formation is associated with transcription activation and initiation (Cho et al., 2018). When transcriptional

condensates are formed, the RNA-mediated condensate dissolution model predicts that inhibition of elongation should increase the size and lifetime of transcriptional condensates (Figure 5A). We used the physics-based model (Figure 4) to simulate the effects of elongation inhibition on transcriptional condensates and performed experiments to test the predictions from these simulations in cells (Figures 5B–5I). To account for the locus-dependent ability to recruit the transcriptional machinery and RNA Pol II, we performed these simulations at a range of total protein concentrations (as in Figures 4D and 4E) but for conditions where the effective rate of RNA synthesis (k_p) was high (corresponding to elongation) and low (corresponding to inhibited elongation). The results of the simulations predict that a reduced effective rate of RNA synthesis should increase the size and lifetime of transcriptional condensates across a range of total protein concentrations (Figure 5B and 5F).

To experimentally test these predictions from the simulations, mESCs engineered with an endogenous, GFP-tagged subunit of Mediator (Med1-GFP) (Sabari et al., 2018) were treated for 30 min with Actinomycin-D (ActD) or 5,6-dichlorobenzimidazole riboside (DRB) (Figure 5C), which disrupt transcription elongation through DNA intercalation and inhibition of CDK9-mediated RNA Pol II pause release, respectively (Singh and Padgett, 2009; Sobell, 1985; Steurer et al., 2018). Consistent with the model predictions, after inhibition of elongation, Med1-GFP condensates increased in volume by ~2-fold as measured by 3D super-resolution microscopy (Figures 5D and 5E). Condensate lifetime could not be assessed in these cells because of the long duration of image acquisition and consequent photobleaching, so we turned to time-correlated photo-activation localization microscopy (tcPALM) super-resolution microscopy in mESCs with an endogenous Med19-Halo tag (Cho et al., 2018; Cisse et al., 2013) to investigate the effects of elongation inhibition on condensate lifetime (Figure 5G). Cells were treated for 30 min with DRB to disrupt transcription elongation, and the lifetime of Med19 condensates was quantified. When transcription elongation was inhibited by DRB treatment, Med19 condensates exhibited significantly longer lifetimes than mock-treated cells (Figures 5H and 5I), and when DRB-treated cells were washed with fresh medium, the lifetimes of the Med19 condensates recovered to those of the mock-

Figure 5. Inhibition of RNA Elongation Leads to Enhanced Condensate Size and Lifetime in Cells

- (A) Scheme for preventing condensate dissolution upon transcriptional burst by treatment with small molecules that inhibit transcriptional elongation.
 (B) Simulation predictions show variation of normalized condensate radius with total protein amount (abscissa) in the absence (black, $k_p = 0.1$) and presence (red, $k_p = 0.05$) of RNA synthesis inhibition (2D simulation grid). The radius is normalized by the radius at $k_p = 0.05$, $\langle P_0 \rangle = 0.115$.
 (C) Experimental design to test the effect of transcriptional inhibition on the size of Mediator condensates. MED1-GFP mESCs are imaged by 3D super-resolution microscopy after treatment with small molecules.
 (D) Maximum intensity projection images of single nuclei tagged with endogenous Med1-GFP in the presence of the indicated transcriptional inhibitors or DMSO control (scale bars = 5 μm).
 (E) Volumes of Med1-GFP condensates in (D) (p for DMSO versus ActD < 0.0001 and p for DMSO versus DRB < 0.0001, one-way ANOVA).
 (F) Simulation predictions show variation of condensate lifetime with total protein amount (abscissa) in the absence (black, $k_p = 0.1$) and presence (red, $k_p = 0.05$) of RNA synthesis inhibition (2D simulation grid). The lifetime is presented in units of simulation time.
 (G) Experimental design to test the effect of DRB on the lifetime of Mediator clusters in Med19-tagged mESCs. Lifetimes are quantified by time-correlated PALM.
 (H) Representative heatmap of Med19-Halo localizations in a single nucleus upon addition of the transcriptional inhibitor DRB, DRB wash, or DMSO control (top scale bars = 5 μm ; bottom scale bars = 0.5 μm).
 (I) Cumulative distribution frequency plot of condensate lifetime in response to the indicated treatments (p < 0.0001, one-way ANOVA).
 See also Figure S7.



(legend on next page)

treated condition (Figures 5H and 5I). The *in silico* and experimental results show that suppression of elongation in cells leads to increased condensate size and lifetime, consistent with a model where a burst of RNA synthesis can promote dissolution of transcriptional condensates in cells.

Increasing the Levels of Local RNA Synthesis Reduces Condensate Formation and Transcription in Cells

The RNA-mediated feedback model suggests that modifying the concentration or size of RNA molecules should have a predictable effect on transcriptional output. We developed complementary experimental and simulation approaches (Figure 6) where the levels of putative “feedback RNAs” could be increased artificially. We first used the physics-based model (Figure 4) to simulate the effect of increasing effective rates of RNA synthesis as well as varying lengths for the synthesized RNA on condensates (STAR Methods). The simulations predicted that increases in the production rate of shorter RNAs initially enhance and subsequently suppress transcriptional condensate size, whereas increases in the production rate of longer RNAs lead to reduced condensate size with increasing synthesis rates (Figure 6E).

To test this prediction, we investigated the effect of artificially increasing the levels of feedback RNAs on transcription of an adjacent luciferase reporter gene in cells (Figures 6A and 6B; Kirk et al., 2018). DNA molecules specifying RNAs of a range of sizes were cloned into this system to allow doxycycline (Dox)-inducible expression of these RNAs, and mESC lines were generated with clones of integrated constructs. Feedback RNAs were observed at loci of Mediator puncta under low-Dox stimulation, suggesting that these actively transcribed genes are associated with transcriptional condensates (Figures 6C and 6D). Elevated expression of feedback RNAs under higher-Dox stimulation reduced their colocalization with Mediator puncta, consistent with the model of RNA-mediated feedback on condensates (Figures 6C, 6D, and S7D). To study the effect of local RNA levels on transcription, additional cell lines harboring diverse feedback RNAs were then treated with increasing doses of Dox to induce feedback RNA expression (Figure 6F), and reporter expression was measured by luminescence (Figure 6G). The results were consistent with model predictions (Figure 6E); increases in the levels of short feedback RNAs initially enhanced reporter expression and then suppressed this, whereas progressive increases in the levels of the longer feedback RNAs reduced reporter expression more strongly (Figure 6G). We confirmed that changes in reporter

expression arise from *cis* RNA-mediated effects by modifying the constructs, controlling for the global effects of Dox, and perturbing local RNA concentration (Figures S7E–S7K). These results support a role for RNA-mediated feedback in the control of transcriptional condensates.

DISCUSSION

The results described here indicate that transcription is a non-equilibrium process that provides dynamic feedback through its RNA product. The results support a model where RNA provides positive and negative feedback on transcription via regulation of electrostatic interactions in transcriptional condensates. Transcriptional condensates, whose formation involves crowding of transcription factors by enhancer DNA (Shrinivas et al., 2019) and electrostatic and other interactions between the IDRs of transcription factors and coactivators (Boija et al., 2018; Sabari et al., 2018), engage RNA to both promote and dissolve the condensates. In this RNA feedback model, low levels of short RNAs produced during transcription initiation promote formation of transcriptional condensates, whereas high levels of the longer RNAs produced during elongation can cause condensate dissolution (Figure 7).

An RNA-mediated feedback model for transcriptional regulation provides a potential explanation for the roles of enhancer and promoter-associated RNAs, which are evolutionarily conserved features of eukaryotes. These low-abundance short RNAs, transcribed bidirectionally from enhancers and promoters, have been reported to affect transcription from their associated genes through diverse postulated mechanisms. The diversity of sequences present in these short RNA species has made it difficult to postulate a common molecular mechanism for their effects on transcription. In this context, a model for RNA-mediated feedback regulation of condensates is attractive for several reasons. RNA molecules are known components of other biomolecular condensates, including the nucleolus, nuclear speckles, paraspeckles, and stress granules, where they are known to play regulatory roles (Fay and Anderson, 2018; Roden and Gladfelter, 2020). RNA is a powerful regulator of condensates that are formed by electrostatic forces because it has a high negative charge density due to its phosphate backbone (Drobot et al., 2018; Frankel et al., 2016), explaining why the effects of diverse RNAs on transcriptional condensates are sequence independent. The functions of most noncoding RNAs remain a mystery, and this model suggests a mechanism

Figure 6. Increasing the Levels of Local RNA Synthesis Reduces Condensate Formation and Transcription in Cells

- (A) Scheme depicting the reporter system (left) where local RNA expression near a luciferase reporter gene can be induced by Dox.
(B) Experimental design to test the effect of increasing local RNA levels on condensate formation and reporter gene expression.
(C) Live-cell imaging showing localization of Mediator condensates and MS2-tagged RNA expressed near the reporter gene with the indicated Dox stimulations. Med1-GFP mESCs have an integrated reporter system and 2×-MS2 coat protein (MCP)-mCherry to visualize MS2-tagged RNA (2,456 nt). Representative images are maximum projections that have been subtracted by a median filter and smoothed (STAR Methods) (top scale bars = 5 μm; bottom scale bars = 0.5 μm).
(D) Average density of the MED1 signal centered at the RNA signal with the indicated Dox stimulations ($p = 0.066$, 10 ng/mL versus 100 ng/mL Dox; $p = 0.013$, 10 ng/mL versus 1000 ng/mL Dox; $p = 0.315$, 100 ng/mL versus 1000 ng/mL; 2-way Kolmogorov-Smirnov test).
(E) Simulations predict the variation of condensate size with increasing effective rates of RNA synthesis (abscissa) (2D simulation grid). The condensate radius is normalized by value at rate = 1, and RNA synthesis rates are normalized to $k_p = 0.02$ (STAR Methods).
(F) qRT-PCR of various “feedback RNAs” with increasing Dox concentrations. Markers show the mean of at least 3 replicates, and error bars depict the SD.
(G) Luciferase luminescence with increasing Dox concentrations. Markers show the mean of at least 3 replicates, and error bars depict the SD.
See also Figure S7.

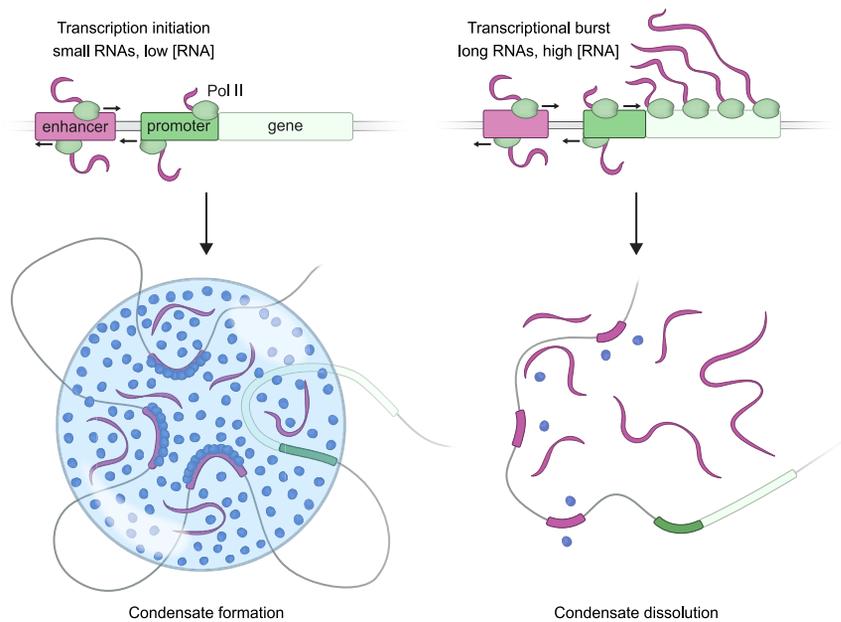


Figure 7. A Model for RNA-Mediated Feedback Control of Transcriptional Condensates

The cartoon depicts a model where low levels of RNA present at transcription initiation promote condensate formation, whereas high levels of RNA present during a transcriptional burst promote condensate dissolution.

by which some of these might participate in tuning local gene expression.

Recent studies indicate that transcription occurs in periodic bursts (~1–10 min in duration), where multiple molecules of RNA Pol II can be released from promoters within a short time frame and produce multiple molecules of mRNA (~1–100 molecules per burst). Multiple models explain such periodic bursts through stochastic gene activation events (Chen and Larson, 2016; Larsson et al., 2019; Raj et al., 2006; Rodriguez and Larson, 2020; Suter et al., 2011; Tunnaclyffe and Chubb, 2020) but are often agnostic to the underlying mechanism or attribute these to rate-limiting transcription factor binding events. We suggest that a rapid and spatially localized change in charge balance, due to increased RNA synthesis at pause release of active RNA Pol II, may contribute to dissolution of transcriptional condensates and thus dynamic loss of the pool of transcriptional apparatus in those condensates. This would provide negative feedback to arrest transcription and a mechanism to account for the dynamic bursty behavior observed for transcription.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
 - Lead Contact
 - Materials Availability
 - Data and Code Availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
 - Cell lines
 - Cell culture conditions
- **METHOD DETAILS**
 - ChIP-seq analysis
- GRO-seq analysis
- RNA-seq analysis
- Calculation number of RNA molecules in cells
- *In vitro* droplet assay
- Fluorescence recovery after photobleaching (FRAP)
- *In vitro* droplet analysis
- Synthesis of RNA by *in vitro* transcription
- Recombinant protein purification
- Purification of human Mediator complex from HeLa nuclear extract.
- Reconstituted *in vitro* transcription assay
- Constructing a free-energy for RNA-protein phase behavior
- Landau model
- Flory-Huggins model
- Numerical phase-field simulations
- Design of Simulations to vary RNA features and rates of RNA synthesis
- Calculation of number of charged molecules in condensates
- Reactive/diffusive time-scales and estimates in cells
- Calculation of charge balance
- Transcription inhibition by small molecules
- Condensate size
- Condensate lifetime
- Nascent RNA imaging
- Reporter assay to determine the effect of local RNA synthesis on transcription
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
 - Statistical analysis of *in vitro* condensate assays:
 - Statistical analysis of *in vitro* transcription assays:
 - Statistical analysis of transcription inhibition experiments
 - Statistical analysis of luciferase reporter and EU experiments:

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.cell.2020.11.030>.

ACKNOWLEDGMENTS

We are grateful to Dora Tang and Ankur Jain for discussions regarding complex coacervates and their regulation by RNA. We also thank Mehdaan Kardar and Young lab members for discussions; D. Reinberg, J.M. Calabrese, R. Jaenisch, and I. Cheeseman for reagents; W. Salmon of the W.M. Keck Microscopy Facility; and E. Diel of the Harvard Center for Biomedical Imaging. This work was supported by NIH grants GM123511 (to R.A.Y.), CA155258 (to R.A.Y.), 1F32CA254216-01 (to J.E.H.), and GM134734 (to I.I.C.); NSF grant PHY-1743900 (to A.K.C., R.A.Y., and P.A.S.); by the NSF-Simons Center for Mathematical and Statistical Analysis of Biology at Harvard (Award Number #1764269) and the Harvard Quantitative Biology Initiative (to K.S.); and the Gruss-Lipper postdoctoral fellowship and the Rothschild postdoctoral fellowship (to I.S.).

AUTHOR CONTRIBUTIONS

Conceptualization, J.E.H., O.O., K.S., T.I.L., P.A.S., A.K.C., and R.A.Y.; Methodology, J.E.H., O.O., K.S., N.M.H., I.S., M.M.Z., J.O.A., A.V.Z., C.L., and A.K.C.; Software, J.E.H. and K.S.; Formal Analysis, J.E.H., O.O., and K.S.; Investigation, J.E.H., O.O., K.S., I.S., M.M.Z., J.O.A., and A.V.Z.; Resources, N.M.H., G.L., I.I.C., A.K.C., and R.A.Y.; Writing – Original Draft, J.E.H., O.O., K.S., A.K.C., and R.A.Y.; Visualization, J.E.H., O.O., and K.S.; Supervision, A.K.C. and R.A.Y.; Funding Acquisition, P.A.S., A.K.C., and R.A.Y.

