A role for phase separation to prevent R-loops during transcription

Valery F. Thompson[1,2], Daniel R. Wieland[3], Vivian Mendoza-Leon[3], Helen I. Janis[3], Michelle A. Lay[1,2,3], Lucas M. Harrell[3] and Jacob C. Schwartz[1,2,*]

[1] Department of Pharmacology, University of Arizona, Tucson, AZ, 85724, USA
[2] University of Arizona Cancer Center, Tucson, AZ, 85724
[3] Department of Chemistry and Biochemistry, University of Arizona, Tucson, AZ, 85721, USA

* To whom correspondence should be addressed. Email: jcschwartz@arizona.edu

ABSTRACT

The protein FUS (FUSed in sarcoma) binds nascent RNA transcripts and all three nuclear RNA polymerases in metazoans. FUS undergoes phase separation to bind RNA polymerase II, which affects phosphorylation of the polymerase, recruitment of RNA processing factors, and transcription elongation and termination. This study investigates an activity of FUS targeting transcription but does not involve binding to the polymerase. An *in vitro* transcription assay revealed FUS prevented R-loops formed between the RNA transcript and DNA template, which could enhance transcription productivity for a non-eukaryote polymerase that FUS does not bind. The ability to bind RNA and phase separate was required for FUS activity. DRIP-seq analysis of human cells found endogenous activity matched shown by *in vitro* assays. FUS prevented R-loops at genomic locations and with sequence compositions consistent with previously characterized targets of FUS in cells. The model interpreted from these findings is that the abilities for RNA-binding and phase separation combine in FUS to prevent nucleic acid structures that disrupt chromosome stability and gene transcription.

INTRODUCTION

A prominent member of the heterogeneous nuclear ribonucleoprotein (hnRNP) family is FUS (FUsed in Sarcoma), known for mutations to its gene that can cause amyotrophic lateral sclerosis or drive pediatric sarcomas (1-3). FUS is conserved throughout metazoan species and among the highest expressed proteins in human tissues (3). FUS activity has been found to affect transcription by all three nuclear polymerases in eukaryotes (4-7). FUS is also an enigmatic protein. Its structure is mostly intrinsically disordered. Among its disordered domains, the arginine and glycine-rich (RGG/RG) and low complexity (LC) domains of FUS bind RNA Pol II. FUS binds RNA Pol I, RNA Pol III, and other transcription and RNA-processing factors (8-10). ChIP-seq analysis of FUS localization on chromatin shows enrichment at the transcription start sites (TSS) of most expressed protein-coding genes. By a mechanism not fully understood, a loss of FUS leads to increased RNA Pol II phosphorylation at Ser-2 (Ser2P) in the C-terminal domain (CTD) heptad repeat. FUS can block phosphorylation of CTD by kinases P-TEFb and CDK12 in *in vitro* assays (7,11). FUS is not observed to affect phosphorylation of the nearby Ser-5 position, which is crucial to transcription initiation (7).

FUS affinity for RNA is provided by its RNA-recognition motif (RRM) and three RGG/RG domains (12-16). FUS specificity for RNA-binding is degenerate but prefers G-rich sequences, RNA stem-loops, and complex RNA structure (13,16). The number of messenger and noncoding RNAs in a cell that bind FUS is large (3). RNA binding causes FUS to oligomerize (14,17). FUS self-assembly results in phase separation, yielding a large molecular scaffold or condensate. FUS condensates bind the CTD of RNA Pol II and those formed in cells have been found to contain >80% of the polymerase (4,8,18). This nuclear body has been observed at sizes 10 times the diameter of the polymerase. Assembly of FUS and RNA Pol II into the granule is transcription-dependent (19) The role of protein phase separation during eukaryotic transcription has drawn a high level of interest due to its profound implications to current paradigms for regulation of gene expression (20-22).

We set out to use *in vitro* transcription assays to better understand how isolated FUS functions contribute to its activity during transcription. We anticipated the most influential FUS activities would involve direct binding to a eukaryotic RNA polymerase and RNA processing factors (6,7,9,12). Our study instead revealed an effect of FUS on nucleic acid structures that influence transcription. We now show that FUS can alter transcription without direct interactions to the polymerase or other transcription factors.

MATERIAL AND METHODS

Recombinant protein expression and purification

FUS constructs used in this study are available upon request. All recombinant proteins were expressed and purified from *E. coli* BL21(DE3) cells. All FUS constructs were N-terminally fused to 6x His and maltose-binding protein (MBP) to improve purification and solubility. Protein expression was

induced after cell growth reached an OD600 of 0.8 and then allowed to continue overnight at 17 ºC with shaking (200 rpm). Purification was made using 1 to 2 g of frozen pellets of induced *E. coli*, lysates incubated with Ni-NTA Sepharose beads (Cytiva Lifesciences, 17531802), and eluted in FUS-SEC buffer (1 M urea, 0.25 M KCl, 50 mM Tris-HCl pH 8.0) with 250 mM imidazole and either 1.5 mM β-mercaptoethanol or 1 mM DTT added (4,13,15).

Plasmids for protein expression of T7(P266L) polymerase fused with a 6xHis tag were provided by A. Berman (University of Pittsburgh) (23). Expression in *E. coli* BL21(DE3) cells and purification were done essentially as previously published (24). Protein expression was induced by 0.5 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) when transformed cells grew to an OD600 of 0.6 then continued for 3 hours at 37 ºC. Induced cells stored at -80 ºC were thawed, lysed in T7 Buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 5% v/v glycerol, and 5 mM β-mercaptoethanol added just before use) with 1 mM imidazole then incubated with Ni-NTA beads. T7 Buffer with increasing imidazole was used to wash beads 4 times (1 mM imidazole), 2 times (10 mM imidazole), and to elute protein (100 mM imidazole).

Size-exclusion chromatography (SEC) was performed with a 10/300 Superdex® 200 column (Cytiva Lifesciences, 17517501) for FUS using FUS-SEC buffer and for T7 Pol with T7-SEC buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl, 0.1 mM EDTA, and 1 mM DTT). FUS protein was stored at room temperature for up to 4 weeks or until activity appeared lost. T7 Pol activity was assessed by titrating protein into transcription assays as described below. The absence of nuclease activity was confirmed by comparing plasmid DNA and transcribed RNA incubated overnight with or without protein and electrophoresis using a Novex™ TBE-urea, 6% polyacrylamide gel (Invitrogen, EC68655BOX) and stained with SYBR-Gold (Invitrogen, S11494).

Nucleic Acid preparation and purification

DNA substrates used in T7 transcription assays were prepared from pcDNA3 vector plasmid and gene inserted for RNA transcription was Flag-FUS (13). Bsu36I restriction enzyme (50 U, NEB, R0524S) was used to linearize plasmid (10 µg) by incubating at 37 ºC for a minimum of 2 hours in 1x CutSmart buffer (NEB, B7204S). Linearized plasmid was purified by extraction with an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1, pH 6.7/8.0) (Fisher Scientific, BP1752I-100) followed by two additional chloroform extractions (VWR, 97064-680) and ethanol precipitated. Plasmid concentrations were measured by UV absorption (260 nm) using a Biotek Epoch 2 plate reader with a TAKE3 plate. RNA TET456 was provided by T.R. Cech (University of Colorado Boulder) (25). Additional nucleic acids were yeast tRNA (Invitrogen, 15401011) and a commercially synthesized ssDNA (Millipore-Sigma). Nucleic acid sequences can be found in **Supplemental Table 1**.

T7 Transcription

Transcription reactions were prepared at room temperature in 20 µL transcription buffer (40 µM Tris-HCl pH 8, 5 mM DTT, 10 mM $MgCl_2$) with 2 mM NTP's (NEB, N0450S) and 40 U/µL RNasin™ Plus (Promega, N2615). T7 (P266L) polymerase (26) was added to linearized plasmid DNA (25 ng/µL) at

37 ºC, incubated for up to 2 hours, and halted either by heating to >75 ºC or addition of EDTA (12 mM final concentration). To detect RNA products by electrophoresis, samples were heated (90 ºC, 5 min) in formamide loading dye (49% formamide, 5 mM EDTA, 0.05% xylene cyanol, 0.05% bromophenol blue) and loaded to a Novex™ TBE-Urea, 6% polyacrylamide gels (Invitrogen, EC68655BOX) to run in TBE (VWR, 97061-754) at 180 V for 50 minutes. A ssRNA ladder (NEB, N0364S) was used for molecular weight standards. Gels were stained with SYBR-Gold (Invitrogen, S11494) in TBE and imaged using a Bio-Rad ChemiDoc™ MP Imaging System. Densitometry was performed with Bio-Rad Image Lab™ Software v. 6.0.1. The ratio of RNA/DNA was calculated for each sample to control for loading differences. RNA and DNA were also quantified by Quant-iT™ Assay (Thermo Fisher Scientific, Q333140 and Q33120) according to manufacturer's instructions.

