Full-length direct RNA sequencing uncovers stress-granule dependent RNA decay upon cellular stress

Showkat A. Dar[1,*], Sulochan Malla[1,*], Vlastimil Martinek[1,2,3], Matthew J. Payea[1], Christopher T. Lee, Jessica Martin[1], Aditya J. Khandeshi[1], Jennifer L. Martindale[1], Cedric Belair[1], Manolis Maragkakis[1,$]

[1] Laboratory of Genetics and Genomics, National Institute on Aging, Intramural Research Program, National Institutes of Health, Baltimore, MD 21224, USA
[2] Central European Institute of Technology, Masaryk University, 625 00 Brno, Czech Republic
[3] National Centre for Biomolecular Research, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic

* Equal contribution
[$] To whom correspondence should be addressed

Contact: maragkakis@nih.gov

Abstract

Cells react to stress by triggering response pathways, leading to extensive alterations in the transcriptome to restore cellular homeostasis. The role of RNA metabolism in shaping the cellular response to stress is vital, yet the global changes in RNA stability under these conditions remain unclear. In this work, we employ direct RNA sequencing with nanopores, enhanced by 5' end adaptor ligation, to comprehensively interrogate the human transcriptome at single-molecule and nucleotide resolution. By developing a statistical framework to identify robust RNA length variations in nanopore data, we find that cellular stress induces prevalent 5' end RNA decay that is coupled to translation and ribosome occupancy. Unlike typical RNA decay models in normal conditions, we show that stress-induced RNA decay is dependent on XRN1 but does not depend on removal of the poly(A) tail. We observed that RNAs undergoing decay are predominantly enriched in the stress granule transcriptome. Inhibition of stress granule formation via genetic ablation of G3BP1 and G3BP2 fully rescues RNA length and suppresses stress-induced decay. Our findings reveal RNA decay as a key determinant of RNA metabolism upon cellular stress and dependent on stress-granule formation.

Introduction

Cells respond to stress by triggering extensive transcriptome remodeling to restore cellular homeostasis (Galluzzi et al., 2018; Payea et al., 2023). Cellular stress responses have been under intense study as the capacity of cells to activate adaptive pathways deteriorates with age (Derisbourg et al., 2021; Haigis and Yankner, 2010; Maragkakis et al., 2022) and has been associated with many diseases, particularly age-related neurodegeneration (Beckman and Ames, 1998; Stadtman and Berlett, 1997).

The response to different stressors converges at the induction of the Integrated Stress Response (ISR), a common adaptive molecular pathway (Harding et al., 2003; Pakos-Zebrucka et al., 2016) that induces phosphorylation of eukaryotic translation initiation factor 2 alpha (eIF2α) to inhibit translation initiation and decrease global protein synthesis (Hinnebusch, 1994). In turn, mRNAs exiting the translational pool get localized in stress granules (SGs), membraneless organelles that are formed upon stress when translation of mRNAs gets arrested (Panas et al., 2016; Protter and Parker, 2016). SGs comprise several different types of biomolecules, particularly mRNAs and RNA-binding proteins, but their precise functions in mRNA storage or decay are not yet fully understood (Anderson and Kedersha, 2008; Marcelo et al., 2021).

Short- and long-read sequencing following tagging of RNA 5' ends has revealed that RNA content in cells is a mix of intact RNA molecules and RNA decay intermediates that exist in a transient, partially degraded state (Ibrahim et al., 2021, 2018; Pelechano et al., 2015). However much less is known about the global transcriptome-wide state of RNA under stress and whether stress leads to RNA decay or stabilization. Evidence have been found in both directions with stress having been associated with mRNA stabilization in a translation-dependent and independent way (Gowrishankar et al., 2006; Hilgers et al., 2006; Horvathova et al., 2017; Kharel et al., 2023). On the other hand, in yeast, ribosomal and other RNA binding proteins, which under certain physiological conditions contact mRNAs to facilitate translation, show diminished mRNA association under heat shock stress and decreased mRNA abundance mediated by the 5′-3′ exoribonuclease Xrn1 (Bresson et al., 2020). Similarly, under oxidative stress, deletion of Xrn1 can lead to accumulation of oxidized RNAs bearing nucleotide adducts such as 8-oxoguanosine (8-oxoG), in turn inhibiting translation and triggering No-Go decay (Yan et al., 2019).