DECLARATION OF INTERESTS

R.A.Y. is a founder and shareholder of Syros Pharmaceuticals, Camp4 Therapeutics, Omega Therapeutics, and Dewpoint Therapeutics. T.I.L. is a shareholder of Syros Pharmaceuticals and a consultant to Camp4 Therapeutics. A.K.C. and P.A.S. are shareholders and consultants to Dewpoint Therapeutics. A.K.C. is a SAB member of Omega Therapeutics. The Whitehead Institute has filed a patent application based on this study.

Received: April 9, 2020

Revised: August 9, 2020

Accepted: November 16, 2020

Published: December 16, 2020

REFERENCES

- Adelman, K., and Lis, J.T. (2012). Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat. Rev. Genet.* *13*, 720–731.
- Adhikari, S., Leaf, M.A., and Muthukumar, M. (2018). Polyelectrolyte complex coacervation by electrostatic dipolar interactions. *J. Chem. Phys.* *149*, 163308.
- Andersson, R., Gebhard, C., Miguel-Escalada, I., Hoof, I., Bornholdt, J., Boyd, M., Chen, Y., Zhao, X., Schmidl, C., Suzuki, T., et al. (2014). An atlas of active enhancers across human cell types and tissues. *Nature* *507*, 455–461.
- Andrews, J.O., Conway, W., Cho, W.-K., Narayanan, A., Spille, J.-H., Jayanth, N., Inoue, T., Mullen, S., Thaler, J., and Cissé, I.I. (2018). qSR: a quantitative super-resolution analysis tool reveals the cell-cycle dependent organization of RNA Polymerase I in live human cells. *Sci. Rep.* *8*, 7424.
- Aumiller, W.M., Jr., Pir Cakmak, F., Davis, B.W., and Keating, C.D. (2016). RNA-Based Coacervates as a Model for Membraneless Organelles: Formation, Properties, and Interfacial Liposome Assembly. *Langmuir* *32*, 10042–10053.
- Azofeifa, J.G., Allen, M.A., Hendrix, J.R., Read, T., Rubin, J.D., and Dowell, R.D. (2018). Enhancer RNA profiling predicts transcription factor activity. *Genome Res.* *28*, 334–344.
- 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.
- Banerjee, P.R., Milin, A.N., Moosa, M.M., Onuchic, P.L., and Deniz, A.A. (2017). Reentrant Phase Transition Drives Dynamic Substructure Formation in Ribonucleoprotein Droplets. *Angew. Chem. Int. Ed. Engl.* *56*, 11354–11359.
- Bergot, M.O., Diaz-Guerra, M.J., Puzenat, N., Raymondjean, M., and Kahn, A. (1992). Cis-regulation of the L-type pyruvate kinase gene promoter by glucose, insulin and cyclic AMP. *Nucleic Acids Res.* *20*, 1871–1877.
- Blair, D.G.R. (1985). Activation of mammalian RNA polymerases by polyamines. *Int. J. Biochem.* *17*, 23–30.
- Boeynaems, S., Holehouse, A.S., Weinhardt, V., Kovacs, D., Van Lindt, J., Larabell, C., Van Den Bosch, L., Das, R., Tompa, P.S., Pappu, R.V., and Gitler, A.D. (2019). Spontaneous driving forces give rise to protein-RNA condensates with coexisting phases and complex material properties. *Proc. Natl. Acad. Sci. USA* *116*, 7889–7898.
- Boija, A., Klein, I.A., Sabari, B.R., Dall’Agnese, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* *175*, 1842–1855.e6.
- Brandman, O., and Meyer, T. (2008). Feedback loops shape cellular signals in space and time. *Science* *322*, 390–395.
- Bruhat, A., Jousse, C., Carraro, V., Reimold, A.M., Ferrara, M., and Fournoux, P. (2000). Amino acids control mammalian gene transcription: activating transcription factor 2 is essential for the amino acid responsiveness of the CHOP promoter. *Mol. Cell. Biol.* *20*, 7192–7204.
- Cambridge, S.B., Gnad, F., Nguyen, C., Bermejo, J.L., Krüger, M., and Mann, M. (2011). Systems-wide proteomic analysis in mammalian cells reveals conserved, functional protein turnover. *J. Proteome Res.* *10*, 5275–5284.
- Carey, M.F., Peterson, C.L., and Smale, S.T. (2009). *Transcriptional Regulation in Eukaryotes: Concepts, Strategies, and Techniques*, Second Edition (Cold Spring Harbor Laboratory Press).
- Catarino, R.R., and Stark, A. (2018). Assessing sufficiency and necessity of enhancer activities for gene expression and the mechanisms of transcription activation. *Genes Dev.* *32*, 202–223.
- Chen, H., and Larson, D.R. (2016). What have single-molecule studies taught us about gene expression? *Genes Dev.* *30*, 1796–1810.
- Chen, W., Smeekens, J.M., and Wu, R. (2016). Systematic study of the dynamics and half-lives of newly synthesized proteins in human cells. *Chem. Sci. (Camb.)* *7*, 1393–1400.
- Chiu, A.C., Suzuki, H.I., Wu, X., Mahat, D.B., Kriz, A.J., and Sharp, P.A. (2018). Transcriptional Pause Sites Delineate Stable Nucleosome-Associated Premature Polyadenylation Suppressed by U1 snRNP. *Mol. Cell* *69*, 648–663.e7.
- Cho, W.-K., Jayanth, N., English, B.P., Inoue, T., Andrews, J.O., Conway, W., Grimm, J.B., Spille, J.-H., Lavis, L.D., Lionnet, T., and Cisse, I.I. (2016). RNA Polymerase II cluster dynamics predict mRNA output in living cells. *eLife* *5*, e13617.
- Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I.I. (2018). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* *361*, 412–415.
- Chubb, J.R., Trcek, T., Shenoy, S.M., and Singer, R.H. (2006). Transcriptional pulsing of a developmental gene. *Curr. Biol.* *16*, 1018–1025.
- Cisse, I.I., Izeddin, I., Causse, S.Z., Boudarene, L., Senecal, A., Muresan, L., Dugast-Darzacq, C., Hajj, B., Dahan, M., and Darzacq, X. (2013). Real-time dynamics of RNA polymerase II clustering in live human cells. *Science* *341*, 664–667.
- Core, L., and Adelman, K. (2019). Promoter-proximal pausing of RNA polymerase II: a nexus of gene regulation. *Genes Dev.* *33*, 960–982.
- Core, L.J., Martins, A.L., Danko, C.G., Waters, C.T., Siepel, A., and Lis, J.T. (2014). Analysis of nascent RNA identifies a unified architecture of initiation regions at mammalian promoters and enhancers. *Nat. Genet.* *46*, 1311–1320.

- Cramer, P. (2019). Organization and regulation of gene transcription. *Nature* 573, 45–54.
- Delaney, K.T., and Fredrickson, G.H. (2017). Theory of polyelectrolyte complexation—Complex coacervates are self-coacervates. *J. Chem. Phys.* 146, 224902.
- Di Tommaso, P., Chatzou, M., Floden, E.W., Barja, P.P., Palumbo, E., and Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nat. Biotechnol.* 35, 316–319.
- Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* 11, 1475–1489.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- Drobot, B., Iglesias-Artola, J.M., Le Vay, K., Mayr, V., Kar, M., Kreysing, M., Mutschler, H., and Tang, T.D. (2018). Compartmentalised RNA catalysis in membrane-free coacervate protocells. *Nat. Commun.* 9, 3643.
- Dunlap, J.C. (1999). Molecular bases for circadian clocks. *Cell* 96, 271–290.
- Ebert, B.L., and Bunn, H.F. (1999). Regulation of the erythropoietin gene. *Blood* 94, 1864–1877.
- Elowitz, M.B., and Leibler, S. (2000). A synthetic oscillatory network of transcriptional regulators. *Nature* 403, 335–338.
- Ewels, P.A., Peltzer, A., Fillinger, S., Patel, H., Alneberg, J., Wilm, A., Garcia, M.U., Di Tommaso, P., and Nahnsen, S. (2020). The nf-core framework for community-curated bioinformatics pipelines. *Nat. Biotechnol.* 38, 276–278.
- Fay, M.M., and Anderson, P.J. (2018). The role of RNA in biological phase separations. *J. Mol. Biol.* 430, 4685–4701.
- Flores, O., Lu, H., and Reinberg, D. (1992). Factors involved in specific transcription by mammalian RNA polymerase II. Identification and characterization of factor IIIH. *J. Biol. Chem.* 267, 2786–2793.
- Flory, P.J. (1942). Thermodynamics of High Polymer Solutions. *J. Chem. Phys.* 10, 51–61.
- Frankel, E.A., Bevilacqua, P.C., and Keating, C.D. (2016). Polyamine/Nucleotide Coacervates Provide Strong Compartmentalization of Mg²⁺, Nucleotides, and RNA. *Langmuir* 32, 2041–2049.
- Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer Control of Transcriptional Bursting. *Cell* 166, 358–368.
- Gardini, A., and Shiekhattar, R. (2015). The many faces of long noncoding RNAs. *FEBS J.* 282, 1647–1657.
- Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403, 339–342.
- Gu, B., Swigut, T., Spencley, A., Bauer, M.R., Chung, M., Meyer, T., and Wysocka, J. (2018). Transcription-coupled changes in nuclear mobility of mammalian cis-regulatory elements. *Science* 359, 1050–1055.
- Guo, Y.E., Manteiga, J.C., Henninger, J.E., Sabari, B.R., Dall'Agnese, A., Hannett, N.M., Spille, J.-H., Afeyan, L.K., Zamudio, A.V., Shrinivas, K., et al. (2019). Pol II phosphorylation regulates a switch between transcriptional and splicing condensates. *Nature* 572, 543–548.
- Guyer, J.E., Wheeler, D., and Warren, J.A. (2009). FiPy: Partial Differential Equations with Python. *Comput. Sci. Eng.* 11, 6–15.
- Henriques, T., Scruggs, B.S., Inouye, 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.
- Hnisz, D., Shrinivas, K., Young, R.A., Chakraborty, A.K., and Sharp, P.A. (2017). A Phase Separation Model for Transcriptional Control. *Cell* 169, 13–23.
- Hohenberg, P.C., and Halperin, B.I. (1977). Theory of dynamic critical phenomena. *Rev. Mod. Phys.* 49, 435–479.
- Jangi, M., and Sharp, P.A. (2014). Building robust transcriptomes with master splicing factors. *Cell* 159, 487–498.
- Jin, Y., Eser, U., Struhl, K., and Churchman, L.S. (2017). The Ground State and Evolution of Promoter Region Directionality. *Cell* 170, 889–898.e10.
- Kardar, M. (2007). *Statistical physics of fields* (Cambridge Univ. Press).
- Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* 32, D493–D496.
- 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., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465, 182–187.
- Kirk, J.M., Kim, S.O., Inoue, K., Smola, M.J., Lee, D.M., Schertzer, M.D., Wooten, J.S., Baker, A.R., Sprague, D., Collins, D.W., et al. (2018). Functional classification of long non-coding RNAs by k-mer content. *Nat. Genet.* 50, 1474–1482.
- Klein, I.A., Bojia, A., Afeyan, L.K., Hawken, S.W., Fan, M., Dall'Agnese, A., Oksuz, O., Henninger, J.E., Shrinivas, K., Sabari, B.R., et al. (2020). Partitioning of cancer therapeutics in nuclear condensates. *Science* 368, 1386–1392.
- Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A.J., Elowitz, M.B., and Alon, U. (2004). Dynamics of the p53-Mdm2 feedback loop in individual cells. *Nat. Genet.* 36, 147–150.
- Lai, F., Orom, U.A., Cesaroni, M., Beringer, M., Taatjes, D.J., Blobel, G.A., and Shiekhattar, R. (2013). Activating RNAs associate with Mediator to enhance chromatin architecture and transcription. *Nature* 494, 497–501.
- Landau, L.D. (1937). ON THE THEORY OF PHASE TRANSITIONS. *Zh. Eksp. Teor. Fiz* 11, 19.
- Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 10, R25.
- Larsson, A.J.M., Johnsson, P., Hagemann-Jensen, M., Hartmanis, L., Faridani, O.R., Reinius, B., Segerstolpe, Å., Rivera, C.M., Ren, B., and Sandberg, R. (2019). Genomic encoding of transcriptional burst kinetics. *Nature* 565, 251–254.
- Leinonen, R., Sugawara, H., and Shumway, M.; International Nucleotide Sequence Database Collaboration (2011). The sequence read archive. *Nucleic Acids Res.* 39, D19–D21.
- LeRoy, G., Rickards, B., and Flint, S.J. (2008). The Double Bromodomain Proteins Brd2 and Brd3 Couple Histone Acetylation to Transcription. *Mol. Cell* 30, 51–60.
- LeRoy, G., Oksuz, O., Descostes, N., Aoi, Y., Ganai, R.A., Kara, H.O., Yu, J.-R., Lee, C.-H., Stafford, J., Shilatifard, A., and Reinberg, D. (2019). LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells. *Sci. Adv.* 5, eaay3068.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- Li, W., Notani, D., and Rosenfeld, M.G. (2016). Enhancers as non-coding RNA transcription units: recent insights and future perspectives. *Nat. Rev. Genet.* 17, 207–223.
- Li, C.H., Coffey, E.L., Dall'Agnese, A., Hannett, N.M., Tang, X., Henninger, J.E., Platt, J.M., Oksuz, O., Zamudio, A.V., Afeyan, L.K., et al. (2020). MeCP2 links heterochromatin condensates and neurodevelopmental disease. *Nature* 586, 440–444.
- Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923–930.
- Lin, Y., McCarty, J., Rauch, J.N., Delaney, K.T., Kosik, K.S., Fredrickson, G.H., Shea, J.-E., and Han, S. (2019). Narrow equilibrium window for complex coacervation of tau and RNA under cellular conditions. *eLife* 8, e42571.
- Lin, Y., Fichou, Y., Zeng, Z., Hu, N.Y., and Han, S. (2020). Electrostatically Driven Complex Coacervation and Amyloid Aggregation of Tau Are Independent Processes with Overlapping Conditions. *ACS Chem. Neurosci.* 11, 615–627.