Co-immunoprecipitation assay

Recombinant expressed FUS protein (1.3 µg) and anti-FUS (4H11) antibody (11 µg; Santa Cruz Biotechnology, sc-47711) were immobilized by incubating together for 1 hour at room temperature in phosphate buffered saline (PBS, pH 7.4) (Sigma, P3813) and 0.1 mg/mL bovine serum albumin protein (VWR, 97062-508), then 1 hour with Pierce™ Protein A/G agarose beads (Thermo Fisher Scientific, 20421). Beads were washed 4 times with 0.1% NP-40 substitute (VWR, 97064-730) in tris-buffered saline (TBS). Beads with FUS were incubated with T7 Pol for 1 hour at room temperature with rotation, then washed twice in high salt, twice in TBS with 0.1% NP-40 substitute, and one with TBS. Elution of protein bound to beads was incubated at 95 ºC for ten minutes in 2x LDS-SB (Life Technologies, NP0008) with 100 mM DTT. Nucleic acids used were prepared as described above and **Supplemental Table 1**. Detection of eluted proteins was made by western blot assay using anti-FUS (A300-294A, Bethyl Laboratories) to identify FUS protein and anti-His (HIS.H8, Novus Biologicals, NBP2-31055) to identify His tags for both FUS and T7 Pol proteins.

RNA:DNA hybrid dot blot assays

RNA:DNA hybrids were detected and quantified by dot blot assay. From transcription assays, 10% of reactions were spotted on N+ membranes (Fisher Scientific, 45-000-850), dried and UV crosslinked. Blots were blocked with 5% nonfat dried milk (NFDM) in TBS-T, then incubated overnight at 4 ºC with an anti-RNA:DNA hybrid antibody S9.6 against RNA:DNA hybrids (1/2000 dilution) in TBS-T with 2.5% NFDM. Blots were washed 5 minutes 4 times in TBS-T, incubated in secondary antibody (goat anti-mouse IgG-HRP, 1/20000) for 1 hr at RT. Images were taken after incubation with SuperSignal™ West Pico PLUS (Fisher Pierce, PI34578) and imaged using a Chemidoc MP system (BioRad).

DRIP-seq assay and data analysis

DRIP-seq assays were performed essentially as a previously published protocol (27). HEK293T/17 cells (ATCC, CRL-11268) grown and passaged in DMEM (5% FBS) were transfected by siRNA using RNAiMax™ (Invitrogen, 13778150) in Optimem (Invitrogen, 31985070). Sequences of siRNAs used are included in **Supplemental Table 1**. Between 7 and 9 x 10^6 cells were harvested 72 hours post-siRNA treatment, and cell pellets were frozen in liquid nitrogen. Pellets were later resuspended in 1.6

mL of TE buffer (10 mM Tris-HCl pH 8, 1 mM EDTA) that was added to include 0.6% sodium dodecyl sulfate (SDS) and 60 µg/mL Proteinase K. After Proteinase K digested lysates overnight at 37 ºC, DNA was extracted using the MaXtract High Density tubes (Qiagen, 129065). DNA was spooled by glass Pasteur pipette, transferred, then ethanol precipitated.

DNA was digested overnight in TE buffer with 1x NEBuffer™ 2 (New England Biolabs Inc., #B6002), BSA (95 µg/mL), spermidine (1 mM), and 30 U each of restriction enzymes BsrGI, EcoRI, HindIII, SspI, and XbaI. Restriction digest was confirmed by agarose (0.8% w/v) gel electrophoresis. After restriction enzyme digestions, DNA was extracted by Phase Lock Gel™ (VWR, 10847-800) per manufacturer instructions. Negative control samples were digested by 40U RNaseH (Fisher Scientific, 50-811-717) per 10 µg DNA. 50 µg of DNA was incubated overnight at 4 ºC with 20 µg S9.6 antibody and in 500 µL DRIP buffer (1x TE buffer, 10 mM Na phosphate pH 7, 140 mM NaCl, 0.05% Triton X-100). 5 µg of DNA was kept for input samples. Antibody precipitated with 100 µL Protein A/G beads (Millipore, IP05-1.5ML) was washed twice in 700 µL DRIP buffer. Hybrids were eluted by incubating beads with 300 µL DRIP Elution buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 0.5% SDS, 0.5 µg/µL Proteinase K) for 45 minutes at 55 ºC, then extracted by Phase Lock Gel™. Samples were sonicated for 30 cycles (30 sec on / 30 sec off) follow by ethanol precipitation.

Sample libraries were prepared and sequenced by Novogene Corporation Inc. using a NovoSeq6000 with 150 base paired-end reads (8 G raw data per sample). Reads were trimmed using TrimGalore (version 0.6.6) and aligned to GRCh38 using STAR (version_2.7.8a), essentially as previously described. Bam files were processed using the deepTools2.0 suite (version 3.5.1) to normalize to 1x depth (reads per genome coverage) with bamCoverage, merge replicates by bigwigCompare, and tags binned by computeMatrix, which was then used to compute scaling factors and generate heat maps and profiles. Merged bigwig files were converted to bedgraph and then peaks by Macs2 (version 2.1.1.20160309). HOMER (version v4.11.1) was used to calculate GC-content and annotate peaks. Data was visualized, and figures were generated using Integrated Genomics Viewer (version 2.5.0). DRIP-seq data is available (GSE206740) from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/).

RESULTS

FUS can increase transcription without binding the polymerase

One model for FUS activity on RNA Pol II transcription would be to form protein assemblies at active genes that can concentrate polymerases, transcription machinery, or RNA processing factors that aid in transcription initiation and elongation. We have previously shown that RNA Pol II transcription of a naked DNA template is greatly enhanced by recombinant FUS protein added in concentrations near those of the cell, >2 µM (4). If protein-protein interactions with RNA Pol II were required for activity, we hypothesized that FUS would not act on a polymerase it cannot bind.

Multiple domains in FUS can bind RNA Pol II (8,14,28,29). FUS binding and activity has been observed for all three nuclear RNA polymerases (6,30). We turned to a DNA-dependent RNA polymerase from T7 phage (T7 Pol) to test FUS activity in the absence of binding to the polymerase (**Supplemental Figure 1A**). The P266L mutant of T7 Pol was chosen, which has higher stability during promoter clearance and produces fewer abortive transcripts (26). Run-off transcription of a linear DNA template was assessed by acrylamide gel electrophoresis (PAGE) and SYBR staining of the DNA and RNA present over time (**Figure 1A**, top). Contrary to expectation, transcription assays in the presence of freshly purified FUS (2 μM) yielded more RNA transcript (**Figure 1A**). Titrating FUS concentration in T7 Pol transcription assays correspondingly increased RNA yield (**Figure 1B**).

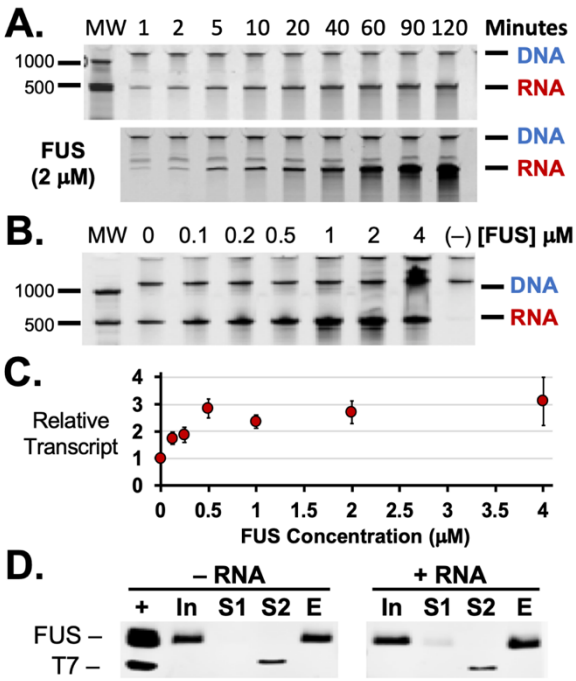

Figure 1. FUS increases RNA yields from transcription by T7 phage polymerase. (A) RNA produced by T7 Pol transcription of linear plasmid DNA seen by urea-PAGE and SYBR staining. Increased RNA yields were seen with addition of FUS (2 μM, lower gel) relative to without FUS (top gel). **(B)** RNA product seen by urea-PAGE and SYBR staining after T7 Pol transcription with FUS protein added. At 4 μM FUS, aggregation caused RNA and DNA to shift to the well. NTPs were not added to negative control (–) samples. **(C)** RNA concentrations were measured by RNA-specific fluorescent dye from transcription by T7 Pol (2 hours at 37 ºC) with an increasing concentration of FUS. Results were averaged (N=5) and relative to no FUS added. Error bars represent ±SEM. **(D)** Western blots of co-IP assays test FUS binding to T7 Pol. FUS protein (In) incubated with anti-FUS was pulled down, leaving no protein in the supernatant (S1). T7 Pol incubated with FUS-bound beads remained in supernatants (S2) and did not elute with FUS from beads (E). The presence of RNA ("+ RNA") (12 nM) did not cause FUS to bind T7 Pol. A lane with FUS and T7 Pol (+) served as a positive control for detection by western blot assay. See also **Supplemental Figure 1C** for non-specific IgG immunoprecipitation control.

At FUS concentrations above 2 μM, a mobility shift was observed, suggesting an interference from protein aggregation on the detection of RNA by PAGE. Consistent with this interpretation, the well shift was observed when transcription was prevented by omitting NTP. The shift was also reduced by treating samples of run-off transcripts with proteinase K (**Supplement Figure 1B**). To eliminate this factor from transcript measurements, we also quantified run-off transcription assays using fluorescent dyes specific for RNA or DNA, which found that the increase to RNA yield could reach ≥ 3-fold with FUS added at concentrations ≥ 2 μM (**Figure 1C**).