Here, we globally profile the integrity of individual RNA molecules in human cells under stress by full-length direct RNA sequencing with nanopores. We develop a new computational framework for statistical interrogation of RNA decay at single molecule and nucleotide resolution. We find that upon cellular stress, RNAs globally exhibit marked shortening at their 5′ end. This stress-induced decay is mediated by XRN1, but is not dependent on prior de-adenylation, as classical models would posit (Passmore and Coller, 2021). Importantly, we find that the decaying RNA molecules are preferentially enriched in the SG transcriptome. Prevention of SG formation by ablation of cells for G3BP1 and G3BP2 rescues RNA length upon stress, suggesting SG formation as a required step for stress-induced RNA decay. Our results identify 5' RNA decay as a hallmark of cellular stress response that defines the transcriptome and that is dependent on stress-granule formation.

Results

Stress induces widespread decay at the 5' end of RNAs

To explore the state of RNA under stress, we treated HeLa cells with sodium arsenite (0.5 mM) for 60 minutes to induce oxidative stress. We verified the formation of SGs and increased eIF2α phosphorylation (discussed in subsequent section) and monitored cell viability up to 24 hours after treatment with no major change observed (**Sup. Fig. 1a**). We collected arsenite-treated and untreated cells and performed direct RNA-Seq using the TERA-Seq protocol (Ibrahim et al., 2021) (**Fig. 1a**). Direct RNA-Seq profiles RNA at single molecule resolution without cDNA conversion or PCR amplification (Workman et al., 2019). TERA-Seq additionally incorporates unique adapters that are ligated to the 5' phosphorylated end of individual RNA molecules, which are then sequenced along with the native sequence. Nanopore-based direct RNA sequencing proceeds in the 3' to 5' direction thus the inclusion, and subsequent computational identification of an adapter, at the 5' end of a read acts as internal control ensuring an RNA molecule has been sequenced in its entirety. Notably, the method does not biochemically select for ligated molecules, thus the resultant sequencing datasets comprise both ligated and non-ligated molecules. Ligated molecules can be computationally recognized and separated to be analyzed jointly or independently to non-ligated reads. Samples were prepared in triplicates and sequenced on a MinION device, yielding approximately 2 million long reads per library (**Sup. Table 1**). Replicates displayed high expression correlation and high alignment efficiency to the human genome **(Sup. Fig. 1b, Sup. Table 1)**. Consistent with oxidative stress induction, differential gene expression analysis identified several stress response genes being upregulated in arsenite while gene ontology enrichment analysis of differentially expressed genes showed expected associations to cellular response to stress (**Fig. 1b, Sup. Fig 1c, Sup. Table 3)**. These results show that our long-read data accurately and reproducibly capture the expression status of the transcriptome upon oxidative stress.

To test the decay status of RNA for each individual transcript, we quantified the average read length aligning to each transcript ("transcript" refers to annotated loci corresponding to gene isoforms). Surprisingly, we found RNAs in oxidative stress were globally shorter than in unstressed cells, independently of transcript differential expression status, indicating that RNAs in stressed cells were globally shorter than unstressed (**Fig. 1c**). Quantification of all transcripts combined, revealed that RNAs under oxidative stress were significantly shorter by an average of 115.3 nucleotides compared to unstressed (Mann–Whitney: p-value $< 10^{-123}$, **Fig. 1d**). Similar results were observed independently of the transcript coding potential (**Sup. Fig. 1d and 1e**). To exclude potential library preparation artifacts or failure to capture long transcripts by the sequencing device, we repeated the analysis by computationally selecting only reads with ligated 5′ adaptors, ensuring RNA molecules were fully sequenced. The subset of 5′ adaptor-ligated reads confirmed our findings and again showed global RNA shortening upon oxidative stress (**Fig. 1e**). Since both ligated and non-ligated reads showed comparable results, all reads were used in subsequent analysis, unless otherwise mentioned.

To test whether stress induced RNA shortening depends on preexisting steady-state fragmentation levels, we performed analyses in a length-normalized space (meta-length) that represents read length as a percent of annotated transcript length. We calculated normalized meta-length by splitting each annotated transcript in 20 equal bins, assigning the mapped read ends into individual bins, and defining the transcript meta-length as the difference of the 5' and 3'

end bins (**Fig. 1f**). Our data showed a consistent reduction of transcript meta-length upon oxidative stress, independent of the position on the x-axis, indicating that transcripts are being shortened independently of steady-state fragmentation levels (**Fig. 1g, Sup. Fig. 1f**).