- Maharana, S., Wang, J., Papadopoulos, D.K., Richter, D., Pozniakovsky, A., Poser, I., Bickle, M., Rizk, S., Guillén-Boixet, J., Franzmann, T.M., et al. (2018). RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **360**, 918–921.
- Maiuri, P., Knezevich, A., De Marco, A., Mazza, D., Kula, A., McNally, J.G., and Marcello, A. (2011). Fast transcription rates of RNA polymerase II in human cells. *EMBO Rep.* **12**, 1280–1285.
- Mikhaylichenko, O., Bondarenko, V., Harnett, D., Schor, I.E., Males, M., Viales, R.R., and Furlong, E.E.M. (2018). The degree of enhancer or promoter activity is reflected by the levels and directionality of eRNA transcription. *Genes Dev.* **32**, 42–57.
- Milín, A.N., and Deniz, A.A. (2018). Reentrant Phase Transitions and Non-Equilibrium Dynamics in Membraneless Organelles. *Biochemistry* **57**, 2470–2477.
- Monod, J., and Jacob, F. (1961). Teleonomic mechanisms in cellular metabolism, growth, and differentiation. *Cold Spring Harb. Symp. Quant. Biol.* **26**, 389–401.
- Moruzzi, G., Barbiroli, B., Moruzzi, M.S., and Tadolini, B. (1975). The effect of spermine on transcription of mammalian chromatin by mammalian deoxyribonucleic acid-dependent ribonucleic acid polymerase. *Biochem. J.* **146**, 697–703.
- Mountain, G.A., and Keating, C.D. (2020). Formation of Multiphase Complex Coacervates and Partitioning of Biomolecules within them. *Bio-macromolecules* **21**, 630–640.
- Muthukumar, M. (2016). Electrostatic correlations in polyelectrolyte solutions. *Polym. Sci. Ser. A Chem. Phys.* **58**, 852–863.
- Nair, S.J., Yang, L., Meluzzi, D., Oh, S., Yang, F., Friedman, M.J., Wang, S., Suter, T., Alshareedah, I., Gamliel, A., et al. (2019). Phase separation of ligand-activated enhancers licenses cooperative chromosomal enhancer assembly. *Nat. Struct. Mol. Biol.* **26**, 193–203.
- Niewidok, B., Igaev, M., Pereira da Graca, A., Strassner, A., Lenzen, C., Richter, C.P., Piehler, J., Kurre, R., and Brandt, R. (2018). Single-molecule imaging reveals dynamic biphasic partition of RNA-binding proteins in stress granules. *J. Cell Biol.* **217**, 1303–1318.
- Nott, T.J., Petsalaki, E., Farber, P., Jervis, D., Fussner, E., Plochowitz, A., Craggs, T.D., Bazett-Jones, D.P., Pawson, T., Forman-Kay, J.D., and Baldwin, A.J. (2015). Phase transition of a disordered nucleosome protein generates environmentally responsive membraneless organelles. *Mol. Cell* **57**, 936–947.
- Orphanides, G., LeRoy, G., Chang, C.-H., Luse, D.S., and Reinberg, D. (1998). FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92**, 105–116.
- Overbeek, J.T.G., and Voorn, M.J. (1957). Phase separation in polyelectrolyte solutions; theory of complex coacervation. *J. Cell. Physiol. Suppl.* **49 (Suppl 1)**, 7–22, discussion, 22–26.
- Pai, D.A., Kaplan, C.D., Kweon, H.K., Murakami, K., Andrews, P.C., and Engelke, D.R. (2014). RNAs nonspecifically inhibit RNA polymerase II by preventing binding to the DNA template. *RNA* **20**, 644–655.
- Pefanis, E., Wang, J., Rothschild, G., Lim, J., Kazadi, D., Sun, J., Federation, A., Chao, J., Elliott, O., Liu, Z.-P., et al. (2015). RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. *Cell* **161**, 774–789.
- Peran, I., and Mittag, T. (2020). Molecular structure in biomolecular condensates. *Curr. Opin. Struct. Biol.* **60**, 17–26.
- Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842.
- Rahnamoun, H., Lee, J., Sun, Z., Lu, H., Ramsey, K.M., Komives, E.A., and Lauberth, S.M. (2018). RNAs interact with BRD4 to promote enhanced chromatin engagement and transcription activation. *Nat. Struct. Mol. Biol.* **25**, 687–697.
- Raj, A., and van Oudenaarden, A. (2008). Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226.
- Raj, A., Peskin, C.S., Tranchina, D., Vargas, D.Y., and Tyagi, S. (2006). Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.* **4**, e309.
- Reinberg, D., and Roeder, R.G. (1987). Factors involved in specific transcription by mammalian RNA polymerase II. Transcription factor IIS stimulates elongation of RNA chains. *J. Biol. Chem.* **262**, 3331–3337.
- Roden, C., and Gladfelter, A.S. (2020). RNA contributions to the form and function of biomolecular condensates. *Nat. Rev. Mol. Cell Biol.* Published online July 6, 2020. <https://doi.org/10.1038/s41580-020-0264-6>.
- Rodriguez, J., and Larson, D.R. (2020). Transcription in Living Cells: Molecular Mechanisms of Bursting. *Annu. Rev. Biochem.* **89**, 189–212.
- Roeder, R.G. (2019). 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms. *Nat. Struct. Mol. Biol.* **26**, 783–791.
- Sabari, B.R., Dall’Agnese, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958.
- Sabari, B.R., Dall’Agnese, A., and Young, R.A. (2020). Biomolecular Condensates in the Nucleus. *Trends Biochem. Sci.* **45**, 961–977.
- Schaukowitz, K., Joo, J.-Y., Liu, X., Watts, J.K., Martinez, C., and Kim, T.-K. (2014). Enhancer RNA facilitates NELF release from immediate early genes. *Mol. Cell* **56**, 29–42.
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682.
- Scruggs, B.S., Gilchrist, D.A., Nechaev, S., Muse, G.W., Burkholder, A., Fargo, D.C., and Adelman, K. (2015). Bidirectional Transcription Arises from Two Distinct Hubs of Transcription Factor Binding and Active Chromatin. *Mol. Cell* **58**, 1101–1112.
- Seila, A.C., Calabrese, J.M., Levine, S.S., Yeo, G.W., Rahl, P.B., Flynn, R.A., Young, R.A., and Sharp, P.A. (2008). Divergent transcription from active promoters. *Science* **322**, 1849–1851.
- Sellick, C.A., and Reece, R.J. (2003). Modulation of transcription factor function by an amino acid: activation of Put3p by proline. *EMBO J.* **22**, 5147–5153.
- Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382.
- Shin, Y., and Brangwynne, C.P. (2017). Liquid phase condensation in cell physiology and disease. *Science* **357**, eaaf4382.
- Shrinivas, K., Sabari, B.R., Coffey, E.L., Klein, I.A., Boija, A., Zamudio, A.V., Schuijers, J., Hannett, N.M., Sharp, P.A., Young, R.A., and Chakraborty, A.K. (2019). Enhancer Features that Drive Formation of Transcriptional Condensates. *Mol. Cell* **75**, 549–561.e7.
- Sigova, A.A., Abraham, B.J., Ji, X., Molin, B., Hannett, N.M., Guo, Y.E., Jangi, M., Giallourakis, C.C., Sharp, P.A., and Young, R.A. (2015). Transcription factor trapping by RNA in gene regulatory elements. *Science* **350**, 978–981.
- Sing, C.E. (2017). Development of the modern theory of polymeric complex coacervation. *Adv. Colloid Interface Sci.* **239**, 2–16.
- Singh, J., and Padgett, R.A. (2009). Rates of in situ transcription and splicing in large human genes. *Nat. Struct. Mol. Biol.* **16**, 1128–1133.
- Smith, K.N., Miller, S.C., Varani, G., Calabrese, J.M., and Magnuson, T. (2019). Multimodal Long Noncoding RNA Interaction Networks: Control Panels for Cell Fate Specification. *Genetics* **213**, 1093–1110.
- Sobell, H.M. (1985). Actinomycin and DNA transcription. *Proc. Natl. Acad. Sci. USA* **82**, 5328–5331.
- Srivastava, S., and Tirrell, M.V. (2016). Polyelectrolyte complexation. *Adv. Chem. Phys.* **161**, 499–544.
- Steuere, B., Janssens, R.C., Geverts, B., Geijer, M.E., Wienholz, F., Theil, A.F., Chang, J., Dealy, S., Pothof, J., van Cappellen, W.A., et al. (2018). Live-cell analysis of endogenous GFP-RPB1 uncovers rapid turnover of initiating and promoter-paused RNA Polymerase II. *Proc. Natl. Acad. Sci. USA* **115**, E4368–E4376.
- Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2020). Cellpose: a generalist algorithm for cellular segmentation. *bioRxiv*. <https://doi.org/10.1101/2020.02.02.931238>.
- Strom, A.R., and Brangwynne, C.P. (2019). The liquid nucleome - phase transitions in the nucleus at a glance. *J. Cell Sci.* **132**, jcs235093.

- Struhl, K. (2007). Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat. Struct. Mol. Biol.* *14*, 103–105.
- Suter, D.M., Molina, N., Gatfield, D., Schneider, K., Schibler, U., and Naef, F. (2011). Mammalian genes are transcribed with widely different bursting kinetics. *Science* *332*, 472–474.
- Taylor, N.O., Wei, M.-T., Stone, H.A., and Brangwynne, C.P. (2019). Quantifying Dynamics in Phase-Separated Condensates Using Fluorescence Recovery after Photobleaching. *Biophys. J.* *117*, 1285–1300.
- Tunnacliffe, E., and Chubb, J.R. (2020). What Is a Transcriptional Burst? *Trends Genet.* *36*, 288–297.
- Umbarger, H.E. (1956). Evidence for a negative-feedback mechanism in the biosynthesis of isoleucine. *Science* *123*, 848.
- 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.
- Zamudio, A.V., Dall'Agnese, A., Henninger, J.E., Manteiga, J.C., Afeyan, L.K., Hannett, N.M., Coffey, E.L., Li, C.H., Oksuz, O., Sabari, B.R., et al. (2019). Mediator Condensates Localize Signaling Factors to Key Cell Identity Genes. *Mol. Cell* *76*, 753–766.e6.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* *9*, R137.
- Zhang, P., Shen, K., Alsaifi, N.M., and Wang, Z.-G. (2018). Salt Partitioning in Complex Coacervation of Symmetric Polyelectrolytes. *Macromolecules* *51*, 5586–5593.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
MED1/TRAP220	Abcam	ab64965; RRID:AB_1142301
RNAP2-S2	Millipore Sigma	04-1571; RRID:AB_10627998
Goat anti-rabbit AlexaFluor-488	ThermoFisher	A11008; RRID:AB_143165
Goat anti-rat AlexaFluor-647	Invitrogen	A21247; RRID:AB_141778
Bacterial and Virus Strains		
LOBSTR cells	Cheeseman Lab (WI/MIT)	N/A
Chemicals, Peptides, and Recombinant Proteins		
DRB	Sigma	D1916-50MG
Actinomycin D	Sigma	A1410-5MG
Doxycycline	Sigma	D9891-5G
Poly-L-ornithine	Sigma	P4957-50ML
Laminin (Mouse)	ThermoFisher	CB40232-1MG
Hoechst 33342	Life Technologies	62249
MED1-IDR-GFP	Sabari et al., 2018	N/A
MED1-IDR-GFP (RHK > A)	This study	N/A
BRD4-IDR-GFP	Sabari et al., 2018	N/A
GFP	Sabari et al., 2018	N/A
OCT4-GFP	Boija et al., 2018	N/A
Heparin	Sigma	H3393
Spermine	Sigma	S4264
Critical Commercial Assays		
Luciferase Assay System	Promega	E1500
Power SYBR Green mix	Life Technologies	4367659
NEBuilder HiFi DNA Assembly Master Mix	NEB	E2621S
RNeasy Plus Mini Kit	QIAGEN	74136
Phusion polymerase	NEB	M0531S
Qiaquick Gel Extraction Kit	QIAGEN	28706
MEGAscript T7 Transcription Kit	Invitrogen	AM1334
MEGAscript SP6 Transcription Kit	Invitrogen	AM1330
MEGAclear Transcription Clean-Up Kit	Invitrogen	AM1908
Cy5-labeled UTP	Enzo LifeSciences	ENZ-42506
RNeasy Mini Kit	QIAGEN	74106
Superscript III Reverse Transcriptase	Invitrogen	18080093
BCA protein assay kit	Invitrogen	23227
Lipofectamine 3000 Transfection Agent	ThermoFisher	L3000001
Click-iT RNA Alexa Fluor 594 Imaging Kit	Invitrogen	C10330
Deposited Data		
GROseq of mESCs	Sigova et al., 2015	GSM1665566
ChIP-seq for MED1, PolII, BRD4	Sabari et al., 2018	GSE112808
ChIP-seq for OCT4	Whyte et al., 2013	GSM1082340
Experimental Models: Cell Lines		
V6.5 murine embryonic stem cells	Jaenisch laboratory	N/A
Med1-GFP mESCs	Sabari et al., 2018	N/A

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Med19-Halo mESCs	This study	N/A
TETRIS_eOct4_02 mESCs	This study	N/A
TETRIS_eOct4_03 mESCs	This study	N/A
TETRIS_eTrim28_02 mESCs	This study	N/A
TETRIS_Oct4mRNA mESCs	This study	N/A
TETRIS_24xMS2_Oct4mRNA mESCs	This study	N/A
TETRIS_tandem_Empty mESCs	This study	N/A
TETRIS_tandem_Oct4mRNA mESCs	This study	N/A
TETRIS_divergent_Empty mESCs	This study	N/A
TETRIS_divergent_Oct4mRNA mESCs	This study	N/A
TETRIS_split_Empty mESCs	This study	N/A
TETRIS_split_Oct4mRNA mESCs	This study	N/A
TETRIS_split_Luciferase	This study	N/A
TETRIS_EMPTY mESCs	This study	N/A
Oligonucleotides		
See Table S1 for all primers	This study	N/A
Recombinant DNA		
pETEC-GFP	Sabari et al., 2018.	N/A
pETEC-MED1-IDR-GFP	Sabari et al., 2018.	N/A
pETEC-OCT4-GFP	Boija et al., 2018.	N/A
tJH139_pTETRIS_cargo-STOP_eOct4_02	This study	N/A
tJH138_pTETRIS_cargo-STOP_eOct4_03	This study	N/A
tJH140_pTETRIS_cargo-STOP_eTrim28_02	This study	N/A
tJH135_pTETRIS_cargo-STOP_Oct4mRNA	This study	N/A
tJH097_pTETRIS_cargo-STOP_6xMS2_EMPTY	Modified from Kirk et al., 2018	N/A
tJH098_pTETRIS_cargo-STOP_EMPTY	Modified from Kirk et al., 2018	N/A
tJH150_TETRIS_LUC_PURO	This study	N/A
tJH153_TETRIS_OCT4mRNA_rtta	This study	N/A
tJH157_TETRIS_STOPpA_InverseLuc_Empty	This study	N/A
tJH160_TETRIS_STOPpA_24xMS2-Oct4mRNA	This study	N/A
tJH163_TETRIS_STOPpA_InverseLuc_Oct4mRNA	This study	N/A
tJH164_TETRIS_STOPpA_Divergent_Empty	This study	N/A
tJH165_TETRIS_STOPpA_Divergent_Oct4mRNA	This study	N/A
rtTA-cargo	Kirk et al., 2018	N/A
pJH135_pb_MCPx2_mCherry_rTTA	This study	N/A
Software and Algorithms		
GPP sgRNA Designe	Broad Institute	https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design
Fiji image processing package	Schindelin et al., 2012	https://fiji.sc/
Prism	GraphPad	https://www.graphpad.com/scientific-software/prism/
Code generated by the study	This study	https://github.com/krishna-shrinivas/2020_Henninger_Oksuz_Shrinivas_RNA_feedback
Bowtie2	Langmead and Salzberg, 2012	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Samtools	Li et al., 2009	http://samtools.sourceforge.net/

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bowtie1	Langmead et al., 2009	http://bowtie-bio.sourceforge.net/index.shtml
MACS2	Zhang et al., 2008	https://github.com/mac3-project/MACS
featureCounts v1.6.2	Liao et al., 2014	http://subread.sourceforge.net/
SRA toolkit	Leinonen et al., 2011	https://github.com/ncbi/sra-tools
BEDTools v.2.26.0	Quinlan and Hall, 2010	https://bedtools.readthedocs.io/en/latest/index.html
nf-core RNA-seq pipeline v.1.4.2	Ewels et al., 2020	https://github.com/nf-core/rnaseq
STAR v2.6.1d	Dobin et al., 2013	https://github.com/alexdobin/STAR
Nextflow v20.01.0	Di Tommaso et al., 2017	https://www.nextflow.io/index.html
Cellpose v0.7.2	Stringer et al., 2020	http://www.cellpose.org
Other		
35 mm glass-bottom imaging dishes	Mattek Corporation	P35G-1.5-20-C

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Richard A. Young (young@wi.mit.edu)

Materials Availability

All unique/stable reagents generated in this study are available from the Lead Contact upon reasonable request with a completed Materials Transfer Agreement.

Data and Code Availability

The code generated during this study is available at:

https://github.com/krishna-shrinivas/2020_Henninger_Oksuz_Shrinivas_RNA_feedback.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines

The Jaenisch laboratory of the Whitehead Institute gifted the V6.5 murine embryonic stem cells. These cells are male cells derived from a cross of C57BL/6(F) x 129/sv(M).