We considered whether FUS had an unforeseen ability to bind T7 Pol. Co-immunoprecipitation (co-IP) assays revealed that a FUS-specific antibody could deplete FUS out of solution (S1) but not T7 Pol (S2) (**Figure 1D**). Likewise, only FUS was detected in eluates of beads bound to anti-FUS antibodies (E, **Figure 1D**). Co-IP assays that included an RNA known to cause FUS binding to RNA Pol II (RMRP) did not induce detectable binding to T7 Pol (**Figure 1D**) (7,14). Linear plasmid DNA also had no effect (**Supplemental Figure 1C**). In conclusion, we found no evidence of direct binding between T7 Pol and FUS that might explain or contribute to the activity observed.

Activity on transcription is specific to FUS and not competed by RNA

We considered whether high protein concentrations might increase transcription by non-specific molecular crowding. T7 Pol activity has been reported to increase through molecular crowding by agents that include glycerol and PEG (31-33). We tested a well-folded soluble protein, bovine serum albumin (BSA), and detected no increase in RNA product (**Figure 2A** and **2B**). We also saw FUS activity diminished and was lost for protein stored at -80 ºC. Activity was stable up to 4 weeks for FUS kept at room temperature (**Supplemental Figure 2A**). We noted the temperature effect on FUS matched what is for many proteins a result of structural instability during freezing.

Though RNA did not cause FUS to bind T7 Pol, we questioned whether RNA could increase or sequester activity on transcription. An RNA that is commonly used for a non-specific competitor to RNA-binding is tRNA. A second RNA, TET456, has tertiary structure that includes parallel stacked helices and is previously known to cause FUS binding to RNA Pol II. FUS activity was unchanged by the addition of tRNA at the same concentration by mass as the RNA produced by the assay (**Figure 2C, Supplemental Figure 2B**). Titrating TET456 RNA to concentrations up to 12.8 μM also did not affect FUS activity (**Supplemental Figure 2C**). Lastly, single-stranded DNA had no effect on FUS activity on T7 Pol transcription (**Supplemental Figure 2D**).

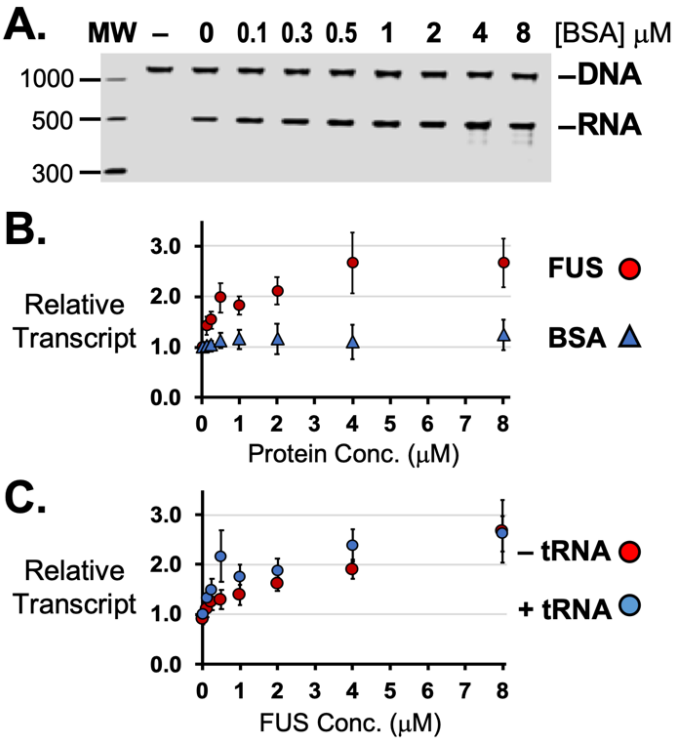

Figure 2. FUS effects on T7 transcription are specific. (**A**) RNA produced by T7 polymerase was not affected by BSA protein added up to 8 μM concentration. Negative control samples (–) were without NTP in the assay. (**B**) Fluorescence-based quantification of RNA transcribed in the presence of increasing concentration of FUS or BSA protein. (**C**) Quantification by densitometry of SYBR-stained PAGE analysis of RNA transcribed in the presence of increasing concentration of FUS protein with or without added yeast tRNA (6.4 μM). All results are averaged from 3 or 4 experiments relative to samples without FUS or BSA added. Error bars represent ±SEM.

FUS inhibits RNA:DNA hybrids formed during transcription

We considered whether the activity observed may involve nucleic acid structures. One structure is an RNA:DNA hybrid, formed by a nascent transcript displacing the non-template strand to bind its complementary template DNA strand (34,35). RNA:DNA hybrids internal to larger DNA molecules are called R-loops, which are found in cells at low but consistent levels (34,36-39). R-loops can affect transcription and chromosome stability, which may threaten many DNA functions, and multiple mechanisms exist in cells to prevent or remove R-loops (34,40).

We used a well-characterized antibody, S9.6, to detect RNA:DNA hybrids by dot blot assays (27,41). The results showed hybrids increase during run-off transcription (**Figure 3A**). Digestion of RNA:DNA hybrids by RNaseH confirmed that the signals detected were specific to hybrids. Controls that lacked NTPs or T7 Pol showed that hybrids were a product of transcription, and the S9.6

antibody did not detect hybrids when transcription assays included the TET456 RNA, which lacks complementarity to the plasmid (**Supplemental Figures 3A** and **3B**). While FUS (2 μM) did not affect the level of hybrids detected after 2 hours of transcription, hybrids were decreased between 20 and 40 minutes of transcription (**Supplemental Figure 3C**). As FUS concentration increased, RNA:DNA hybrids detected were reduced (**Figure 3B**). To test that transcripts made could still form hybrids, we performed dot blot assay with samples that had been heated and saw similar levels of hybrids across all samples (**Figure 3B**, anneal). Lastly, quantitating dot blot assays confirmed a 2-fold reduction in hybrids at the 20 minute timepoint for transcription (**Figure 3C**).

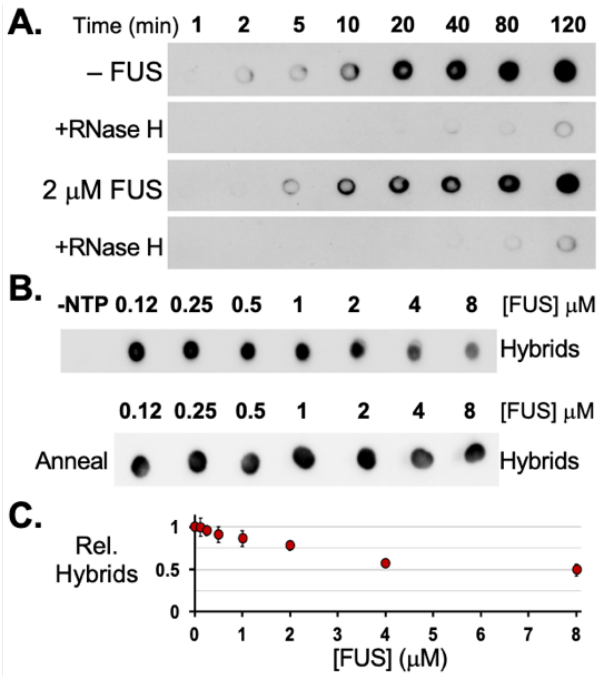

Figure 3. FUS prevents R-loop formation during transcription. (**A**) A dot blot assay using the S9.6 antibody detected RNA:DNA hybrids formed during T7 Pol transcription with or without the presence of FUS (2 μM). Equal volumes of samples incubated with RNaseH (+RNaseH) were probed to test antibody specificity. (**B**) Dot blot assays detect RNA:DNA hybrids for T7 Pol transcription after 20 minutes with increasing concentrations of FUS protein. Negative controls (–) did not have NTP in the assay. Half of each sample collected was annealed by heating (95 ºC) and then cooled to form hybrids from RNA transcripts as a positive control. (**C**) Dot blot assays quantified by densitometry show reduced hybrids formed with FUS present relative to without FUS. All results are averaged from 3 experiments. Error bars represent ±SEM.

Requirements for RNA-binding and phase separation for FUS activity

We sought to uncover a mechanism for activity by investigating the role of individual FUS domains. The LC domain contributes to protein binding, oligomerization, and phase separation (4,42-44). The

three RGG/RG domains are capable of both phase separation and RNA-binding (8,14,45-47). We compared FUS activity to that of the LC domain alone (FUS-LC) or RNA-binding domains (del-LC) (**Figure 4A**). Neither truncated protein had significant activity on transcript yields or hybrids (**Figure 4B** and **C**). This suggested that both phase separation and RNA-binding were necessary for activity.