To exclude potential effects of the arsenite chemical itself, we further tested oxidative stress induction with 0.3 mM hydrogen peroxide (H_2O_2) for 2 hours followed by TERA-Seq. H_2O_2-induced oxidative stress also resulted in significant transcript shortening (Mann–Whitney: p-value < 10^{-96}), similar to that observed with arsenite, indicating that oxidative stress results in RNA shortening irrespective of inducer (**Sup. Fig. 1g**). To explore whether RNA shortening is unique to oxidative stress or constitutes a general stress response, we re-examined publicly available direct RNA-Seq data for human K562 cells that were subjected to heat shock at 42°C for 60 min (Maier et al., 2020). The results showed that heat-shocked cells also had significantly shorter RNAs compared to unstressed cells (Mann–Whitney: p-value < 10^{-76}), again independently of differential expression (**Fig. 1h, Sup. Fig. 1h, i**).

To identify transcripts with systemic RNA length changes, we developed NanopLen, a statistical tool that uses linear mixed models, and modeled the library as a random effect to adjust for variation across replicates. To test the model, we simulated sequencing data over varying RNA shortening proportions and expression counts. The simulated data showed that the model accurately predicted the true length difference for each tested shortening proportion and had a well-controlled false positive rate (**Fig. 2a**). As expected, NanopLen reported lower significance for smaller length differences while higher expression counts resulted in higher statistical power. We compared unstressed and arsenite-treated cells and identified 2730 significantly shortened transcripts (p-value < 0.05) (**Fig. 2b, Sup. Table 4)**. Gene Ontology analysis of these transcripts highlighted processes relevant to RNA catabolism, translation initiation, and protein localization to the endoplasmic reticulum (**Fig. 2c**).

The presence of the poly(A) tail is required for direct RNA-Seq. Therefore, we anticipated that our method would only capture fragmentation occurring at the 5' end of RNAs. To confirm this, we compared the heterogeneity observed at the 5' end of transcripts and at the terminal templated end upstream of the poly(A) tail. As expected, contrary to non-significant transcripts, significantly shortened transcripts showed that their 3' templated end were almost identical in oxidative stress and unstressed conditions whereas the corresponding 5' ends were more heterogenous (**Fig. 2d, Sup. Fig. 2a, b)**. Quantification at the transcriptome-wide level further confirmed this observation (Mann Whitney p-value < 10^{-256}, **Sup. Fig. 2c**). We reasoned that RNA shortening at the 5' end, should also be reflected in short-read RNA-Seq as reduction in read coverage at the 5' end of RNAs. We thus re-analyzed recent data from HEK293 cells treated with a battery of stressors i.e. 42 °C heat shock for 1 h; 0.6 mM H_2O_2 for 2 h and 300 µM $NaAsO_2$ for 2 h followed by short-read RNA-Seq (Watkins et al., 2022). Our analysis showed that, compared to unstressed, all stress conditions resulted in significantly reduced read coverage at the 5' end (**Fig. 2e, Sup. Fig. 2d**).

Finally, we considered the possibility that selection of alternative transcription start sites (TSSs) downstream of annotated TSSs could be contributing to RNA shortening. We have previously established that positions with high read density upstream of coding sequences in TERA-Seq, accurately represent TSSs matching those defined by CAGE-Seq (Ibrahim et al., 2021). We used

this method to independently calculate TSSs in stress and unstressed conditions. Our results showed non-significant changes between conditions (Mann–Whitney: p-value = 0.35) with no preference towards shorter or longer transcripts (**Fig. 2f**) indicating that the observed shortening is not affected by TSS selection. Conclusively, our results, while not excluding 3' end decay for molecules not sequenced due to lack of poly(A) tails, provide strong evidence for stress-induced decay occurring at the 5' end of RNA molecules that harbor poly(A) tails.

Stress-induced RNA decay is XRN1-mediated but independent of de-adenylation

XRN1 is the primary 5' to 3' exonuclease in cells, and thus a likely candidate for mediating 5' transcript shortening upon stress. We hypothesized that in the absence of XRN1, RNAs would be stabilized, and their observed length would be restored. Silencing of XRN1 with siRNAs (siXRN1) resulted in substantial reduction of protein level compared to non-targeted control (siCTRL) (**Fig. 3a**, **Sup. Fig. 3a**). We subsequently performed TERA-Seq for siCTRL- and siXRN1-transfected cells in the presence or absence of oxidative stress. Our results showed that transfection with siXRN1 resulted in significantly longer transcripts than mock transfection, largely rescuing RNA length (**Fig. 3b, c**). Quantifying this length difference at the transcript level showed that indeed most expressed transcripts were generally longer after XRN1 downregulation, again independently of differential expression (**Fig. 3d, Sup. Fig. 3b**). While silencing of *XRN1* might be expected to have a general effect on RNA length at the 5' end, our results showed that RNA length rescue was predominant for stress-induced significantly shortened RNAs (**Fig 3e**), strongly implicating XRN1 and 5' exonucleolytic decay with transcript shortening during oxidative stress.