Cell culture conditions

ES cells were maintained at 37°C with 5% CO₂ in a humidified incubator on 0.2% gelatinized (Sigma, G1890) tissue-culture plates in 2i medium with LIF, which was made according to the following recipe: 960 mL DMEM/F12 (Life Technologies, 11320082), 5 mL N2 supplement (Life Technologies, 17502048; stock 100X), 10 mL B27 supplement (Life Technologies, 17504044; stock 50X), 5 mL additional L-glutamine (GIBCO 25030-081; stock 200 mM), 10 mL MEM nonessential amino acids (GIBCO 11140076; stock 100X), 10 mL penicillin-streptomycin (Life Technologies, 15140163; stock 104 U/mL), 333 μL BSA fraction V (GIBCO 15260037; stock 7.50%), 7 μL β-mercaptoethanol (Sigma M6250; stock 14.3 M), 100 μL LIF (Chemico, ESG1107; stock 107 U/mL), 100 μL PD0325901 (Stemgent, 04-0006-10; stock 10 mM), and 300 μL CHIR99021 (Stemgent, 04-0004-10; stock 10 mM). For confocal and PALM imaging, cells were grown on glass coverslips (Carolina Biological Supply, 633029) that had been coated with the following: 5 μg/mL of poly-L-ornithine (Sigma P4957) at 37°C for at least 30 minutes followed by 5 μg/mL of laminin (Corning, 354232) at 37°C for at least 2 hours. Cells were passaged by washing once with 1X PBS (Life Technologies, AM9625) and incubating with TrypLE (Life Technologies, 12604021) for 3–5 minutes, then quenched with serum-containing media made by the following recipe: 500 mL DMEM KO (GIBCO 10829-018), MEM nonessential amino acids (GIBCO 11140076; stock 100X), penicillin-streptomycin (Life Technologies, 15140163; stock 104 U/mL), 5 mL L-glutamine (GIBCO 25030-081; stock 100X), 4 μL β-mercaptoethanol (Sigma M6250; stock 14.3 M), 50 μL LIF (Chemico, ESG1107; stock 107 U/mL), and 75 mL of fetal bovine serum (Sigma, F4135). Cells were passaged every 2 days.

METHOD DETAILS

ChIP-seq analysis

As described in Sabari et al., 2018, ChIP-seq browser tracks for MED1, Pol II, BRD4, and OCT4 were generated by aligning reads to NCBI37/mm9 using Bowtie with the following settings: “-p 4–best -k 1 -m 1–sam -l 40.” WIG files represent counts (in reads per million, floored at 0.1) of aligned reads within 50 bp bins. Each read was extended by 200nt in the direction of the alignment.

(Source: <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112808>)

GRO-seq analysis

For generation of the GRO-seq browser tracks, GRO-seq reads were processed as described in Sigova et al. (2015). The GRO-seq .sra file corresponding to GEO accession number GEO: GSM1665566 (Sigova et al., 2015) was converted to .fastq using the SRA toolkit (Leinonen et al., 2011). Reads were aligned to the mouse genome (NCBI37/mm9) using Bowtie v1.2.2 (Langmead et al., 2009) with the following settings “-e 70 -k 1 -m 10 -n 2–best.” The reads corresponding to each one of the features (super-enhancers, typical enhancers, proximal promoter regions, genes) were counted using featureCounts v1.6.2 (Liao et al., 2014) with default settings. The coordinates for typical enhancers and super-enhancers in mouse embryonic stem cells (mESCs) were acquired from Whyte et al. (2013). The coordinates for genes (transcription start and end sites) were acquired using the UCSC Table Browser (Karolchik et al., 2004). The upstream antisense promoter regions were defined as genomic areas containing 1 kb upstream of each TSS. Their coordinates were retrieved by using BEDTools v.2.26.0 (Quinlan and Hall, 2010) and the TSS coordinates as input (to the slop function). Reads were normalized with the size of the corresponding feature they aligned to.

RNA-seq analysis

The RNA-seq .sra file corresponding to GEO accession number GEO: GSM2686137b (Chiu et al., 2018) was converted to .fastq using the SRA Toolkit RNA-seq analysis was performed using the nf-core RNA-seq pipeline (v1.4.2) (Ewels et al., 2020) with default settings and NCBI37/mm9 as reference genome. Nextflow v20.01.0 was used as a workflow tool on an LSF High-Performance Computing environment (Di Tommaso et al., 2017). STAR v2.6.1d (Dobin et al., 2013) was used for the alignment of reads. Aligned reads were assigned to the aforementioned intervals (typical enhancers, super-enhancers, proximal promoter regions and genes) by using featureCounts v1.6.4, with the default settings.

Calculation number of RNA molecules in cells

Known concentrations of *in vitro* transcribed enhancer RNAs and pre-mRNAs from *Trim28* and *Pou5f1* loci are used as standards to approximate the number of molecules in cells. These RNAs are converted to cDNAs by reverse-transcription and mixed at equal concentrations. For each RNA species, a standard curve of qRT-PCR Ct value to RNA amount was generated using serial dilutions, with two different primer sets in technical duplicates. Next, qRT-PCR reactions using the same primer sets were performed for biological duplicates of mESCs. Actb-normalized Ct values were then used to determine the amount of RNA species in the reaction based on the standard curves above. To calculate the number of RNA molecules per cell, the amount of RNA (g) was divided by the molar weight of each species ($\sim 350 \text{ (g mol}^{-1} \text{ nt}^{-1}) \times \text{length of } in vitro \text{ transcribed RNA (nt)}$), multiplied by Avogadro's number ($6.022 \times 10^{23} \text{ mol}^{-1}$), and divided by the approximate number of cells used in each reaction (10,000 cells). Melting curves were analyzed to confirm primers specificity. Non-reverse-transcribed (–RT) controls were included to rule out the amplification of genomic DNA. Primer sequences are indicated in Table S1.

In vitro droplet assay

Recombinant GFP fusion proteins were concentrated to a desired protein concentration using Amicon Ultra centrifugal filters (30K MWCO, Millipore). Droplet reactions with the recombinant proteins were performed in 10 μ l volumes in PCR tubes under the following buffer condition: 30 mM Tris HCl pH 7.4, 100 mM NaCl, 2% Glycerol and 1 mM DTT. The same buffer containing 55 mM NaCl was used for BRD4-IDR-GFP. Droplet reactions with the Mediator complex were performed under the following buffer condition: 30 mM HEPES pH 7.4, 65 mM NaCl, 2% Glycerol and 1 mM DTT. For all droplet reactions, protein and buffer were mixed first and RNA or ssDNA or heparin (Sigma, H3393) was added later. The reactions were incubated at room temperature for 1 hr without any shaking or rotating. The reactions were then individually transferred into 384 well-plate (Cellvis P384-1.5H-N) by using a micropipette (2–20 μ L) 5 minutes prior to imaging by confocal microscopy at 150X magnification or prior to turbidity measurements on a plate reader (Tecan) at 350 nm absorbance at room temperature (Banerjee et al., 2017). The concentration of proteins and RNAs in the droplet reactions are indicated in the figure legends. For brightfield Mediator experiments (Figure 1), representative images were subtracted by a median filtered image (px = 15) using ImageJ to remove camera artifacts discovered by taking images of blank wells.

Fluorescence recovery after photobleaching (FRAP)

FRAP was performed on an Andor Revolution Spinning Disk Confocal microscope with 488-nm laser. Droplets were bleached using 30% laser power with 20 μ s dwell time for 5 pulses, and images were collected every second for 60 s. Fluorescence intensity at the bleached spot and a control unbleached spot was measured using ImageJ. Values are normalized to the unbleached spot to control for photobleaching during image acquisition and then normalized to the first time point intensity.

MATLAB scripts were written to process the intensity data, and post bleach FRAP recovery data was normalized to pre-bleach intensity (FRAP(t)) and fit to:

$$\text{FRAP}(t) = M(1 - \exp(-t/\tau))$$

Where, M (mobile fraction) and τ (half-life of recovery) are inferred from fitting the data using in-built MATLAB functions. These values are inferred for each replicate and averaged to provide a range for the apparent diffusion coefficients, which is computed as:

$$D_{\text{app}} = (\text{Bleach radius})^2 / \tau$$

In vitro droplet analysis

To analyze *in vitro* droplet experiments, we used a previously reported pipeline (Guo et al., 2019). The code for this analysis is available at the Github link in the Data/Code Availability section. All droplets were segmented from average images of captured channels on various criteria: (1) an intensity threshold that was three s.d. above the mean of the image; (2) size thresholds (20 pixel minimum droplet size); and (3) a minimum circularity (circularity = $4\pi \cdot (\text{area}) / (\text{perimeter}^2)$) of 0.8 (1 being a perfect circle). After segmentation, mean intensity for each droplet was calculated while excluding pixels near the phase interface. Hundreds of droplets identified in (typically) ten independent fields of view were quantified. The mean intensity within the droplets (C-in) and in the bulk (C-out) were calculated for each channel. The partition ratio was computed as (C-in)/(C-out).

Droplet size, partition ratio, and condensed fraction measure distinct properties of droplet formation, and these three metrics show similar trends upon RNA-mediated reentrant phase transitions. When a protein or RNA is fluorescently-labeled in our experiments, we favor measuring the partition ratio. This is because the partition ratio can be measured on a per-droplet basis, and unlike condensed fraction, which varies depending on the number of droplets per field, the partition ratio is more independent of the field that is imaged.

For the size analysis of droplets formed in the reconstituted transcription assays (Figure 3), brightfield images were subtracted by a median-filtered image (px = 21), and droplets were manually segmented and their areas measured using ImageJ.

Synthesis of RNA by *in vitro* transcription

Enhancer and promoter sequences for RNAs were obtained from super-enhancer-regulated genes *Pou5f1*, *Nanog*, and *Trim28*. For promoter sequences, the first 475-490bp from the first exon were selected from mm10. For enhancer sequences, GROseq reads (Sigova et al., 2015) from both + and - strands aligned to mm9 were overlapped with called super-enhancers (Whyte et al., 2013). Contiguous regions of read density above background were manually selected (Figure S1). Primers were designed to amplify the selected promoter and enhancer sequences from genomic DNA isolated from V6.5 mESCs (Table S1). The following sequences were added to the forward and reverse primers to add the bacterial polymerase promoters:

T7 (add to 5' of sense or forward primer): 5'-TAATACGACTCACTATAGGG-3'

SP6 (add to 5' of antisense or reverse primer): 5'-ATTTAGGTGACACTATAGAA-3'

Phusion polymerase (NEB) is used to amplify the products with the bacterial promoters, and products are run on a 1% agarose gel, gel-purified using the Qiaquick Gel Extraction Kit (QIAGEN), and eluted in 40 μ L H₂O. Templates were sequenced to verify their identity. A volume of 8 μ L of each template (10-40 ng/ μ L) was transcribed using the MEGAscript T7 (Invitrogen; sense) or MEGAscript SP6 (Invitrogen; antisense) kits according to the manufacturer's instructions. For visualization of the RNA by microscopy, reactions included a Cy5-labeled UTP (Enzo LifeSciences ENZ-42506) at a ratio of 1:10 labeled UTP:unlabeled UTP. The *in vitro* transcription was incubated overnight at 37°C, then 1 μ L TURBO DNase (supplied in kit) was added, and the reaction was incubated for 15 minutes at 37°C. The MEGAclean Transcription Clean-Up Kit (Invitrogen) was used to purify the RNA following the manufacturer's instructions and eluting in 40 μ L H₂O. RNA was diluted to 2 μ M and aliquoted to limit freeze/thaw cycles, and RNA was run on 1% agarose gels in TBE buffer to verify a single band of correct size.

Recombinant protein purification

Recombinant protein purifications were performed as previously reported (Boija et al., 2018; Guo et al., 2019; Sabari et al., 2018; Shrinivas et al., 2019; Zamudio et al., 2019). pET expression plasmids containing 6xHIS tag and genes of interest or their IDRs tagged with either mEGFP or mCherry were transformed into LOBSTR cells (gift of I. Cheeseman Lab). Expression of proteins was induced by addition of 1mM IPTG either at 16°C for 18 hours or at 37°C for 5 hours. Extracts were prepared as previously described (Boija et al., 2018). Cell pellets were resuspended in 15 mL of denaturing buffer (50 mM Tris 7.5, 300 mM NaCl, 10 mM imidazole, 8M Urea) with protease inhibitors (Roche, 11873580001). After complete resuspension, the lysates were sonicated for ten cycles (15 s on and 60 s off) and subjected to centrifugation at 12,000 g for 30 minutes. The supernatant was transferred into fresh tube and the lysates were incubated with 1 mL of pre-equilibrated Ni-NTA agarose beads (Invitrogen, R901-15) with denaturing buffer at 4°C for 1.5 hours. After washing the beads with 15 volumes of the denaturing buffer, proteins were eluted with 50mM Tris pH 7.4, 500mM NaCl, 250mM imidazole buffer containing complete protease inhibitors (Roche, 11873580001). Proteins were dialyzed against 50mM Tris pH 7.4,

125 mM NaCl, 10% glycerol and 1mM DTT at 4°C for BRD4-IDR-GFP, OCT4-GFP and GFP alone and the same buffer containing 500 mM NaCl for MED1-IDR-GFP.

Purification of human Mediator complex from HeLa nuclear extract.

HeLa nuclear protein extract (4g) was prepared as described in [Dignam et al. \(1983\)](#). Nuclear extract was dialyzed against BC100: BC buffer, pH 7.5 + 100mM KCl (20 mM Tris-HCl, 20 mM β -Mercaptoethanol, 0.2 mM PMSF, 0.2 mM EDTA, 10% glycerol (v/v) and 100 mM KCl). The extract was fractionated on a phosphocellulose column (P11) with BC buffer containing 0.1, 0.3, 0.5 and 1M KCl. The Mediator complex eluted in the 0.5M KCl (BC500) fraction. This fraction was dialyzed against BC100 and loaded on a DEAE Cellulose column and sequentially fractionated with BC buffer containing 0.1, 0.3 and 0.5M KCl. The Mediator did not bind the DEAE Cellulose resin and was collected in the flow through fraction 0.1M KCl (BC100). This fraction was then directly loaded onto a DEAE-5PW column (TSK) and eluted with a linear KCl gradient from 0.1 to 1M KCl in BC buffer. The Mediator complex eluted between 0.4 and 0.6M KCl. The fractions containing Mediator were pooled and dialyzed against BD700: BD buffer, (20 mM HEPES pH 7.5, 20 mM β -Mercaptoethanol, 0.2 mM PMSF, 0.2 mM EDTA, 10% glycerol, and 700 mM (NH₄)₂SO₄). This fraction was then loaded onto a Phenyl-Sepharose Hydrophobic Interaction Chromatography (HIC) column and eluted with a linear reverse gradient from 0.7 to 0.025M (NH₄)₂SO₄ in BD buffer. The Mediator complex eluted between 0.3 and 0.1M (NH₄)₂SO₄. The Mediator-containing fractions were again pooled and dialyzed against BA100: BA buffer, pH 7.5 + 100 mM NaCl (20 mM HEPES, 20 mM β -Mercaptoethanol, 0.2 mM PMSF, 0.2 mM EDTA, 10% glycerol and 100 mM NaCl) and loaded onto a Heparin Agarose column. The column was washed with BA100 and step-eluted with BA buffer containing 0.25, 0.5, 1M and 1M NaCl. The Mediator complex eluted in the 0.5M NaCl (BA500) fraction. A portion of this fraction was then loaded on a Superose-6 (gel filtration column) that was equilibrated and run in BC100. The Mediator complex eluted from the gel filtration column with a mass range between 1-2MDa.

Reconstituted *in vitro* transcription assay

The reconstituted *in vitro* transcription by RNA polymerase II was performed as previously described ([Flores et al., 1992](#); [LeRoy et al., 2008, 2019](#); [Orphanides et al., 1998](#)) with some modifications. A 1000 bp template DNA (unlabeled or Cy-3 labeled at 3' end) containing adenovirus major late promoter, five Gal4 binding sites, TATA-box sequence and 561 bp from eGFP sequence was used. First, pre-initiation complex was assembled at RT for 15 min by mixing the following components: 50 nM RNA polymerase II enriched for hypophosphorylated CTD, 50 nM general transcription factors (TFIIA-B-D-E-F-H), and 5.75 nM template DNA, in a buffer containing 10 mM HEPES pH 7.5, 65 mM NaCl, 6.25 mM MgCl₂, and 6.25 mM Sodium butyrate. Next, 10 nM Mediator complex and 10 nM GAL4 (Gal4 DNA binding domain fused to activation domain of VP16) were added to the reaction. Last, nucleotide mix containing 0.375 mM ATP, CTP, UTP, GTP (Invitrogen), 0.01 U RNase Inhibitor (Invitrogen), 1.25% PEG-8000 were added together with one of the following: a) various amounts of purified exogenous *Pou5f1* RNA (0-500 nM) b) spermine (Sigma, S4264) c) extra NTPs (Invitrogen) d) extra NaCl e) heparin (Sigma, H3393). The reaction was incubated at 30°C for 2 hr. RNA isolation was performed using RNeasy kit (QIAGEN) by including a spike-in RNA control and an RNA carrier. Purified RNAs were treated with ezDNase (Invitrogen) for 30 min at 37°C to eliminate the template DNA. Reverse transcription was performed using Superscript IV (Invitrogen) and qPCR was performed with SYBR Green Real Time PCR master mix (Invitrogen) to quantify the template derived transcriptional output. The Ct values of the reactions were normalized to the spike-in RNA control. The concentration of template derived transcriptional output was calculated by using a standard curve of qRT-PCR Ct values generated by known amounts of serially diluted GFP RNA. The sequence of primers used for qRT-PCR are indicated in [Table S1](#).