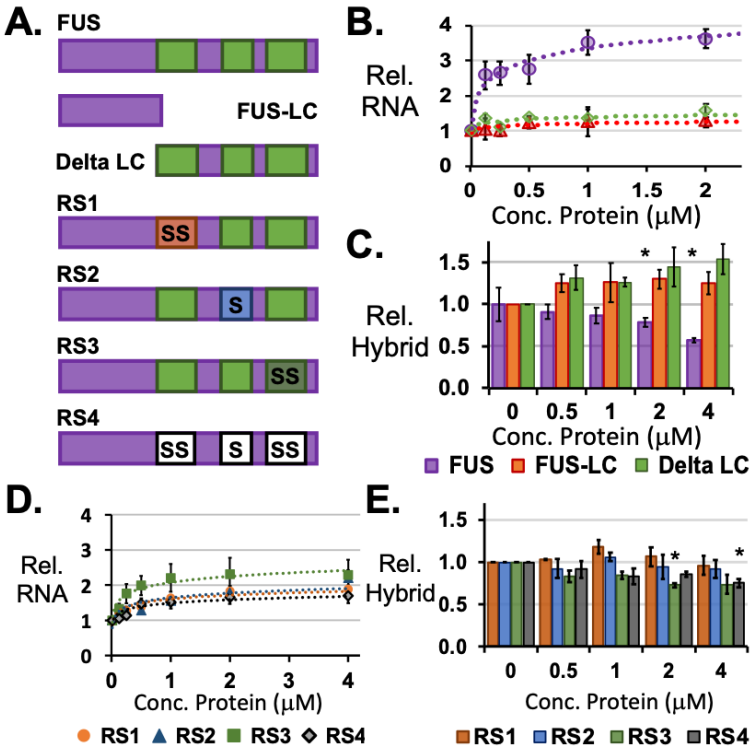

Figure 4. FUS oligomerization and RNA-binding domains are required to prevent R-loops. (**A**) A diagram showing domains of MBP-tagged FUS, truncations, and RS proteins. Arginine residues are replaced by serine in a single RGG/RG domain of RS1, RS2, and RS3 and all RGG/RG domains for RS4. FUS-LC and Delta-LC proteins were unable to (**B**) increase RNA yield measured by fluorescence or (**C**) prevent RNA:DNA hybrids. The titration of FUS protein shown was performed in parallel to assays shown for truncated and RS proteins. Modest activity of RS proteins was indicated by (**D**) increases to RNA yield quantified by fluorescence assay and (**E**) reduced signals for RNA:DNA hybrids measured by dot blot assay. All results were averaged from 3 or 4 experiments. Error bars represent standard error about the mean. An asterisk (*) indicates $p < 0.05$.

Next, we tested FUS with reduced RNA-binding by substituting serine for arginine residues in one (RS1, RS2, RS3) or all (RS4) the RGG domains (**Figure 4A**) (13). RS1 and RS2 were previously found to retain most their RNA-binding affinity measured by EMSA but lost activity in cells by 2-fold. Addition of RS1 or RS2 could increase transcript yields by 2-fold (**Figure 4D**) and did not reduce hybrids at 20 minutes (**Figure 4E**). RS3 increased the transcript T7 Pol produced by >2-fold, while decreasing RNA:DNA hybrids a small but significant amount ($p < 0.05$, student t-test assuming equal variances, **Figure 4D–E**). RS4 was previously shown to have >30-fold lower affinity for RNA than

FUS, with no RNA-binding observed in cells (13). RS4 activity on T7 Pol transcript levels equalled that of RS1 and RS2, and a small reduction of hybrids was observed (**Figure 4D–E**).

We next investigated transcription assay conditions for evidence of phase separation by FUS and RS proteins. To preserve conditions close to the transcription assays while aiding the detection of condensates by microscopy, transcription assays were incubated 24 hours at room temperature. Visible condensates up to 0.5 µm in diameter were found for FUS, RS1, RS2, and RS3 proteins (2 µM), but no RS4 condensates were observed (**Supplemental Figure 4A**). Conditions also tested were incubation at 4 ºC, increased protein concentrations (20 µM), and with the addition of RNA. While FUS formed many condensates of ≥1 µm in diameter, none were seen for RS4 (**Supplemental Figure 4B** and **4C**). FUS at high concentrations (85 µM) formed condensates large enough to clearly show liquid characteristics, but phase separation by RS4 (100 µM) remained undetected (**Supplemental Figure 4D**). Since condensate sizes were invisible to microscopy at early timepoints for the transcription assay, RS4 may have reached this undetectable level in phase separation but continued no further. We therefore concluded that RS4 was less competent to phase separate than FUS and other RS proteins.

FUS prevents R-loops in human cells

We next asked whether the FUS activity found *in vitro* was also present in cells. If true, a measurable increase in R-loops may result from a knockdown of FUS protein. We used a next-generation sequencing approach to quantify R-loops enriched by the S9.6 antibody (DRIP-seq) (27,48,49). We used western blot assay to verify protein levels were reduced in HEK293T/17 cells by transfection of an siRNA targeted for FUS (siFUS) relative to an siRNA control of scrambled sequence (SCR) (**Supplemental Figure 5A**). Half of each lysate was treated with RNaseH as a negative control. Dot blot assays with the S9.6 antibody confirmed the presence of R-loops in eluted DRIP samples (**Supplemental Figure 5B**). Published real-time PCR controls also indicated R-loop enrichment by comparison of signals for R-loop positive gene promoters, *RPL13A* and *TFPT*, to those of RNaseH-treated negative controls and an R-loop negative promoter, *EGR1* (**Supplemental Figure 5C**) (27).

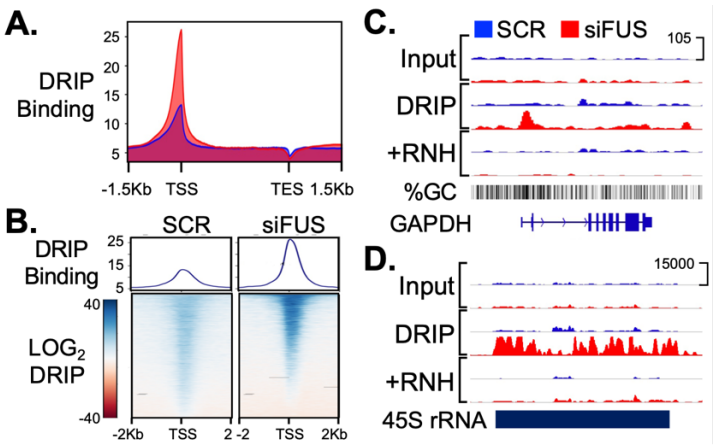

Figure 5. Knockdown of FUS increases R-loop abundance in human cells. DRIP-seq was performed on HEK293T/17 cells with knockdown of FUS protein, siFUS, or negative control, SCR. (**A**) Normalized DRIP-seq signal profiles at human genes were averaged and plotted for siFUS (red) or SCR (blue) samples. Gene lengths were scaled to align transcription start (TSS) and end (TES) sites. (**B**) Heat maps show DRIP-seq signals for SCR (left) or siFUS (right) treatments. (**C**) DRIP-seq showing increases to R-loop signal at the *GAPDH* gene promoter resulting from siFUS treatment. (**D**) The largest R-loop signals were found at rRNA genes in siFUS treated samples.

We immunoprecipitated DNA fragments and generated libraries for sequencing. DRIP-seq analysis revealed R-loop enrichment at transcription start sites (TSS) for expressed genes (**Figure 5A**). For samples with FUS knockdown, R-loop signals measured at promoters were increased by >2-fold in siFUS samples for 29% (N=5260) of protein coding genes (N=18188) (**Figure 5B**). For the representative and well-expressed *GAPDH*, increased R-loop signals overlapped previous reports of FUS enrichment determined by ChIP-seq (**Figure 5C**) (7). The largest increase in R-loop signals was found in rRNA genes (**Figure 5D**). Since rRNA genes are transcribed by RNA Pol I, this result was consistent with the prediction for FUS activity to extend to any RNA polymerase.

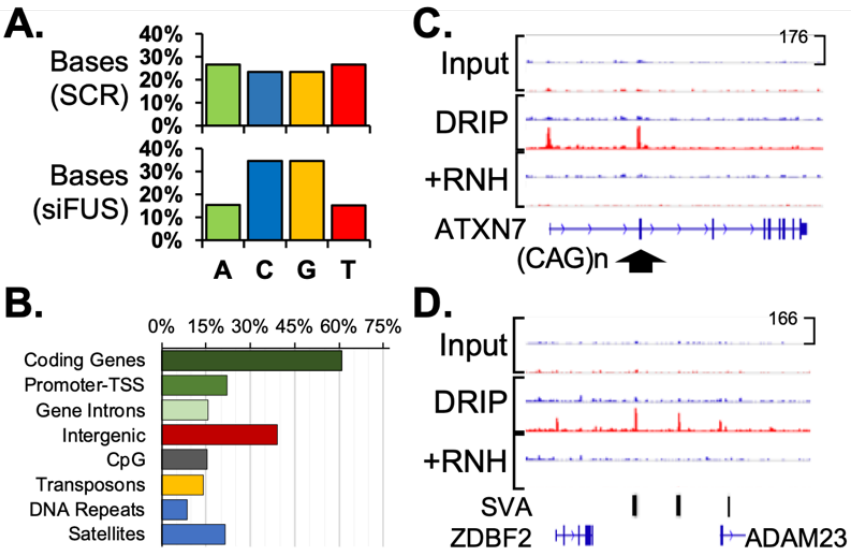

Figure 6. R-loops affected by FUS have high GC-content. (**A**) The base composition of DNA was quantified within peaks found in DRIP-seq signals. Sequences of peaks detected in siFUS-treated samples were comprised on average of >60% GC content. (**B**) Protein-coding genes comprised the largest percentage of peaks detected in the FUS knockdown results, in contrast to SCR-treated sample results (see also **Supplemental 6A**). (**C**) DRIP-seq analysis of siFUS-treated samples reveals a peak in R-loop signal arising at a CAG-repeat DNA sequence at the *ATXN7* gene. (**D**) DRIP-seq results for siFUS-treated samples were enriched for R-loops at SVA sequences, a SINE retrotransposon sub-type.

More peaks were detected in R-loop signals from siFUS (n=16184) relative to SCR (n=5046) samples. FUS knockdown also preferentially increased R-loop signals at sites in DNA with high GC content (**Figure 6A**). The greatest number of R-loop peaks shifted from intergenic and satellite DNA in SCR samples to protein-coding genes in siFUS samples (**Figure 6B**). A representative simple DNA repeat affected by FUS knockdown was seen in the CAG-repeat of the *ATXN7* gene, which when expanded can cause neurodegenerative disease (**Figure 6C**) (50). FUS knockdown also caused R-loops to increase at retrotransposons (n=2036). The most affected were SINE retrotransposons (n=1237), with nearly half (48%, n=987) being the SINE-VNTR-Alu (SVA) sub-type (**Figure 6D**) (51). We noted that SINE elements are expected to be transcribed by RNA Pol III. Inspection of other RNA Pol III genes found R-loop signals increased in an 5S rRNA gene cluster on chromosome 1 (**Supplemental Figure 6B**). R-loops did not appear increased at RNA Pol III genes *RMRP*, *RN7SK*, or tRNA gene clusters found in chromosomes 1 and 6 (**Supplemental Figure 6C**).

In summary, DRIP-seq analysis revealed FUS activity did suppress R-loops in HEK293T/17. Enrichment in GC content was notable due to the previous reported preference of FUS to bind GC-rich RNA sequences. We concluded that the characteristics of FUS activity on R-loops in human cells matched expectations based on that seen in T7 Pol run-off transcription assays.

DISCUSSION

Here we explore a mechanism for a FUS activity on nucleic acid structures and transcription that did not involve binding to the polymerase. This activity did require binding to RNA and was correlated with competency for phase separation. Our working model is that FUS binding prevents hybridization of nascent RNA transcripts to their complementary DNA template. DRIP-seq analysis showed in human cells that FUS prevents R-loops at expressed genes and other loci throughout the genome. This study highlights the role R-loops can have for biology and disease. Its results also offer new insights that unite common themes to FUS activities seen in and out of cells.

FUS has been found to influence the activity of all three nuclear RNA polymerases (3,6). We add to this list the RNA polymerase of T7 phage virus. This surprising result appears to be allowed because the activity does not involve binding to the polymerase (**Figure 1D**). The *in vitro* activity was seen at micromolar protein concentrations for FUS, which is within the range of FUS concentrations (4 to 8 μM) measured in the nucleus of cells (4,52,53). The largest effect of FUS found in cells was at rRNA genes, which may be explained by the high activity of RNA Pol I relative to other polymerases that can produce and sustain high levels of R-loops detected by DRIP-seq (54,55). Evidence for FUS activity on RNA Pol III was suggested by R-loops detected near retrotransposons and 5S rRNA genes. Interestingly, the FUS activity on R-loops at RNA Pol II promoters matched well with previous reports of FUS localization on chromatin measured by ChIP-seq (7). This may indicate that FUS activity is most pronounced early during transcription at regions that span promoters and pause sites (7,11).

However, if R-loops found at the simple DNA repeats in gene bodies were formed by the nascent transcript originating from the distant mRNA promoter, it could be argued that FUS activity extends throughout the gene. The FUS knockdown also produced R-loops at sites not previously reported to have detectable FUS enrichment, such as retrotransposons.

Like other activities of FUS, RNA-binding was needed to prevent R-loops (14). FUS is implicated in RNA processing factor recruitment to RNA Pol II and nascent transcripts (9,12,56). Failure to recruit snRNPs, SR proteins, and hnRNPs has been reported to give rise to R-loops (48,57-61). These effects could also be seen during T7 Pol transcription, which further supports the model interpreted from this study (60). This study has highlighted a feature that competing RNA or DNA could not interfere with activity (**Figure 2C, Supplemental Figure 2B–C**). At least two considerations tie this observation to cell biology. First, high concentration of RNA is expected in the nucleus. Nascent transcripts make more secondary structure available for binding that resemble substrates preferred by FUS determined by *in vitro* studies (62-64). Once transcription is complete, many secondary structures become inaccessible as RNA forms tertiary structure, which is also true of RNAs in our competition assays with tertiary structures that are known from x-ray crystallography (25,62). Second, FUS concentrations measured in cells are >100-fold above the binding K_D of FUS measured for a wide range in RNA substrates (13,15). While amount of FUS produced by cells may appear excessive for binding to RNA, it provides ideal conditions for phase separation (4). Therefore, the phase separation component to FUS activity has likely imposed a helpful condition to counter the diluting effect of RNA competing for FUS in the cell.

An ability to phase separate correlated to activity on transcription and R-loops. FUS phase separation has been hypothesized to recruit RNA Pol II to gene promoters, but recruitment is irrelevant to activity on T7 Pol, which FUS does not bind (3,14,17). Instead, phase separation likely increases FUS affinity for sites adjacent along the RNA, providing FUS an advantage to compete RNA from its nearby DNA template. Another aspect of phase separation considered by this study was the striking reduction of this activity in RS4. Two models may reconcile the observations of high phase separation activity for the LC domain, while FUS activity is limited without RNA (14). In one model, FUS C-terminal domains bind and sequester the LC domain whenever RNA is not bound (65). In a second model, the bulk of full-length FUS adds a steric hindrance to phase separation that must be overcome by the concentrating effect of binding to RNA (66-68). A challenge to the second model is that fusion to a large protein, MBP, affects LC domain activity less than might be expected, and phase separation activity seen in the RGG/RG domains would be predicted to add to activity, not subtract (14,69). If the 25 arginine residues replaced in RS4 formed intramolecular interactions with the LC domain, phase separation activity would be predicted to increase. With so little evidence for activity in RS4, future studies of how FUS activity is controlled might find reason to prioritize the surface residues of the RRM and ZnF domains, as well as the few RGG/RG domain residues not changed in RS4 (13,47,70).

R-loops play important roles in cell biology and disease (34,40,71). A long-standing paradigm has been that R-loops negatively impact transcription, but this assumption has been challenged by evidence of R-loops providing necessary effects in cells (34,37,72,73). If the latter proves correct,

cells must have some mechanism to turn away FUS activity. Our study did not find how FUS activity is sustain past 20 minutes of transcription, but RNA Pol II transcription in cells has been found to occur as bursts (74-76). This provides periodic rests in transcription of 1 to 6 minutes, which may reset FUS activity and avoid the decreases observed in our assays. Finally, if evidence continues to tie DNA instability in ALS or Ewing sarcoma to R-loops, the significance of this FUS activity should rise above others that are not defective in portions of disease without FUS pathology (34,48).

To summarize, these results can improve future research by rephrasing the questions studied to be more specific. For example, if the activity described here extends to any RNA polymerase, can endogenous FUS act on viral or mitochondrial polymerases? The loss of FUS function produced a large increase in R-loops at rRNA genes. Can this lead to a depletion of rRNA and harm translation? Why do R-loops appear at retrotransposons? Is it a backup mechanism to preserve silencing or merely reflect that protection against transposon activity was lost? Concluding with a question about RNA Pol II, do the R-loops prevented by FUS alter splicing or transcription termination, and what relationship do these R-loops have to RNA Pol II phosphorylation? Some of the questions rephrased are longstanding because they are difficult to address through unambiguous experiments (3,7,77,78). If understanding the FUS activity studied here can lessen challenges to investigate other FUS activities, these future studies should provide important new insights on the mechanisms that regulate gene expression, drive neurodegenerative disease, and prevent childhood cancers.

DATA AVAILABILITY

DRIP-seq data is available (GSE206740) from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENT

We thank Dr. Andrea Berman (U. Pittsburgh) for T7 Pol plasmids and protocols; Dr. Thomas R. Cech (CU Boulder) for TET456 RNA; and H. Miller and Dr. Alex Bishop (UTHSSA) for helpful advice and discussion about DRIP-seq protocol and data analysis.

Author contributions: V.F.T. designed and performed most experiments and analyzed data; D.R.W. performed co-IP experiments and accompanying transcription run-off and dot blot assays; V.L-M. performed transcription run-off, dot blot, and phase separation assays; H.I.J. expressed, purified, and confirmed activity of proteins; M.A.L. performed phase separation, dot blot, and western assays for DRIP-seq experiments; L.M.H. expressed, purified, and confirmed activity of proteins, performed

transcription run-off assays, and optimized protocols; J.C.S. secured funding, designed and analyzed data for experiments, and supervised the project. J.C.S. wrote the manuscript. V.F.T., D.R.W., and L.M.H. edited the manuscript and writing portions of the Materials and Methods section. The final manuscript was approved by all authors.

FUNDING

This work was supported by the National Institutes of Health [CA238499 and CA259570 to J.C.S.]; the American Cancer Society [RSG-18-237-01-DMC to J.C.S.]; and a training fellowship from the University of Arizona Cancer Center to M.A.L. Funding for open access charge: National Institutes of Health.

CONFLICT OF INTEREST

There are no conflicts of interest to report relating to this work.