To visualize the droplets formed in the reconstituted transcription assay, using a micropipette (2-20 μ L), 5 μ L of the reactions were loaded onto a homemade chamber, which was prepared by attaching coverslips to a glass slide by parallel strips of double-sided tape ([Sabari et al., 2018](#)). After the droplets were settled on the glass coverslip, the images were collected by using RPI Spinning Disk confocal microscope with a 100x objective. To account for camera artifacts in the images, brightfield images of droplets from reconstituted assays were subjected to a white tophat filter with a disk element radius of 21 using the MorphoLib plugin in ImageJ, then a Gaussian filter (sigma = 1) was applied.

Constructing a free-energy for RNA-protein phase behavior

Our goal in this section is to develop a simplified and coarse-grained model that captures the qualitative physics of RNA-protein mixtures. Based on phenomenological observations of transcriptional proteins and RNA ([Figure 2](#)), such a model must recapitulate the following key features:

- Transcriptional proteins phase separate in the absence of RNA through other types of interactions, albeit at higher concentrations.
- At fixed protein concentrations, addition of RNA initially promotes de-mixing and at higher levels drive a re-entry into the mixed phase.

Motivated by the evidence that transcriptional condensates recruit diverse coactivators, transcription factors, and other proteins of the transcriptional apparatus (Boija et al., 2018; Guo et al., 2019; Sabari et al., 2018; Shrinivas et al., 2019), we define an effective protein component P that lumps together different transcriptional molecules. Similarly, while different species of RNA are likely present within these condensates, we define an effective RNA species (R).

Landau model

First, we approach this problem by constructing a phenomenological free-energy with 2 order-parameters that represent scaled concentrations of protein ($\phi_p(\vec{r}, t)$) and RNA ($\phi_r(\vec{r}, t)$). We define the free-energy (normalized to $k_B T = 1$) as:

$$f[\phi_p, \phi_r] = \int_V d^d V \left(f_{dw}(\phi_p(\vec{r}, t)) + \rho_r \phi_r^2 + \chi_{eff}(\phi_p(\vec{r}, t), \phi_r(\vec{r}, t)) + \frac{\kappa}{2} (\nabla \phi_p)^2 \right)$$

Here, $f_{dw}(\phi_p(\vec{r}, t)) = \rho_s (\phi_p - \alpha)^2 (\phi_p - \beta)^2$ is a standard double-well potential that ensures protein components phase separate without RNA with co-existence concentrations specified by α, β . Choice of $\kappa > 0$ ensures that there is finite surface tension for the protein condensate. The second-order term for RNA ($\rho_r > 0$) states that within this model-framework, RNA cannot phase-separate in the absence of protein. Given that electrostatic interactions at physiological salt conditions are fairly short-ranged (Debye length ~ 1 nm), we capture the non-linear nature of RNA-protein interactions in an effective interaction term χ_{eff} . We define this interaction term in the spirit of the Landau-Ginzburg approach as an expansion in powers of the order parameters:

$$\chi_{eff}(\phi_p, \phi_r) = -\chi \phi_p \phi_r + a \phi_p \phi_r^2 + b \phi_p^2 \phi_r + c \phi_p^2 \phi_r^2 + \dots + H.O.T$$

While symmetry arguments often dictate or exclude certain types of terms (odd powers in Ising models for example) in such an expansion, there are no obvious symmetry constraints for this system. Hence, our modeling approach is to minimize the number of higher-order terms that need to be included to recapitulate the experimentally observed reentrant phase transition. Our experimental results suggest that low concentrations of RNA promote phase separation, and thus the lowest order term ($-\chi \phi_p \phi_r$, $\chi > 0$) lowers the free-energy. However, higher-order terms must counter this and below we outline how we determine which terms to include. In general, the stability of a mixture described by such a free-energy can be ascertained from the Jacobian matrix J . For our model, the elements of this 2×2 matrix are:

$$J_{pp} = \frac{\partial^2 f}{\partial \phi_p^2} = 2\rho_s \left(6\phi_p^2 - 6\phi_p(\beta + \alpha) + (\alpha - \beta)^2 \right) + 2b\phi_r + 2c\phi_r^2, J_{pr} = \frac{\partial^2 f}{\partial \phi_p \partial \phi_r} = -\chi + 2a\phi_r + 2b\phi_p + 4c\phi_p \phi_r$$

$$J_{rr} = \frac{\partial^2 f}{\partial \phi_r^2} = 2\rho_r + 2a\phi_p + 2c\phi_p^2$$

The mixed phase is no longer stable to perturbations when at least one eigen value of J becomes negative (spinodal instability). In the absence of RNA, the spinodal satisfies $J_{pp} = 0$. If only the pairwise interaction terms were considered ($-\chi \phi_p \phi_r$), the spinodal region broadens i.e., phase separation is promoted at lower protein concentrations when RNA is present. We next characterized the effect of an additional higher-order term (only one of a, b or c is non-zero) on the Jacobian matrix. We ascertained that:

- $a > 0$: While the free-energy is dominated by repulsive interactions at higher RNA concentrations, the Jacobian matrix predicts a continuous underlying instability. Instead of suppressing phase separation at higher RNA concentrations and promoting re-entry to dilute phase, this term would instead change the composition of the demixed phases.
- $b > 0$: While this term promotes a reentrant behavior, the resulting regions of instability demix RNA away from protein for most values of b .
- $c > 0$: For values of c that are not too large (i.e., $c < \approx \rho_r$), the resulting phase diagram mirrors a reentrant shape with RNA enrichment in the protein condensate. If c is larger, then a second de-mixing transition (similar to case 2 i.e., $b > 0$) is observed at high values of ϕ_p, ϕ_r . Since we are interested in the limit of relatively low protein/RNA concentrations, and the values of ϕ_p, ϕ_r represent qualitative proxies of protein/RNA concentrations, we choose to explore our model in this parameter regime.

While cubic and higher-order terms are required to recapitulate complete phase-behavior, we explored our model with $c > 0$, assuming the coefficients a, b are small. In the simulations reported in Figures 4, 5, and 6, the free-energy parameters are $\alpha = 0.1, \beta = 0.7, \chi = 1.0, c = 10.0, \kappa = 0.5, \rho_s = 1.0, \rho_r = 10.0, a = b = 0$. All free-energy calculations were performed with *Python* and code is available at:

https://github.com/krishna-shrinivas/2020_Henninger_Oksuz_Shrinivas_RNA_feedback.

Flory-Huggins model

In this approach, rather than employ a phenomenological model, we parametrize a microscopic model motivated by Flory-Huggins polymer-solution theory (Flory, 1942). The simplified F-H model contains 3 components - protein, RNA, and the solvent (s), whose volume fractions are defined as $\phi_p(\vec{r}, t)$, $\phi_r(\vec{r}, t)$, $1 - \phi_p(\vec{r}, t) - \phi_r(\vec{r}, t)$ respectively. The free-energy (normalized as before) is defined as:

$$f = \sum_i \frac{\phi_i}{r_i} \log(\phi_i) + \sum_{i,j>i} \chi_{ij} \phi_i \phi_j$$

Here, r_i are the solvent-equivalent polymerization lengths of the RNA & protein (assumed to be equal for simplicity) and χ_{ij} are the various pairwise interaction terms. As before, we assume these interactions to be short-ranged at physiological salt levels. Choice of $(-\chi_{pr}) > \chi_{ps} > 0$ and $\chi_{rs} < 0$ recapitulate the attractive contributions of protein-protein/protein-RNA interactions and repulsive RNA-RNA interactions. With these choices of constraints, the resulting free-energy looks similar to the phase diagram from the Landau approach with $c > 0$ (Figures S6A and S6B) where the key F-H parameters are $(-\chi_{pr}) = 1.1, \chi_{ps} = 0.75, \chi_{rs} = -0.6$, and $r_p = r_r = 30$.

Numerical phase-field simulations

Numerical investigations of the coupled-equations outlined in Figure 4C were performed with the FiPy package (Guyer et al., 2009). Simulations were performed on a 2-D/3-D square lattice ($L_x = L_y = 200, dx = 0.3; L_x = L_y = L_z = 40, dx = 1.0$) and with adaptive time-stepping ($dt_{min} = 1e-8, dt_{max} = 5e-1$) until steady state is reached (which typically requires ~ 10000 simulation steps).

The chemical potential for the protein components is calculated as:

$$\mu_p = \frac{df}{d\phi_p} = 2\rho_s(\phi_p - \alpha)(\phi_p - \beta)(2\phi_p - \alpha - \beta) - \kappa \nabla^2 \phi_p - \chi \phi_r + 2c\phi_r^2 \phi_p$$

The radius of condensates was inferred from the volume of mesh regions where $\phi_p \geq (\alpha + \beta/2)$. The mobility of RNA and protein were chosen to be 1.0 unless mentioned elsewhere. The raw data for all figures from simulation data are provided along with the manuscript.

Design of Simulations to vary RNA features and rates of RNA synthesis

We designed simulations (Figure 6E) to study the effect of RNA features and rates of effective synthesis on condensate size. The rates of synthesis were changed by increasing k_p by multiplicative factors (see x axis in Figure 6E). Since RNA length is not explicitly incorporated in the model framework, we defined the effective local synthesis rates of longer RNA as a product of k_p and an additional multiplicative factor (1, 2, and 4x for short, medium, and long RNA respectively) to mimic increased local concentrations of RNA.

Calculation of number of charged molecules in condensates

In estimating the number and charge of transcriptional proteins (Figure S1), we use previous estimates (Cho et al., 2018) that suggest key transcriptional proteins such as Mediator are present at 10-100 molecules in transcriptional condensates. Further, molecules such as MED1 or BRD4 contain large disordered domains with net positive charge of +5 to +40. This provides a highly approximate estimate of 25-500 as the effective positive charge. Since there are many more transcriptional proteins and most proteins tend to contain net positive charges, it is likely that this estimate represents lower bounds on the range. Steady-state levels of nascent eRNA (Figure S1) suggest a range of 0.2-10 molecules, and since super-enhancers typically contain clusters of such active enhancers, we approximate the typical range of eRNA molecules at a transcriptional condensate between 1-10. Since RNA carries a charge of around -1 per nt (Banerjee et al., 2017) and eRNAs are short (< 1 kb, ~ 100 nt), we estimate the effective negative charge during initiation to be in the range 100-1000. During productive elongation, mRNAs are produced in bursts ranging from few to tens (1-50) and are typically longer (> 1 kb, $\sim 10,000$ nt), suggesting a conservative estimate of the effective charge to range from (10,000-1,000,000). It is important to stress that our approximations are performed with the aim of obtaining order-of-magnitude estimates and do not account for factors such as local composition of different proteins or extent to which nascent mRNAs may be coated by RNA-binding proteins. With the above numbers, we estimate concentrations based on a typical transcriptional condensate of size $r = 0.25 \mu\text{m}$ (Cho et al., 2018) that suggests that eRNA concentrations range about 10-200 nM and transcriptional proteins range 1-20 μM within the condensate.

Reactive/diffusive time-scales and estimates in cells

As defined in the model (Figure 4B), the key rates of synthesis/degradation reactions are k_p/k_d , which have units of s^{-1} , and thus the relevant time-scales are $t_r = k_p^{-1}$ (or k_d^{-1}). Timescales of RNA transport depend on both the diffusivity as well as the size of the condensate (L) and is defined as $t_d = L^2/M_{ma}$. We approximated the range of diffusivity of the nascent transcript at the lower end by the diffusivity of of chromatin, which ranges from $10^{-3.5} - 10^{-2} \mu\text{m}^2/\text{s}$ (Gu et al., 2018) and on the higher end by those of freely diffusing mRNPs, which can be up to $5 \times 10^{-2} \mu\text{m}^2/\text{s}$ (Niewidok et al., 2018). By assuming a typical eRNA of size 100nt and Pol II transcription rates as $\sim 20 - 70 \text{ nt/s}$ (Maiuri et al., 2011) we inferred typical synthesis rates of $\sim 0.5 \text{ eRNA s}^{-1} \text{ Pol II}^{-1}$. In our previous work (Cho et al., 2018), we have seen that clusters that contain multiple polymerases (> 5), are typically around $r \approx 200 - 400 \text{ nm}$. Since

super-enhancers typically contain clusters of enhancers with multiple sites of eRNA synthesis (~ 5), this gives an effective synthesis rate of $k_p \approx 2.5 \text{ s}^{-1} \text{ Pol II}^{-1}$. This allows us to approximately obtain the ratio of diffusive and reactive time-scales as $t_d/t_r = kr^2/M \approx 2 - 1000$ over the range of parameters including diffusivity and radii of cluster.

Calculation of charge balance

Charge-balance calculations were performed (Figures 2 and S2–S4) employing the following method. Net protein charge per molecule was calculated as $C_p = \#(R, K) - \#(D, E)$ for the relevant sequence including the GFP tag. RNA charge per molecule was calculated as $C_r = -(\# \text{ of bp})$, assuming an approximate charge of -1 per nucleotide (Lin et al., 2019). Next, the charge balance ratio was computed at a particular RNA and protein concentration as:

$$\text{Charge - balance ratio} = \frac{\min(C_p[P], C_r[R])}{\max(C_p[P], C_r[R])}$$

The effective concentration of MED1-IDR in our assays was 1000 nM . Our results were not quantitatively affected by inclusion/exclusion of the partial charge on Histidine residues, partly due to their low frequency on the protein sequences. For Heparin, a charge of roughly -3 per monomer was employed (Lin et al., 2020) and for single-stranded DNA, a charge of -1 per nt was employed. A comprehensive listing of charges of various species employed in this study are provided in Table S2. The Pearson correlation coefficient (r) was calculated between the median droplet partition value at different concentrations and the relevant charge-balance ratios and reported in Figures 2 and S2–S4. A higher correlation implies that experimental data follow a similar qualitative trend as the estimated charge-balance curves. The code for performing these calculations are available at:

https://github.com/krishna-shrinivas/2020_Henninger_Oksuz_Shrinivas_RNA_feedback.

Transcription inhibition by small molecules

For small molecule inhibition experiments, cells were treated with $100 \mu\text{M}$ DRB (Sigma), or $1 \mu\text{M}$ Actinomycin-D (Sigma) in 2i media (detailed above) for 30 minutes, then imaged. For wash-out experiments, media was replaced with fresh 2i media and cells were allowed to recover for 1 hour, then the cells were imaged.

Condensate size

Cells with endogenously-tagged Med1-GFP (Sabari et al., 2018) were plated on glass-bottom dishes (Mattek) coated with poly-L-ornithine (Sigma) and laminin (ThermoFisher). Mock (DMSO) and treated cells were imaged on a LSM 880 Confocal Microscope with Airyscan to obtain super-resolution z stacks for at least 8 different fields containing multiple cells. For quantification, a manual threshold was applied equally across all conditions to remove background, and the size of Med1-GFP puncta was quantified in 3D using the 3D object counter plugin (Fiji/ImageJ).

Condensate lifetime

HaloTag was endogenously knocked into 5' end of Med19 via homology-directed repair (HDR) in mouse embryonic stem cells (R1 mESCs). Three single-guide RNAs (sgRNAs) targeting ± 100 bps from the start codons of Med19 gene were designed using the web-based CRISPR Design tool (<https://zlab.bio/guide-design-resources>) and integrated into a *Streptococcus pyogenes* Cas9 vector (Addgene #62988) for standard CRISPR/Cas9 editing. Single positive colonies were sorted by fluorescence-activated cell sorting (FACS) and validated under the microscope.

Cells were cultured in serum-free 2i medium on poly-L-ornithine (PLO) and Laminin-coated flasks for more than two days and then were transferred onto coated imaging dishes for another day. Before imaging, cells were stained with (PA)-JF549-HaloTAG dye (a gift from Luke Lavis Lab, Janelia Research Campus) of 100 nM concentration for 2 hours followed by a 60-minute wash in fresh 2i medium. Lastly, dishes were filled in with 2 ml Leibovitz's L-15 Medium (no phenol red, Thermo Fisher) and brought to the microscope for imaging.

Photo-activation localization microscopy (PALM) imaging was performed using a Nikon Eclipse Ti microscope with a $100\times$ oil immersion objective (NA 1.40) (Nikon, Tokyo, Japan). A 405 nm beam of 100 mW power (attenuated with 25% AOTF) and a 561 nm beam of 500 mW power were collimated and superposed to perform simultaneous activation and excitation. The combined beam was expanded and re-collimated with an achromatic beam expander (AC254-040-A and AC508-300-A, THORLABS) to improve the uniformity of illumination across the whole region of interest (ROI 2562 pixels). Images were acquired with an Andor iXon Ultra 897 EMCCD camera (gain 1000, exposure time 50ms) interfaced through Micro Manager 1.4. 2400 frames were acquired for each imaging cycle. The cells were maintained at 37°C in a temperature-controlled platform (InVivo Scientific, St. Louis, MO) on the microscope stage during image acquisition. Med19-Halo cluster lifetimes were calculated as previously described using the qSR software (dark time tolerance = 20 frames, min cluster size = 50) (Andrews et al., 2018), and a cumulative distribution was generated using Prism software (GraphPad).

Nascent RNA imaging

For the nascent RNA experiments in Figure S7A-7C, 1.25×10^5 wild-type mESCs were plated on coverslips coated with poly-L-ornithine (Sigma) and Laminin (ThermoFisher). After overnight plating, nascent RNA labeling with 2.5 mM EU was done with the Click-iT RNA Alexa Fluor 594 Imaging Kit (ThermoFisher) according to manufacturer instructions for 10 minutes. After incubation, cells were immediately fixed with 4% paraformaldehyde for 10 minutes, washed 3X with PBS, then permeabilized with 0.5% Triton X-100 in PBS for 15 minutes. After the Click-iT reaction, coverslips were blocked with 4% RNase-free BSA in PBS for 10 minutes at room temperature. Coverslips were incubated with primary antibodies (1:500; rabbit Abcam ab64965 for MED1 and rat Millipore Sigma 04-1571 for Pol II-S2) in 4% BSA/PBS at room temperature overnight. The next day, coverslips were washed 3X with PBS, then incubated in secondary antibody (1:500; goat anti-rabbit AlexaFluor-488 ThermoFisher A11008, goat anti-rat AlexaFluor-647 Invitrogen A21247) for 1 hour at room temperature. After washing 3X with PBS, coverslips were stained with 1:1000 Hoechst 33342 in PBS, incubated for 15 minutes at room temperature, washed 3X with PBS, and mounted on imaging slides with Vectashield Mounting Media. Images were collected on the RPI Spinning Disk confocal. Representative images in Figure S7A are single z-planes of median-subtracted ($px = 10$) and Gaussian smoothed ($\sigma = 1$) channels to correct for uneven illumination and background.

For analysis of these images, nuclei were segmented using the Cellpose algorithm (Stringer et al., 2020) on the 405 Hoechst channel images. For average image analysis in Figure S7B, all channel images were maximally projected, subtracted by median filter ($px = 10$), and Gaussian smoothed ($\sigma = 1$). The center of MED1 and Pol II-S2 puncta were segmented as follows. The Laplace of Gaussian transformation ($\sigma = 3$) was applied to the images using the scikit-image package in python, and puncta were identified above a threshold intensity 3 standard deviations above the mean of the image. All spots were confirmed to be in nuclei. A $1 \mu\text{m}$ by $1 \mu\text{m}$ box was centered on the spots, and the box subimage was collected for that region in both the processed MED1 and Pol II-S2 channel images. These subimages from > 10 imaged fields were stacked and averaged, which was the input for the contour plots in Figure S7B. Radial intensity plots in Figure S7C show the distribution of these averaged signals as a function of the distance from the center of the spot, along with their correlation to EU RNA signal.

Reporter assay to determine the effect of local RNA synthesis on transcription

Vectors used in the reporter assay are modified from pTETRIS-cargo vector, gift from J. M. Calabrese (Kirk et al., 2018). 6X STOP codon sequence was cloned into NotI digested pTETRIS-cargo vector using Gibson cloning strategy by following the manufacturer's instructions (NEB). This vector is called pTETRIS-cargo-STOP. The feedback gene and the reporter gene have their own polyA termination signal (200-300 bp) to terminate transcription. There is 51 bp between these two polyA signals that are facing each other. The reporter gene is regulated by a phosphoglycerate kinase (PGK) promoter. Various versions of the pTETRIS-cargo-STOP using Gibson cloning strategy (NEB): i) the relative orientations of the feedback RNA and luciferase reporter were altered (tandem or divergent orientations) ii) feedback RNAs and luciferase reporter were cloned into separate vectors. Using Gibson cloning strategy (NEB), various RNA sequences were cloned downstream of the 6X STOP sequence to prevent translation of these RNAs. Stable cell lines for individual RNAs were generated by transfecting Med1-GFP mESCs with the following vectors: $1.0 \mu\text{g}$ pTETRIS-cargo-STOP containing individual RNAs, $1.0 \mu\text{g}$ rTTA-cargo, gift from J. M. Calabrese (Kirk et al., 2018), and $1 \mu\text{g}$ piggyBAC transposase (Systems Biosciences). Cells were selected on puromycin ($2 \mu\text{g}/\text{ml}$) and G418 ($200 \mu\text{g}/\text{ml}$) for 1 week for successful integrations. For luciferase assays, 1×10^5 cells of each genotype were plated in triplicate on 0.2% gelatin-coated 24-well plates and allowed to settle overnight. Cells were treated with doxycycline (Sigma) and harvested after 24 h to measure either luciferase activity or to purify RNA. Luciferase activity was measured using the Luciferase Assay System (Promega) according to manufacturer instructions. Luciferase signal was normalized to total protein content, measured by BCA protein assay kit (Invitrogen, #23227), and then normalized to a control not treated with doxycycline. To measure RNA expression, RNA was purified using the QIAGEN RNeasy Mini kit (QIAGEN) according to manufacturer instructions, cDNA was generated by Superscript III (Invitrogen) according to manufacturer instructions, and 10 ng of cDNA was used in a qRT-PCR SYBR-green reaction (Life Technologies) with primers specific to a common sequence shared across the vectors (qPCR_Tetris; Table S1). Ct values were normalized to a housekeeping gene (qPCR_mActb; Table S1) and a control condition with no doxycycline treatment.

For the washout experiments in Figure S7J, reporter cells were plated as described above. After 24 hours of dox treatment, media was replaced with fresh media, whereas control cell media was replaced with dox-containing media. After an additional 24 hours, luciferase levels were measured as described above. For the antisense oligo experiments of Figure S7J, antisense oligos (LNA gapmers, QIAGEN) were designed using the QIAGEN GeneGlobe tool against the feedback RNA. A negative scrambled control was also included. Reporter cells were plated as described above in triplicate, and cells were transfected with 25 nM ASO with Lipofectamine-3000 (and no P3000 enhancer agent). After overnight transfection, cell media was replaced with dox-containing or fresh 2i media as a control. After 24 hour dox treatment, RNA and luciferase levels were quantified by qRT-PCR and luminescence, respectively, as described above. For the analysis of luciferase rescue, luminescence values of the dox conditions were first normalized to the no dox condition for that ASO, and then normalized to the dox condition of the negative scrambled control.

For imaging experiments in Figures 6C and 6D, the reporter construct was modified using Gibson cloning to include a 24X-MS2 hairpin (Cho et al., 2018) at the 5' end of the RNA sequence (2,456 nt total). Cell lines with this construct and double MS2 capsid protein fused to an mCherry tag (2xMCP-mCherry) were generated as detailed above in a mESC background with endogenously-tagged Med1-GFP (Sabari et al., 2018). 1×10^6 reporter cells were plated on glass-bottom dishes (Mattek) coated with poly-L-ornithine (Sigma) and Laminin (ThermoFisher) After overnight plating, cells were treated with 10, 100, or 1000 ng/mL doxycycline for 24 hours.

Cells were imaged on an RPI Spinning Disk Confocal with the following laser powers and exposure times: 488 70% 500 ms, 561 40% 300 ms. Images were maximum projected, median subtracted ($px = 10$), and Gaussian filtered ($\sigma = 1$) to correct for uneven illumination and background subtraction. For analysis of these images in Figure 6D, nuclei were segmented using the Cellpose algorithm (Stringer et al., 2020) on images from the 561 channel that had been subjected to a maximum and median filter ($px = 10$). For average image analysis, both the RNA and MED1-GFP channel images were maximally projected, subtracted by median filter ($px = 10$), and Gaussian smoothed ($\sigma = 1$). The centers of RNA spots in a maximum projection of the 561 channel were manually marked using ImageJ. All spots were confirmed to be in nuclei. A $1\mu\text{m}$ by $1\mu\text{m}$ box was centered on the RNA spot, and the box subimage was collected for that region in both the processed RNA and MED1-GFP channel images. These subimages from > 10 imaged fields were stacked and averaged, which was the input for the contour plots in Figure 6D. Radial intensity plots in Figure S7D show the distribution of these averaged signals as a function of the distance from the center of the spot. To control for global Dox effects, we quantified the size, number, and partition ratio of MED1-GFP condensates in all conditions by using a threshold of 3 standard deviations above the mean intensity of the image to segment condensates. Partition ratio for each condensate was calculated as the average intensity inside the condensate divided by the average intensity of the nucleoplasm.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis of *in vitro* condensate assays:

The Pearson correlation coefficient (r) was calculated between the median droplet partition value at different concentrations and the relevant charge-balance ratios (see Calculation of charge balance under Method Details). The computed values are reported in Figures 2 and S2–S4. In Figure S2A, the correlation coefficient was computed between the median droplet turbidity at different conditions and the relevant charge-balance ratios. A higher correlation implies that experimental data follow a similar qualitative trend as the estimated charge-balance curves. The code for performing these calculations are available at:

https://github.com/krishna-shrinivas/2020_Henninger_Oksuz_Shrinivas_RNA_feedback.

Statistical analysis of *in vitro* transcription assays:

The Student's t test was used to determine whether the addition of spermine to the *in vitro* transcription assays led to statistically different outcomes in the mean of droplet size, partition ratio, and normalized transcription and the corresponding p values are reported in the figure legends (Figures 3D–3F). In the normalized transcription assay (Figure 3F), the values are normalized to the mean of the no spermine condition. The one-way ANOVA test is used to determine whether addition of exogenous RNA or absence of all NTPs leads to statistically different outcomes in droplet area or normalized transcription and the corresponding pairwise p values (compared to control conditions i.e., no addition of exogeneous RNA) are reported in the figure legends (Figures 3H and 3I). These tests were performed using PRISM.

Statistical analysis of transcription inhibition experiments

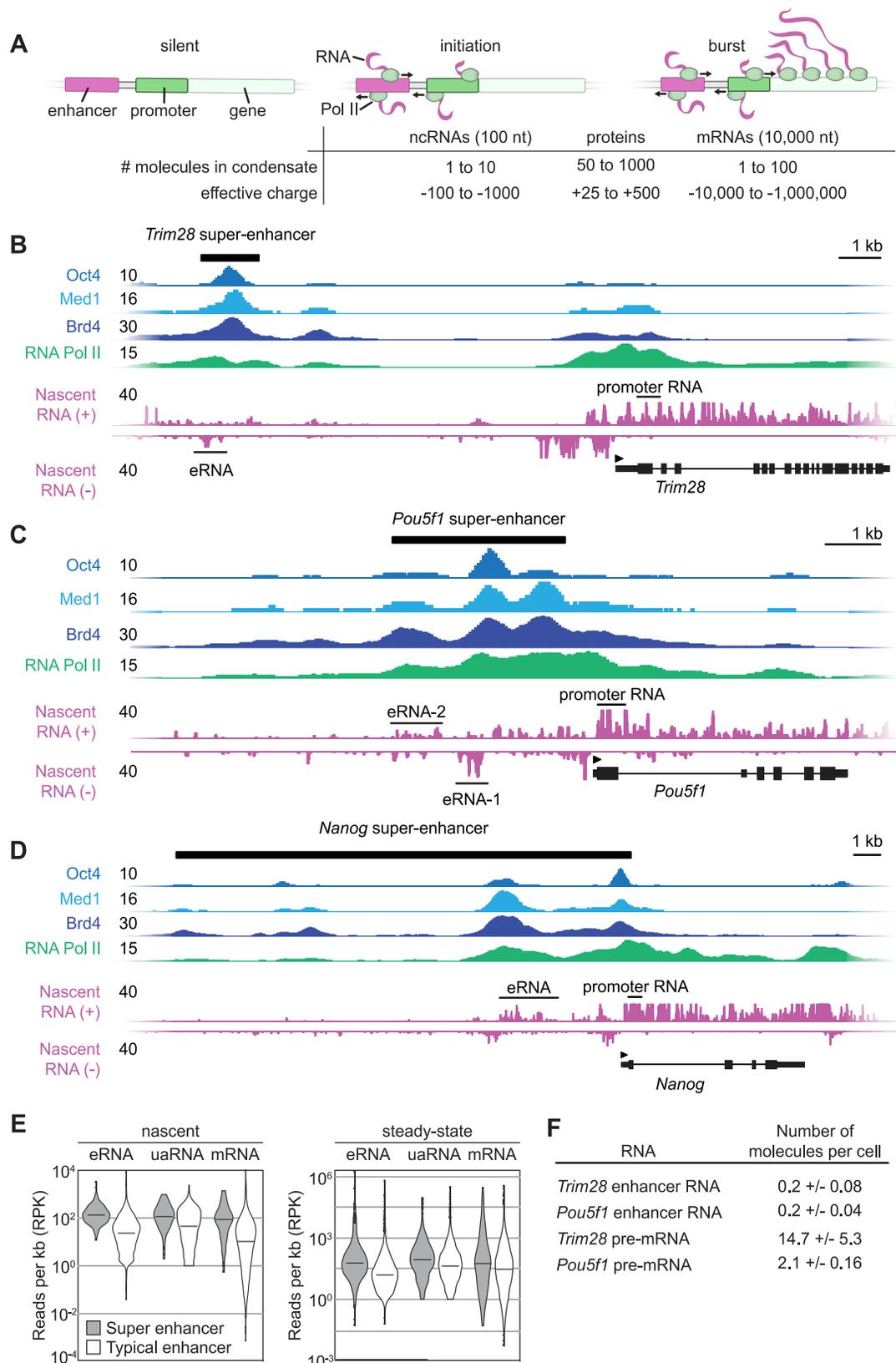
PRISM was used to compute a one-way ANOVA comparison to test whether the mean of the condensate volumes (see Method Details) was statistically different upon inhibition by DRB or ACT-D versus control (DMSO) and the corresponding p value is reported in the Figure legend (Figure 5E). Similarly, a one-way ANOVA comparison was used to test whether the mean of the condensate lifetimes (see Method Details) were statistically different from control conditions (DMSO) and the corresponding p values are reported in the Figure Legend (Figure 5I).

Statistical analysis of luciferase reporter and EU experiments:

A 2-way Kolmogorov-Smirnoff test was used to determine whether the cumulative distribution functions of the average MED1 intensity centered at RNA (see Method Details) arising from replicates (values of n are represented in Figure 6D) at different Dox concentrations were the same distribution. The p values are reported in the Figure legend (Figure 6D). Markers show the mean of at least 3 replicates and error bars depict the SD in Figures 6F, 6G, and S7. The Pearson correlation coefficient between the radial intensity distributions, computed from the averaged signal analysis (see Figure S7B; Method Details) was computed and this value is reported in Figure S7C.