REFERENCES

1. Owen, I., Yee, D., Wyne, H., Perdikari, T.M., Johnson, V., Smyth, J., Kortum, R., Fawzi, N.L. and Shewmaker, F. (2021) The oncogenic transcription factor FUS-CHOP can undergo nuclear liquid-liquid phase separation. *J Cell Sci*, **134**.
2. Rhine, K., Makurath, M.A., Liu, J., Skanchy, S., Lopez, C., Catalan, K.F., Ma, Y., Fare, C.M., Shorter, J., Ha, T. *et al.* (2020) ALS/FTLD-Linked Mutations in FUS Glycine Residues Cause Accelerated Gelation and Reduced Interactions with Wild-Type FUS. *Mol Cell*, **80**, 666-681 e668.
3. Schwartz, J.C., Cech, T.R. and Parker, R.R. (2015) Biochemical Properties and Biological Functions of FET Proteins. *Annu Rev Biochem*, **84**, 355-379.
4. Thompson, V.F., Victor, R.A., Morera, A.A., Moinpour, M., Liu, M.N., Kisiel, C.C., Pickrel, K., Springhower, C.E. and Schwartz, J.C. (2018) Transcription-Dependent Formation of Nuclear Granules Containing FUS and RNA Pol II. *Biochemistry*, **57**, 7021-7032.
5. Conlon, E.G. and Manley, J.L. (2017) RNA-binding proteins in neurodegeneration: mechanisms in aggregate. *Genes Dev*, **31**, 1509-1528.
6. Tan, A.Y. and Manley, J.L. (2010) TLS inhibits RNA polymerase III transcription. *Mol Cell Biol*, **30**, 186-196.
7. Schwartz, J.C., Ebmeier, C.C., Podell, E.R., Heimiller, J., Taatjes, D.J. and Cech, T.R. (2012) FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. *Genes Dev*, **26**, 2690-2695.
8. Murthy, A.C., Tang, W.S., Jovic, N., Janke, A.M., Seo, D.H., Perdikari, T.M., Mittal, J. and Fawzi, N.L. (2021) Molecular interactions contributing to FUS SYGQ LC-RGG phase separation and co-partitioning with RNA polymerase II heptads. *Nat Struct Mol Biol*, **28**, 923-935.
9. Yu, Y. and Reed, R. (2015) FUS functions in coupling transcription to splicing by mediating an interaction between RNAP II and U1 snRNP. *Proc Natl Acad Sci U S A*, **112**, 8608-8613.

10. Kawaguchi, T., Rollins, M.G., Moinpour, M., Morera, A.A., Ebmeier, C.C., Old, W.M. and Schwartz, J.C. (2020) Changes to the TDP-43 and FUS Interactomes Induced by DNA Damage. *J Proteome Res*, **19**, 360-370.
11. Krasnopolsky, S., Marom, L., Victor, R.A., Kuzmina, A., Schwartz, J.C., Fujinaga, K. and Taube, R. (2019) Fused in sarcoma silences HIV gene transcription and maintains viral latency through suppressing AFF4 gene activation. *Retrovirology*, **16**, 16.
12. Jutzi, D., Campagne, S., Schmidt, R., Reber, S., Mechtersheimer, J., Gypas, F., Schweingruber, C., Colombo, M., von Schroetter, C., Loughlin, F.E. *et al.* (2020) Aberrant interaction of FUS with the U1 snRNA provides a molecular mechanism of FUS induced amyotrophic lateral sclerosis. *Nat Commun*, **11**, 6341.
13. Ozdilek, B.A., Thompson, V.F., Ahmed, N.S., White, C.I., Batey, R.T. and Schwartz, J.C. (2017) Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding. *Nucleic Acids Res*, **45**, 7984-7996.
14. Schwartz, J.C., Wang, X., Podell, E.R. and Cech, T.R. (2013) RNA seeds higher-order assembly of FUS protein. *Cell Rep*, **5**, 918-925.
15. Wang, X., Schwartz, J.C. and Cech, T.R. (2015) Nucleic acid-binding specificity of human FUS protein. *Nucleic Acids Res*, **43**, 7535-7543.
16. Loughlin, F.E., Lukavsky, P.J., Kazeeva, T., Reber, S., Hock, E.M., Colombo, M., Von Schroetter, C., Pauli, P., Clery, A., Muhlemann, O. *et al.* (2019) The Solution Structure of FUS Bound to RNA Reveals a Bipartite Mode of RNA Recognition with Both Sequence and Shape Specificity. *Mol Cell*, **73**, 490-504 e496.
17. Kwon, I., Kato, M., Xiang, S., Wu, L., Theodoropoulos, P., Mirzaei, H., Han, T., Xie, S., Corden, J.L. and McKnight, S.L. (2013) Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains. *Cell*, **155**, 1049-1060.
18. Emmanouilidis, L., Esteban-Hofer, L., Damberger, F.F., de Vries, T., Nguyen, C.K.X., Ibanez, L.F., Mergenthal, S., Klotzsch, E., Yulikov, M., Jeschke, G. *et al.* (2021) NMR and EPR reveal a compaction of the RNA-binding protein FUS upon droplet formation. *Nat Chem Biol*, **17**, 608-614.
19. Victor, R.A., Thompson, V.F. and Schwartz, J.C. (2021) Isolating and Analyzing Protein Containing Granules from Cells. *Curr Protoc*, **1**, e35.
20. Ahmed, N.S., Harrell, L.M., Wieland, D.R., Lay, M.A., Thompson, V.F. and Schwartz, J.C. (2021) Fusion protein EWS-FLI1 is incorporated into a protein granule in cells. *RNA*.
21. Wei, M.T., Chang, Y.C., Shimobayashi, S.F., Shin, Y., Strom, A.R. and Brangwynne, C.P. (2020) Nucleated transcriptional condensates amplify gene expression. *Nat Cell Biol*.
22. 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. *et al.* (2019) Enhancer Features that Drive Formation of Transcriptional Condensates. *Mol Cell*, **75**, 549-561 e547.
23. Berman, A.J., Gooding, A.R. and Cech, T.R. (2010) Tetrahymena telomerase protein p65 induces conformational changes throughout telomerase RNA (TER) and rescues telomerase reverse transcriptase and TER assembly mutants. *Mol Cell Biol*, **30**, 4965-4976.
24. Ellinger, T. and Ehricht, R. (1998) Single-step purification of T7 RNA polymerase with a 6-histidine tag. *Biotechniques*, **24**, 718-720.
25. Szewczak, A.A., Podell, E.R., Bevilacqua, P.C. and Cech, T.R. (1998) Thermodynamic stability of the P4-P6 domain RNA tertiary structure measured by temperature gradient gel electrophoresis. *Biochemistry*, **37**, 11162-11170.
26. Guillerez, J., Lopez, P.J., Proux, F., Launay, H. and Dreyfus, M. (2005) A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc Natl Acad Sci U S A*, **102**, 5958-5963.
27. Sanz, L.A. and Chedin, F. (2019) High-resolution, strand-specific R-loop mapping via S9.6-based DNA-RNA immunoprecipitation and high-throughput sequencing. *Nature protocols*, **14**, 1734-1755.

28. Zuo, L., Zhang, G., Massett, M., Cheng, J., Guo, Z., Wang, L., Gao, Y., Li, R., Huang, X., Li, P. *et al.* (2021) Loci-specific phase separation of FET fusion oncoproteins promotes gene transcription. *Nat Commun*, **12**, 1491.
29. Chi, B., O'Connell, J.D., Yamazaki, T., Gangopadhyay, J., Gygi, S.P. and Reed, R. (2018) Interactome analyses revealed that the U1 snRNP machinery overlaps extensively with the RNAP II machinery and contains multiple ALS/SMA-causative proteins. *Sci Rep*, **8**, 8755.
30. Martinez-Macias, M.I., Moore, D.A., Green, R.L., Gomez-Herreros, F., Naumann, M., Hermann, A., Van Damme, P., Hafezparast, M. and Caldecott, K.W. (2019) FUS (fused in sarcoma) is a component of the cellular response to topoisomerase I-induced DNA breakage and transcriptional stress. *Life Sci Alliance*, **2**.
31. Kartje, Z.J., Janis, H.I., Mukhopadhyay, S. and Gagnon, K.T. (2021) Revisiting T7 RNA polymerase transcription in vitro with the Broccoli RNA aptamer as a simplified real-time fluorescent reporter. *J Biol Chem*, **296**, 100175.
32. Sokolova, E., Spruijt, E., Hansen, M.M., Dubuc, E., Groen, J., Chokkalingam, V., Piruska, A., Heus, H.A. and Huck, W.T. (2013) Enhanced transcription rates in membrane-free protocells formed by coacervation of cell lysate. *Proc Natl Acad Sci U S A*, **110**, 11692-11697.
33. Ge, X., Luo, D. and Xu, J. (2011) Cell-free protein expression under macromolecular crowding conditions. *PLoS One*, **6**, e28707.
34. Crossley, M.P., Bocek, M. and Cimprich, K.A. (2019) R-Loops as Cellular Regulators and Genomic Threats. *Mol Cell*, **73**, 398-411.
35. Santos-Pereira, J.M. and Aguilera, A. (2015) R loops: new modulators of genome dynamics and function. *Nat Rev Genet*, **16**, 583-597.
36. Dumelie, J.G. and Jaffrey, S.R. (2017) Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq. *Elife*, **6**.
37. Chen, L., Chen, J.Y., Zhang, X., Gu, Y., Xiao, R., Shao, C., Tang, P., Qian, H., Luo, D., Li, H. *et al.* (2017) R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Mol Cell*, **68**, 745-757 e745.
38. Nadel, J., Athanasiadou, R., Lemetre, C., Wijetunga, N.A., P, O.B., Sato, H., Zhang, Z., Jeddeloh, J., Montagna, C., Golden, A. *et al.* (2015) RNA:DNA hybrids in the human genome have distinctive nucleotide characteristics, chromatin composition, and transcriptional relationships. *Epigenetics Chromatin*, **8**, 46.
39. Ginno, P.A., Lott, P.L., Christensen, H.C., Korf, I. and Chedin, F. (2012) R-loop formation is a distinctive characteristic of unmethylated human CpG island promoters. *Mol Cell*, **45**, 814-825.
40. Garcia-Muse, T. and Aguilera, A. (2019) R Loops: From Physiological to Pathological Roles. *Cell*, **179**, 604-618.
41. Bou-Nader, C., Bothra, A., Garboczi, D.N., Leppla, S.H. and Zhang, J. (2022) Structural basis of R-loop recognition by the S9.6 monoclonal antibody. *Nat Commun*, **13**, 1641.
42. Hughes, M.P., Sawaya, M.R., Boyer, D.R., Goldschmidt, L., Rodriguez, J.A., Cascio, D., Chong, L., Gonen, T. and Eisenberg, D.S. (2018) Atomic structures of low-complexity protein segments reveal kinked beta sheets that assemble networks. *Science*, **359**, 698-701.
43. Murray, D.T., Kato, M., Lin, Y., Thurber, K.R., Hung, I., McKnight, S.L. and Tycko, R. (2017) Structure of FUS Protein Fibrils and Its Relevance to Self-Assembly and Phase Separation of Low-Complexity Domains. *Cell*.
44. Lin, Y., Currie, S.L. and Rosen, M.K. (2017) Intrinsically disordered sequences enable modulation of protein phase separation through distributed tyrosine motifs. *J Biol Chem*, **292**, 19110-19120.
45. Ozdilek, B.A., Thompson, V.F., Ahmed, N.S., White, C.I., Batey, R.T. and Schwartz, J.C. (2017) Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding. *Nucleic Acids Res*.