No methods were used to determine whether the data met assumptions of the statistical approach.

Supplemental Figures



(legend on next page)

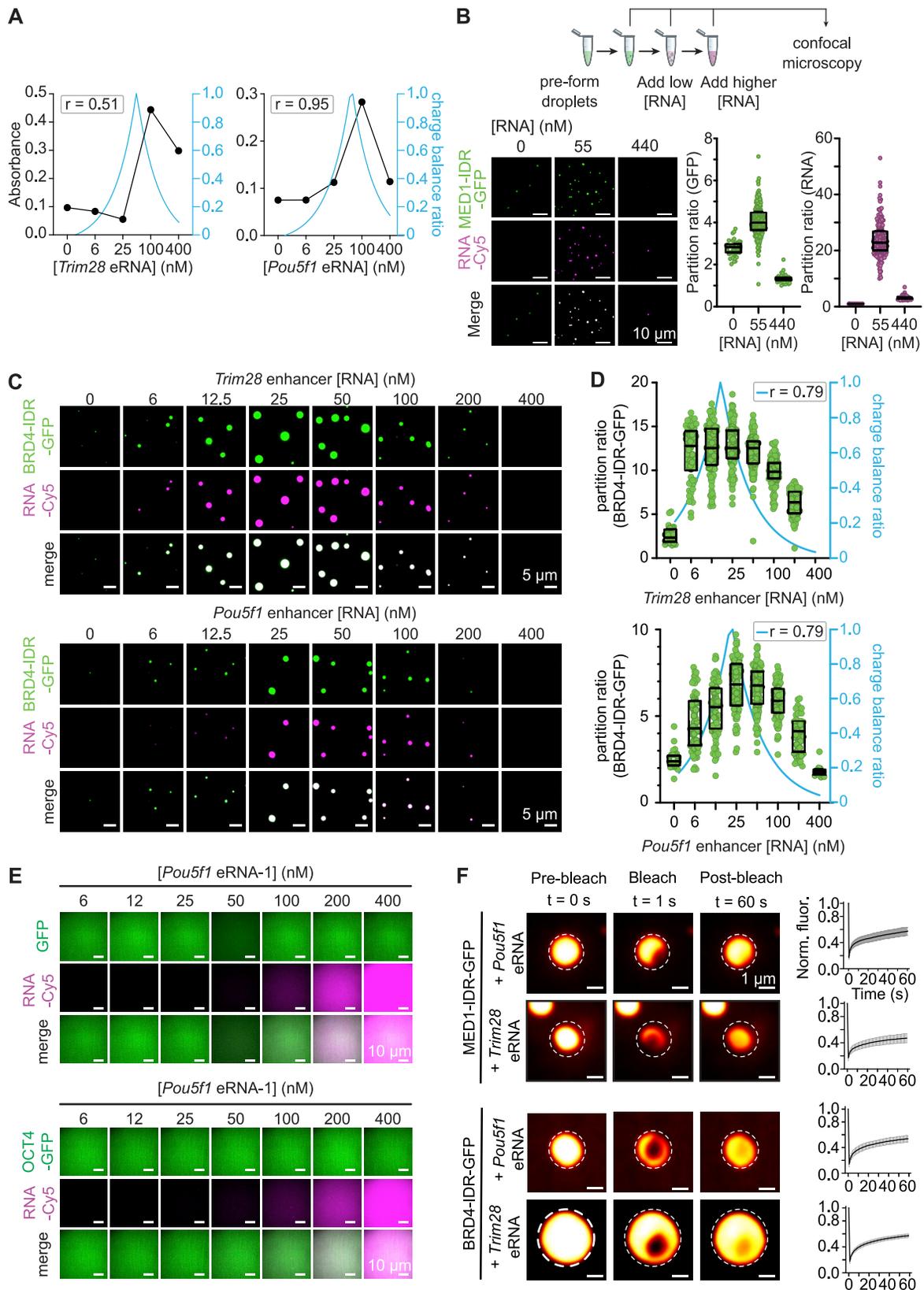
Figure S1. Transcription Machinery and RNA at Active Genes in mESCs, Related to Figure 1

A. A scheme of transcription states and the number of molecules and their corresponding effective charge in a typical transcriptional condensate during initiation and bursts of transcription ([STAR Methods](#))

B. C. D. Enrichment of transcription machinery and RNA at *Trim28* (A), *Pou5f1* (B) and *Nanog* (C) super-enhancers in mESCs. Gene tracks of ChIP-seq and nascent RNA-seq data at the indicated super-enhancers are shown. The enhancer- and promoter-derived (sense) RNAs that are used in this study are annotated in the gene tracks.

E. Nascent (left) or steady-state (right) levels of indicated RNAs at super and typical enhancers (eRNA = enhancer RNAs, uaRNA = upstream antisense promoter-associated RNAs, mRNA = messenger RNA).

F. Quantification of the number of enhancer RNA and pre-mRNA molecules in cells. Calculations are based on two biological replicates ([STAR Methods](#)).



(legend on next page)

Figure S2. Characterization of the Effect of RNA on Droplet Formation and Dissolution, Related to Figure 2

A. Turbidity measurements of droplets formed with MED1-IDR-GFP and indicated RNAs. Correlation between partition ratio and charge balance is determined by Pearson correlation (r).

B. Experimental design to test the effect of RNA on pre-formed MED1-IDR droplets (top). Representative images of MED1-IDR droplets and quantification of partition ratio of protein and RNA (bottom). Indicated concentrations of RNA were added after formation of droplets with 1 μ M of MED1-IDR (scale bars = 10 μ m).

C. Representative images of BRD4-IDR droplets at various RNA concentrations (scale bars = 5 μ m).

D. Quantification of BRD4-IDR partition ratio from (C) and correlation with charge balance (blue lines). Correlation between partition ratio and charge balance is determined by Pearson correlation (r).

E. Purified GFP (top) or OCT4-GFP (bottom) was incubated with an enhancer RNA from the *Pou5f1* locus. Whereas this RNA could stimulate MED1-IDR-GFP condensate formation, it was unable to form droplets with GFP alone or OCT4-GFP. Images were adjusted to show signal and lack of droplet formation (scale bars = 10 μ m).

F. FRAP analysis of droplets formed with MED1-IDR and RNA (top) or BRD4-IDR and RNA (bottom) (scale bars = 1 μ m).

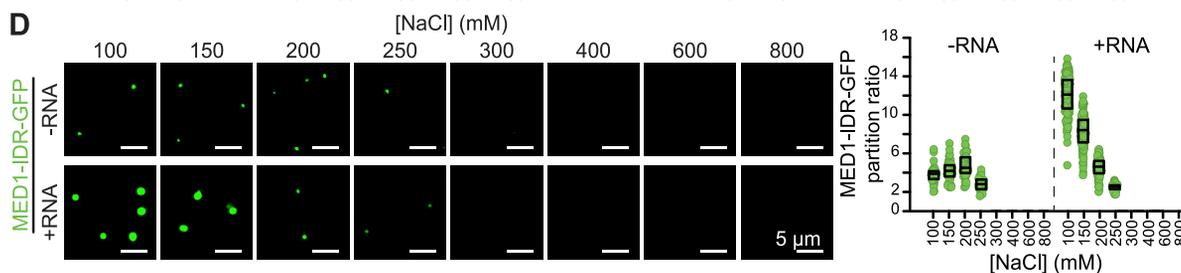
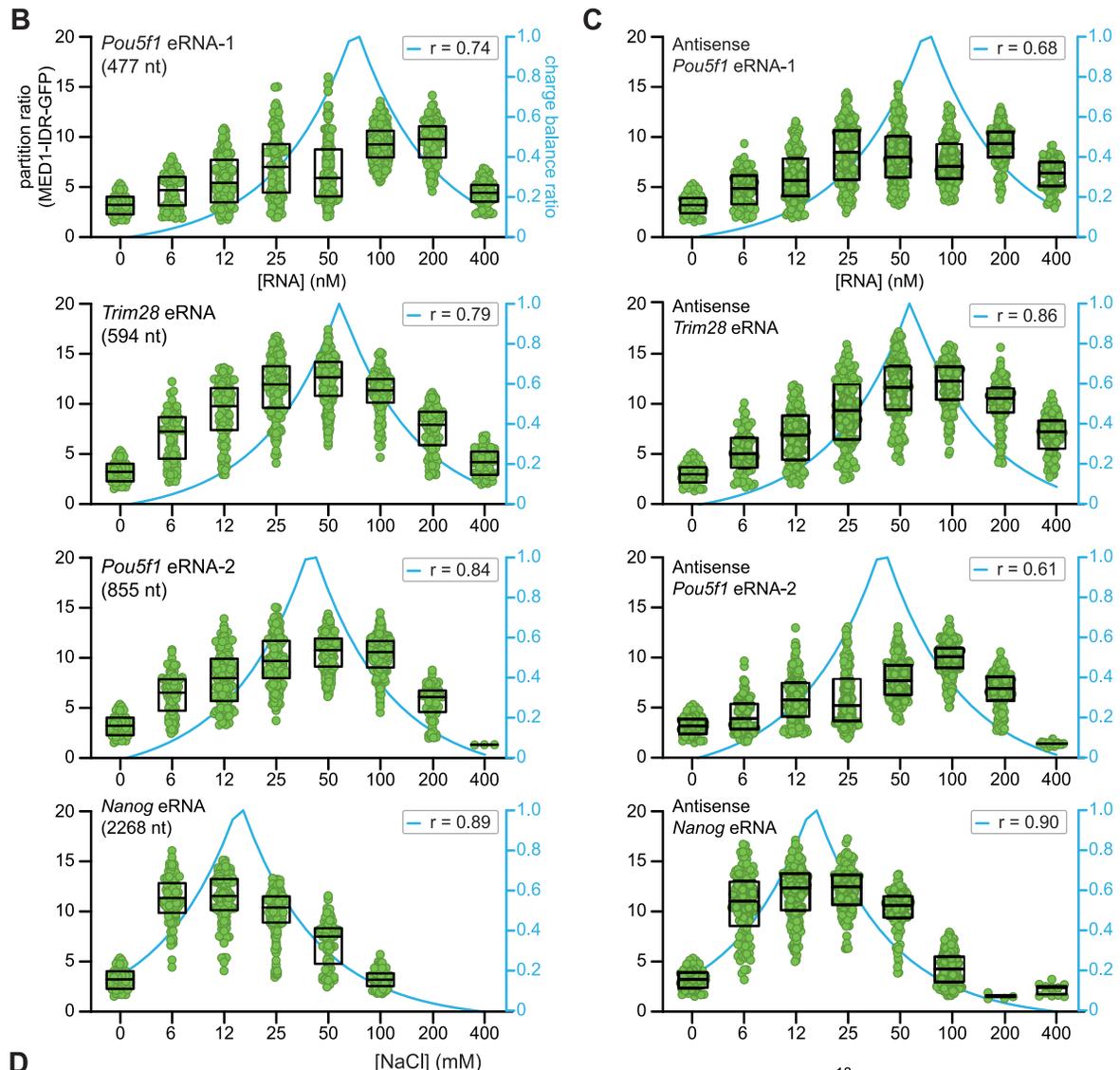
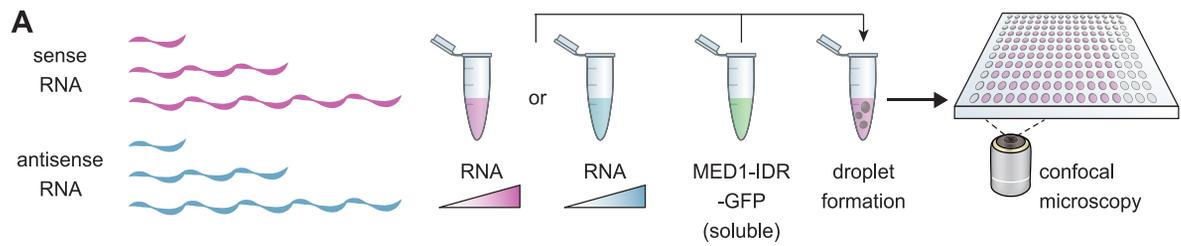
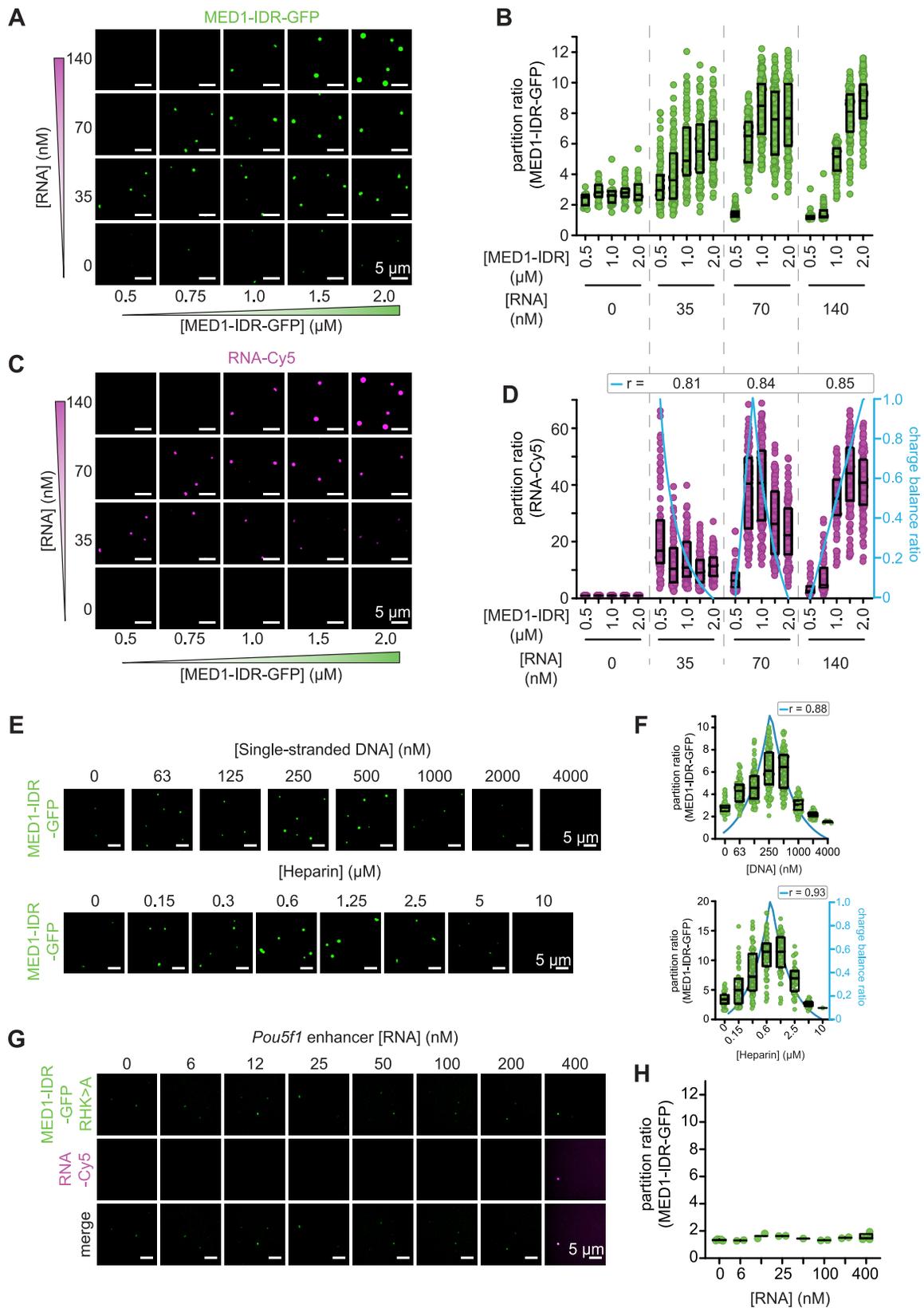


Figure S3. Modulation of Charge Balance between MED1-IDR and RNA Contributes to Stimulation and Dissolution of Condensates, Related to Figure 2

- A. Experimental design for testing diverse sense and antisense RNAs of different lengths on formation of MED1-IDR-GFP droplets.
- B. Quantification of the partition ratios of MED1-IDR-GFP within the droplets when incubated with RNAs of different lengths and sequences. Correlation between partition ratio and charge balance is determined by Pearson correlation (r).
- C. Quantification of the partition ratios of MED1-IDR-GFP within the droplets when incubated with antisense versions of the RNAs in (B). Correlation between partition ratio and charge balance is determined by Pearson correlation (r).
- D. Representative images of MED1-IDR droplets (left), which are formed with or without RNA and are subjected to increasing concentration of monovalent salt (NaCl). Quantification of partition ratios of MED1-IDR-GFP within the droplets are indicated (right) (scale bars = 5 μm).