46. Yoshizawa, T., Ali, R., Jiou, J., Fung, H.Y.J., Burke, K.A., Kim, S.J., Lin, Y., Peeples, W.B., Saltzberg, D., Soniat, M. *et al.* (2018) Nuclear Import Receptor Inhibits Phase Separation of FUS through Binding to Multiple Sites. *Cell*, **173**, 693-705 e622.
47. Wang, J., Choi, J.M., Holehouse, A.S., Lee, H.O., Zhang, X., Jahnel, M., Maharana, S., Lemaitre, R., Pozniakovsky, A., Drechsel, D. *et al.* (2018) A Molecular Grammar Governing the Driving Forces for Phase Separation of Prion-like RNA Binding Proteins. *Cell*, **174**, 688-699 e616.
48. Gorthi, A., Romero, J.C., Loranc, E., Cao, L., Lawrence, L.A., Goodale, E., Iniguez, A.B., Bernard, X., Masamsetti, V.P., Roston, S. *et al.* (2018) EWS-FLI1 increases transcription to cause R-loops and block BRCA1 repair in Ewing sarcoma. *Nature*, **555**, 387-391.
49. Li, L., Matsui, M. and Corey, D.R. (2016) Activating frataxin expression by repeat-targeted nucleic acids. *Nat Commun*, **7**, 10606.
50. Niss, F., Zaidi, W., Hallberg, E. and Strom, A.L. (2021) Polyglutamine expanded Ataxin-7 induces DNA damage and alters FUS localization and function. *Mol Cell Neurosci*, **110**, 103584.
51. Barnada, S.M., Isopi, A., Tejada-Martinez, D., Goubert, C., Patoori, S., Pagliaroli, L., Tracewell, M. and Trizzino, M. (2022) Genomic features underlie the co-option of SVA transposons as cis-regulatory elements in human pluripotent stem cells. *PLoS Genet*, **18**, e1010225.
52. Patel, A., Lee, H.O., Jawerth, L., Maharana, S., Jahnel, M., Hein, M.Y., Stoynov, S., Mahamid, J., Saha, S., Franzmann, T.M. *et al.* (2015) A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation. *Cell*, **162**, 1066-1077.
53. Hein, M.Y., Hubner, N.C., Poser, I., Cox, J., Nagaraj, N., Toyoda, Y., Gak, I.A., Weisswange, I., Mansfeld, J., Buchholz, F. *et al.* (2015) A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell*, **163**, 712-723.
54. Abraham, K.J., Khosraviani, N., Chan, J.N.Y., Gorthi, A., Samman, A., Zhao, D.Y., Wang, M., Bokros, M., Vidya, E., Ostrowski, L.A. *et al.* (2020) Nucleolar RNA polymerase II drives ribosome biogenesis. *Nature*, **585**, 298-302.
55. Shen, W., Sun, H., De Hoyos, C.L., Bailey, J.K., Liang, X.H. and Crooke, S.T. (2017) Dynamic nucleoplasmic and nucleolar localization of mammalian RNase H1 in response to RNAP I transcriptional R-loops. *Nucleic Acids Res*, **45**, 10672-10692.
56. Das, R., Yu, J., Zhang, Z., Gygi, M.P., Krainer, A.R., Gygi, S.P. and Reed, R. (2007) SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol Cell*, **26**, 867-881.
57. Shivji, M.K.K., Renaudin, X., Williams, C.H. and Venkitaraman, A.R. (2018) BRCA2 Regulates Transcription Elongation by RNA Polymerase II to Prevent R-Loop Accumulation. *Cell Rep*, **22**, 1031-1039.
58. Bhatia, V., Barroso, S.I., Garcia-Rubio, M.L., Tumini, E., Herrera-Moyano, E. and Aguilera, A. (2014) BRCA2 prevents R-loop accumulation and associates with TREX-2 mRNA export factor PCID2. *Nature*, **511**, 362-365.
59. Santos-Pereira, J.M., Herrero, A.B., Garcia-Rubio, M.L., Marin, A., Moreno, S. and Aguilera, A. (2013) The Npl3 hnRNP prevents R-loop-mediated transcription-replication conflicts and genome instability. *Genes Dev*, **27**, 2445-2458.
60. Li, X. and Manley, J.L. (2005) Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell*, **122**, 365-378.
61. Tresini, M., Warmerdam, D.O., Kolovos, P., Snijder, L., Vrouwe, M.G., Demmers, J.A., van, I.W.F., Grosveld, F.G., Medema, R.H., Hoeijmakers, J.H. *et al.* (2015) The core spliceosome as target and effector of non-canonical ATM signalling. *Nature*, **523**, 53-58.
62. Scharfen, L. and Neugebauer, K.M. (2021) Transcription Regulation Through Nascent RNA Folding. *J Mol Biol*, **433**, 166975.
63. Mustoe, A.M., Brooks, C.L. and Al-Hashimi, H.M. (2014) Hierarchy of RNA functional dynamics. *Annu Rev Biochem*, **83**, 441-466.
64. Pan, T. and Sosnick, T. (2006) RNA folding during transcription. *Annu Rev Biophys Biomol Struct*, **35**, 161-175.

65. Wang, X., Arai, S., Song, X., Reichart, D., Du, K., Pascual, G., Tempst, P., Rosenfeld, M.G., Glass, C.K. and Kurokawa, R. (2008) Induced ncRNAs allosterically modify RNA-binding proteins in cis to inhibit transcription. *Nature*, **454**, 126-130.
66. Lin, Y., Protter, D.S., Rosen, M.K. and Parker, R. (2015) Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins. *Mol Cell*, **60**, 208-219.
67. Sharp, P.A., Chakraborty, A.K., Henninger, J.E. and Young, R.A. (2022) RNA in formation and regulation of transcriptional condensates. *RNA*, **28**, 52-57.
68. Forman-Kay, J.D., Ditlev, J.A., Nosella, M.L. and Lee, H.O. (2022) What are the distinguishing features and size requirements of biomolecular condensates and their implications for RNA-containing condensates? *RNA*, **28**, 36-47.
69. Burke, K.A., Janke, A.M., Rhine, C.L. and Fawzi, N.L. (2015) Residue-by-Residue View of In Vitro FUS Granules that Bind the C-Terminal Domain of RNA Polymerase II. *Mol Cell*, **60**, 231-241.
70. Qamar, S., Wang, G., Randle, S.J., Ruggeri, F.S., Varela, J.A., Lin, J.Q., Phillips, E.C., Miyashita, A., Williams, D., Strohl, F. *et al.* (2018) FUS Phase Separation Is Modulated by a Molecular Chaperone and Methylation of Arginine Cation-pi Interactions. *Cell*, **173**, 720-734 e715.
71. Gorthi, A. and Bishop, A.J.R. (2018) Ewing sarcoma fusion oncogene: At the crossroads of transcription and DNA damage response. *Mol Cell Oncol*, **5**, e1465014.
72. Belotserkovskii, B.P., Soo Shin, J.H. and Hanawalt, P.C. (2017) Strong transcription blockage mediated by R-loop formation within a G-rich homopurine-homopyrimidine sequence localized in the vicinity of the promoter. *Nucleic Acids Res*, **45**, 6589-6599.
73. Chen, P.B., Chen, H.V., Acharya, D., Rando, O.J. and Fazzio, T.G. (2015) R loops regulate promoter-proximal chromatin architecture and cellular differentiation. *Nat Struct Mol Biol*, **22**, 999-1007.
74. 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.
75. Tantale, K., Mueller, F., Kozulic-Pirher, A., Lesne, A., Victor, J.M., Robert, M.C., Capozi, S., Chouaib, R., Backer, V., Mateos-Langerak, J. *et al.* (2016) A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat Commun*, **7**, 12248.
76. 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. *et al.* (2016) RNA Polymerase II cluster dynamics predict mRNA output in living cells. *Elife*, **5**.
77. Payer, L.M. and Burns, K.H. (2019) Transposable elements in human genetic disease. *Nat Rev Genet*, **20**, 760-772.
78. Bourque, G., Burns, K.H., Gehring, M., Gorbunova, V., Seluanov, A., Hammell, M., Imbeault, M., Izsvak, Z., Levin, H.L., Macfarlan, T.S. *et al.* (2018) Ten things you should know about transposable elements. *Genome Biol*, **19**, 199.