(legend on next page)

Figure S4. Formation and Dissolution of MED1-IDR Droplets through Electrostatic Interactions, Related to Figure 2

- A. Representative images for MED1-IDR-GFP in a two-component phase diagram for MED1-IDR and *Pou5f1* eRNA (scale bars = 5 μm).
- B. Quantification of MED1-IDR partition ratio for (A).
- C. Representative images for RNA-Cy5 in a two-component phase diagram for MED1-IDR and *Pou5f1* eRNA in (A) (scale bars = 5 μm).
- D. Quantification of RNA-Cy5 partition ratio and charge balance ratio for (C). Correlation between partition ratio and charge balance is determined by Pearson correlation (r).
- E. Representative images of droplets formed with single-stranded DNA (top) or heparin (bottom) (scale bars = 5 μm).
- F. Quantification of MED1-IDR partition ratio for images in (E). Correlation between partition ratio and charge balance is determined by Pearson correlation (r).
- G. Representative images of droplets formed with MED1-IDR RHK > A and *Pou5f1* enhancer RNA (scale bars = 5 μm).
- H. Quantification of MED1-IDR partition ratio for images in (G).

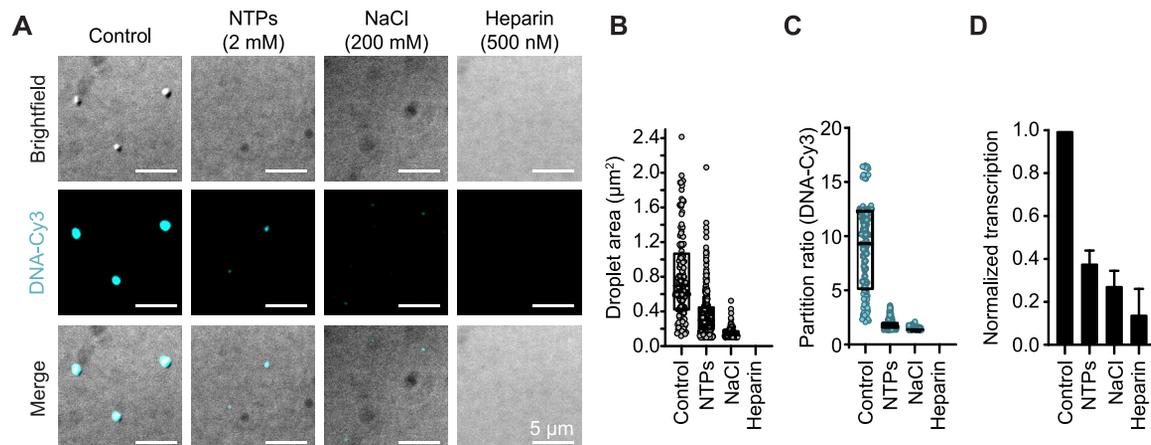


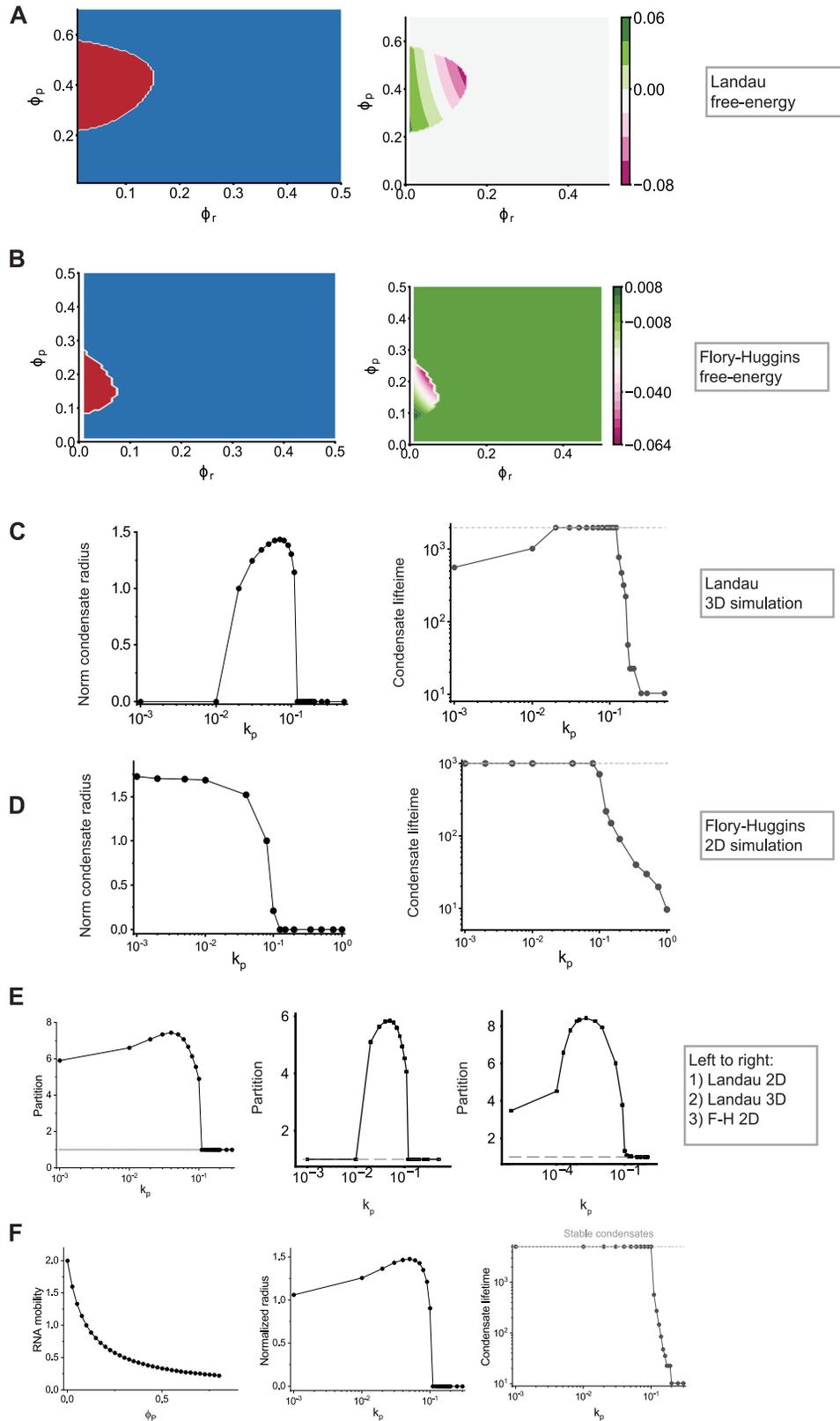
Figure S5. Charged Interactions in Reconstituted Mammalian Transcription and Droplets, Related to Figure 3

A. Representative images of transcription reactions with addition of 2mM NTPs, 200 mM NaCl, and 500 nM Heparin (scale bars = 5 µm).

B. Quantification of droplet area of droplets in (A).

C. Quantification of DNA-Cy3 partitioning from droplets in (A).

D. qRT-PCR measurement of template-derived RNA synthesis in the reconstituted transcription reactions from (A).



(legend on next page)

Figure S6. Computational Model of Non-equilibrium RNA Feedback on Transcriptional Condensates, Related to Figure 4

A. Regions where mixtures of protein and RNA phase separate spontaneously (red, left panel) are calculated from the Landau free-energy (Figure 4C) by analyzing the Jacobian (spinodal analyses; STAR Methods). As expected from the re-entry transition, increasing RNA concentration (abscissa) at fixed protein levels can start from a region promoting phase separation, and beyond a threshold, drive re-entry into dilute phase. The right panel shows the initial direction of the instability (STAR Methods), which indicates the RNA is enriched in protein condensates (value > 0, green shade), while at higher concentrations, RNA de-densifies the condensed phase (value < 0).

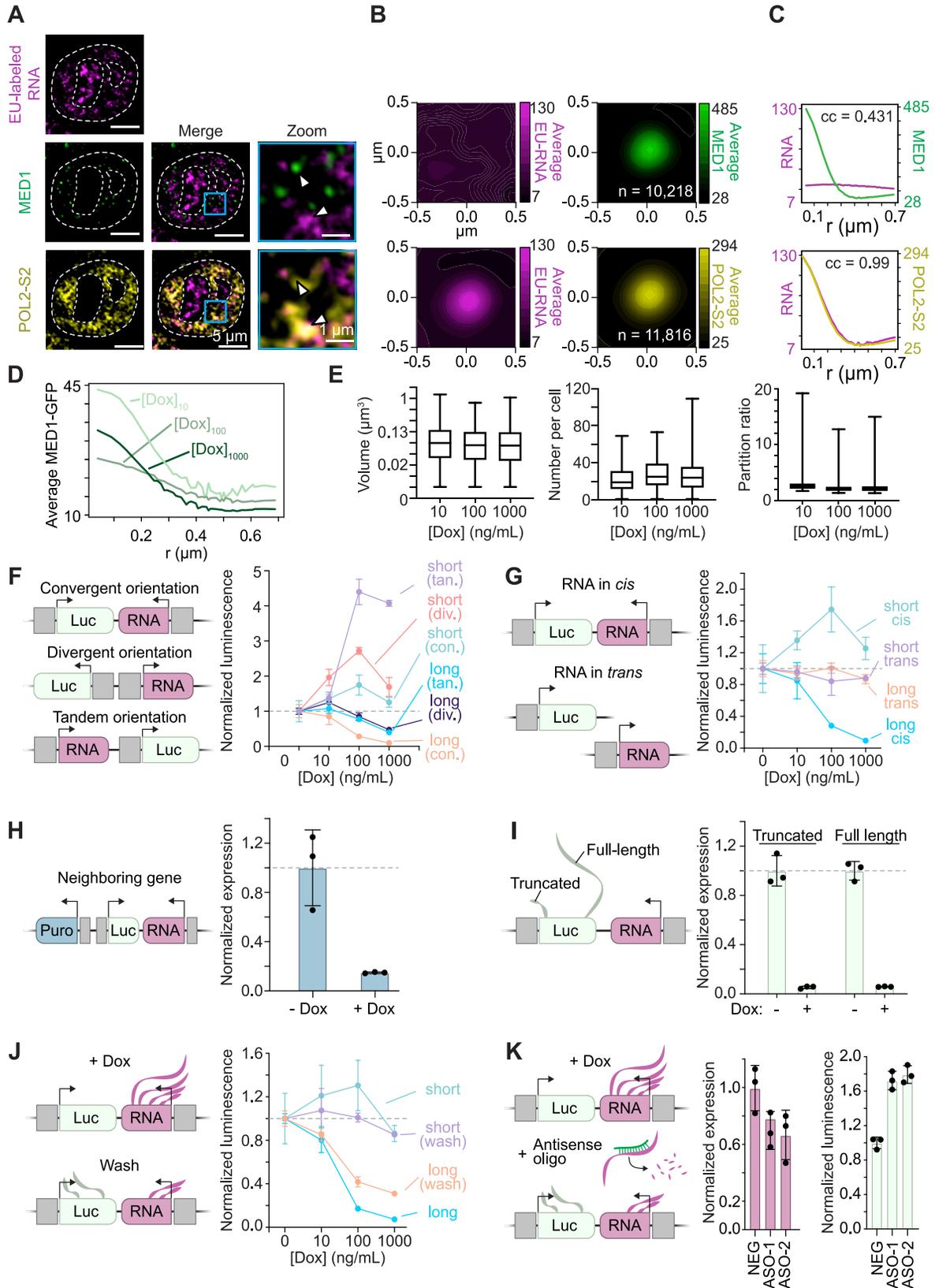
B. Similar analyses as in (A) are performed on a free-energy derived from Flory-Huggins model (STAR Methods).

C. Variation of condensate radius (left panel, normalized to value of R at $k_p = 0.02$) and condensate lifetime (right panel) with effective rates of RNA synthesis (k_p , abscissa) for simulations performed in 3D employing the Landau free-energy (Figure 4C; STAR Methods). Low values of k_p promote condensate stability whereas higher rates drive dissolution. The dashed line in the right-panel represents the conditions under which condensates are stable in the simulations and condensate lifetime is presented in units of simulation time (STAR Methods).

D. Similar analyses as in (C) are performed on a free-energy derived from Flory-Huggins model on a 2D grid (STAR Methods). Values of the condensate radius are normalized to value of R at $k_p = 0.08$.

E. Partition ratio, computed as maximum RNA concentration in condensate divided by dilute phase concentrations, are presented for simulations employing the Landau free-energy in 2 & 3-D as well as those employing the Flory-Huggins model in 2-D (left to right). When condensates are dissolved, the expected value of this ratio is 1 (as depicted by dashed gray lines). These calculations correspond to simulation data from Figures 4D, 4E, S6C, and S6D, respectively (left to right).

F. RNA diffusivity/mobility decreases with increasing protein concentration (abscissa) with an inverse correlation (left panel, $M_{\text{RNA}} \propto \phi_P^{-1}$), consistent with ideal models of crowding. Variation of condensate radius (middle, normalized to $r = 6.0$ units) and lifetime (right) with effective rates of RNA synthesis under conditions of phase-dependent RNA mobility. The simulations recapitulate key trends predicted by models with constant diffusion/mobility that are shown in Figures 4E and 4F.



(legend on next page)

Figure S7. The Effect of Local RNA Synthesis on Transcriptional Condensates and Transcription in Cells, Related to Figures 4, 5, and 6

A. Immunofluorescence of MED1 and POL2-S2 with EU-labeled RNA (10-minute incubation) in WT mESCs. Representative images are a single z-plane that has been subtracted by a median-filtered image and smoothed (STAR Methods) (left scale bars = 5 μm ; right scale bars = 1 μm).

B. Average signal analysis of EU-RNA, MED1 and POL2-S2. Average EU-RNA signal is centered at MED1 puncta (top) or at POL2-S2 puncta (bottom).

C. Radial distribution function with correlation between EU and MED1 (top) or POLII-S2 (bottom) channels for average IF signal in (B).

D. Radial distribution function of MED1-GFP signal for multiple dox concentrations from experiments in Figure 6C.

E. Global quantification of the volume (left), number per cell (middle), and partition ratio (right) of MED1-GFP condensates in reporter cells at multiple Dox concentrations.

F. Quantification of luciferase luminescence with increasing dox concentrations to stimulate expression of short (185 nt) and long (1313 nt) feedback RNAs with indicated orientations of feedback RNA and luciferase gene. The convergent data was collected as part of Figure 6G. Markers show the mean of at least 3 replicates and error bars depict the SD.

G. Quantification of luciferase luminescence with increasing dox concentration to stimulate expression of short (185 nt) and long (1313 nt) feedback RNAs with *cis* or trans integration of feedback RNA and luciferase reporter gene. The *cis* data was collected as part of Figure 6G. Markers show the mean of at least 3 replicates and error bars depict the SD.

H. RT-qPCR of neighboring puromycin resistance marker gene with and without stimulation of long (1313 nt) feedback RNA expression with dox.

I. RT-qPCR of luciferase mRNA gene with primers detecting truncated and full-length RNA with and without stimulation of long (1313 nt) feedback RNA expression with dox.

J. Quantification of luciferase luminescence with increasing dox concentration to stimulate expression of short (185 nt) and long (1313 nt) feedback RNAs, before and after dox washout.

K. The effect of antisense-oligo mediated degradation of feedback RNAs on luciferase luminescence. The expression of the feedback RNAs are measured by RT-qPCR (left bar graphs) and luciferase expression is measured by luminescence (right bar graphs). The luminescence values were first normalized to the no dox condition for that ASO, and then normalized to the dox condition of the negative control ASO.