Figure 1. FUS increases RNA yields from transcription by T7 phage polymerase. (A) RNA produced by T7 Pol transcription of linear plasmid DNA seen by urea-PAGE and SYBR staining. Increased RNA yields were seen with addition of FUS (2 μM, lower gel) relative to without FUS (top gel). **(B)** RNA product seen by urea-PAGE and SYBR staining after T7 Pol transcription with FUS protein added. At 4 μM FUS, aggregation caused RNA and DNA to shift to the well. NTPs were not added to negative control (–) samples. **(C)** RNA concentrations were measured by RNA-specific fluorescent dye from transcription by T7 Pol (2 hours at 37 ºC) with an increasing concentration of FUS. Results were averaged (N=5) and relative to no FUS added. Error bars represent ±SEM. **(D)** Western blots of co-IP assays test FUS binding to T7 Pol. FUS protein (In) incubated with anti-FUS

was pulled down, leaving no protein in the supernatant (S1). T7 Pol incubated with FUS-bound beads remained in supernatants (S2) and did not elute with FUS from beads (E). The presence of RNA ("+ RNA") (12 nM) did not cause FUS to bind T7 Pol. A lane with FUS and T7 Pol (+) served as a positive control for detection by western blot assay. See also **Supplemental Figure 1C** for non-specific IgG immunoprecipitation control.

Figure 2. FUS effects on T7 transcription are specific. (**A**) RNA produced by T7 polymerase was not affected by BSA protein added up to 8 μM concentration. Negative control samples (–) were without NTP in the assay. (**B**) Fluorescence-based quantification of RNA transcribed in the presence of increasing concentration of FUS or BSA protein. (**C**) Quantification by densitometry of SYBR-stained PAGE analysis of RNA transcribed in the presence of increasing concentration of FUS protein with or without added yeast tRNA (6.4 μM). All results are averaged from 3 or 4 experiments relative to samples without FUS or BSA added. Error bars represent ±SEM.

Figure 3. FUS prevents R-loop formation during transcription. (**A**) A dot blot assay using the S9.6 antibody detected RNA:DNA hybrids formed during T7 Pol transcription with or without the presence of FUS (2 μM). Equal volumes of samples incubated with RNaseH (+RNaseH) were probed to test antibody specificity. (**B**) Dot blot assays detect RNA:DNA hybrids for T7 Pol transcription after 20 minutes with increasing concentrations of FUS protein. Negative controls (–) did not have NTP in the assay. Half of each sample collected was annealed by heating (95 ºC) and then cooled to form hybrids from RNA transcripts as a positive control. (**C**) Dot blot assays quantified by densitometry show reduced hybrids formed with FUS present relative to without FUS. All results are averaged from 3 experiments. Error bars represent ±SEM.

Figure 4. FUS oligomerization and RNA-binding domains are required to prevent R-loops. (**A**) A diagram showing domains of MBP-tagged FUS, truncations, and RS proteins. Arginine residues are replaced by serine in a single RGG/RG domain of RS1, RS2, and RS3 and all RGG/RG domains for RS4. FUS-LC and Delta-LC proteins were unable to (**B**) increase RNA yield measured by fluorescence or (**C**) prevent RNA:DNA hybrids. The titration of FUS protein shown was performed in parallel to assays shown for truncated and RS proteins. Modest activity of RS proteins was indicated by (**D**) increases to RNA yield quantified by fluorescence assay and (**E**) reduced signals for RNA:DNA hybrids measured by dot blot assay. All results were averaged from 3 or 4 experiments. Error bars represent standard error about the mean. An asterisk (*) indicates $p < 0.05$.

Figure 5. Knockdown of FUS increases R-loop abundance in human cells. DRIP-seq was performed on HEK293T/17 cells with knockdown of FUS protein, siFUS, or negative control, SCR. (**A**)

Normalized DRIP-seq signal profiles at human genes were averaged and plotted for siFUS (red) or SCR (blue) samples. Gene lengths were scaled to align transcription start (TSS) and end (TES) sites. (**B**) Heat maps show DRIP-seq signals for SCR (left) or siFUS (right) treatments. (**C**) DRIP-seq showing increases to R-loop signal at the *GAPDH* gene promoter resulting from siFUS treatment. (**D**) The largest R-loop signals were found at rRNA genes in siFUS treated samples.

Figure 6. R-loops affected by FUS have high GC-content. (**A**) The base composition of DNA was quantified within peaks found in DRIP-seq signals. Sequences of peaks detected in siFUS-treated samples were comprised on average of >60% GC content. (**B**) Protein-coding genes comprised the largest percentage of peaks detected in the FUS knockdown results, in contrast to SCR-treated sample results (see also **Supplemental 6A**). (**C**) DRIP-seq analysis of siFUS-treated samples reveals a peak in R-loop signal arising at a CAG-repeat DNA sequence at the *ATXN7* gene. (**D**) DRIP-seq results for siFUS-treated samples were enriched for R-loops at SVA sequences, a SINE retrotransposon sub-type.

Supplemental Figure 1. Purification of protein and immunoprecipitation. (A) Coomassie stained gel, left, of fractions collected from the SEC step of protein purification shown by trace, right, for His-MBP-FUS and His-T7pol^{P266L} proteins. (B) Testing the role of protein aggregation in shifting RNA to the well during PAGE analysis. Samples were treated with either proteinase K, Prot K, or sodium dodecyl sulfate, SDS. (C) Co-immunoprecipitation assays for MBP-FUS and T7 Pol as described in Figure 1D. Above is immunoprecipitation by antibody to FUS and below is a non-specific mouse IgG control antibody.

Supplemental Figure 2. FUS activity is not competed by RNA or DNA. (A) Assay of the same FUS preparations used for Figure 1B and 1C, re-assayed after 3 months of storage at room temperature. (B) Titration of yeast tRNA into transcription run-off assays did not change the activity of FUS protein. (C) Titration of a TET456 RNA into transcription run-off assays did not change the activity of FUS protein. (D) Addition of a single-stranded DNA, ssDNA, into transcription run-off assays did not change the titrated activity of FUS protein.

Supplemental Figure 3. Time course of RNA:DNA hybrid formation during run-off transcription assays. (A) Assay as shown in Figure 3A and showing controls of omitting NTP, (–), or T7 Pol, -Pol, or addition of a non-specific RNA, +R. Below are samples treated with Dnase to show that the S9.6 antibody is detecting nucleic acids. (B) PAGE analysis of transcript products from part A. In the left most lane, -Pol+R, can be seen the band of the TET456 RNA added. (C) Quantification of

transcription run-off time course experiments by fluorescence assay. Results are averaged from 3 experiments. Error bars represent standard error about the mean (±SEM).

Supplemental Figure 4. Inspection of phase separation using microscopy. (A) Phase separation of FUS and RS proteins at 2 μM and the same transcription run-off assay conditions after incubation for 24 hours. Additional images of RS2 replicates were intentionally included to help visual comparison. (B) The transcription and incubation as in part A was repeated for RS3 and RS4 proteins at 10 μM, revealing no condensates formed for RS4 protein. (C) FUS and RS4 were incubated to form condensates at 20 μM protein for 24 hours at 4 ºC. To also test for stimulation of phase separation in RS4, TET456 RNA was added at 0.5 μM. Fluorescent images were made by adding GFP-tagged FUS LC-domain (amino acids 1 to 266) at 1/150 times the concentration of FUS. (D) Condensates formed by FUS or RS4 at high protein concentration as indicated. Scale bars are 2 μm for parts A and B and 4 μm in parts C and D.

Supplemental Figure 5. Verification of DRIP assays. (A) Knockdown of FUS protein in HEK293T/17 cells confirmed by measuring FUS protein levels. TDP-43 protein is used as a loading control. An siRNA of scrambled sequence is used as a negative control (SCR) and the knockdown is by siRNA targeting FUS mRNA (siFUS). (B) Verification of restriction enzyme digestion of genomic DNA recovered from HEK293T/17 cells, left, and the presence of RNA:DNA hybrids in genomic DNA recovered by dot blot assay, right, using the S9.6 antibody. An antibody against double-stranded DNA, dsDNA, was also used as a positive control. (C) Realtime PCR is used to confirm the presence of RNA:DNA hybrids after immunoprecipitation for gene promoters expected to have R-loops, RPL13A and TFPT, or a promoter not expected to have R-loops, EGR1. Samples treated with Rnase H, +RNH, are also not expected to have R-loops. In the table to the right are averaged Ct values measured using published primer controls for validating DRIP-seq sample quality. Results shown are from 2 replicates and error bars indicate standard error about the mean (±SEM).

Supplemental Figure 6. DRIP-seq analysis of R-loops formed after FUS knockdown. (A) Percentage of DRIP-seq peaks located at genomic features in SCR-treated samples. See **Figure 6B** for siFUS-treatment results. Genome viewer tracks show changes in RNA:DNA hybrids mapped at (B) a 5S rRNA gene cluster or (C) a tRNA gene cluster.