The yeast Xrn1 has previously been associated with the decay of oxidized RNAs via the No-Go decay pathway following ribosome stalling at nucleotide adducts, particularly 8-oxoguanosine (8-OxoG), and endonucleolytic cleavage (Yan et al., 2019). We reasoned that if endonucleolytic cleavage induced by 8-OxoG was the major contributor to the transcript shortening, then an increase in guanine prevalence should be expected at the vicinity of 5' ends of sequenced RNAs upon stress. However, our data did not support this explanation, as no G-nucleotide enrichment difference between arsenite and unstressed was observed (**Fig. 4a**). Similarly, no difference was observed upon XRN1 silencing (**Sup. Fig. 1a**). These findings, combined with the presence of shortening under heat shock (**Fig. 1h**), indicates that the stress-induced RNA decay described here is not mediated via 8-OxoG ribosome stalling.

Under traditional decay models, de-adenylation is considered the first and rate-limiting step prior to decapping and subsequent exonucleolytic action from the 5'-end via XRN1 (Borbolis and Syntichaki, 2022; Łabno et al., 2016; Mugridge et al., 2018; Passmore and Coller, 2021). We therefore wished to assess the role of the poly(A) tail in stress-induced RNA decay. We used nanopolish to quantify the poly(A) tail length of RNA molecules in the TERA-Seq libraries (Workman et al., 2019). Contrary to the traditional model, our data showed a modest but significant increase of poly(A) tail length (p-value $< 10^{-143}$) in arsenite-treated compared to unstressed cells (**Fig. 4b**). However, no association with gene expression levels was observed (**Fig. 4c, Sup. Fig. 4b**). Similarly, no correlation was observed between poly(A) tail length and RNA decay levels indicating that shortening of the poly(A) tail was not required for stress-induced RNA decay (**Fig. 4d, e, Sup. Fig. 4c-d**). These finding are in agreement with previous studies where increase of poly(A) tail length in response to endoplasmic reticulum, arsenite, heat,

and nutrient starvation stresses was observed (Tudek et al., 2021; Woo et al., 2018; Yamagishi et al., 2014). Finally, we tested a composite metric for transcript length by adding the poly(A) tail length to the corresponding templated read length. Our data again showed that composite transcript length was reduced in arsenite-treated compared to unstressed cells (**Fig. 4f**). In conclusion, our results show that stress-induced RNA decay is mediated by XRN1 but is independent of prior de-adenylation.

Restoring ribosome density inhibits stress-induced RNA decay

To identify gene features that could be driving stress-induced RNA decay, we tested whether the annotated nucleotide content and transcript length were predictive of transcripts prone to stress-induced 5' shortening. Our results showed a significant association with the annotated coding sequence length and GC content, both being significantly lower for significantly shortened transcripts (**Fig. 4g-i, Sup. Fig. 4e, f**). Since the coding sequence is the primary region of ribosome occupancy, we used publicly available ribosome profiling data from unstressed cells (Park et al., 2016) to test whether pre-stress ribosome density could be defining 5' shortening. Our analysis showed no association between ribosome density per RNA and stress-induced shortening (**Sup. Fig. 5a**, **b**) indicating that the pre-stress ribosome levels on RNAs are not linked to RNA decay under stress.

Rather, another possibility is that stress-induced RNA decay is associated with the rate at which RNAs leave the translation pool following translation initiation inhibition upon stress. If true, then alteration of the ribosome dynamics under stress should have a direct effect on observed RNA decay. To test the role of translation initiation inhibition and ribosome run-off upon stress on stress-induced decay, we treated cells with the ISR Inhibitor (ISRIB) that bypasses the effect of eIF2α phosphorylation and facilitates translation initiation under stress (Rabouw et al., 2019; Sidrauski et al., 2015). We confirmed that ISRIB had no noticeable effect on the dose-dependent phosphorylation of eIF2α but inhibited the SG formation under stress, as expected (**Fig. 5a-c**). Polysome fractionation following ISRIB addition showed a clear reduction of the monosome fraction, indicating partial recovery of translation initiation, as previously described (Sidrauski et al., 2013) (**Fig. 5d**). We then performed TERA-Seq on ISRIB-treated and control cells in the presence or absence of arsenite. Our results showed that compared to cells treated with arsenite only, ISRIB treatment resulted in a significant shift of RNA length towards longer molecules (p-value $< 10^{-156}$) essentially restoring transcript length to the level of the unstressed (**Fig 5e**). Importantly, previously identified significantly shortened transcripts showed the greatest recovery of their length compared to non-significant ones (**Fig 5f**). To test whether this could be simply attributed to a possible length imbalance between up- and down-regulated genes, we again plotted the average transcript length stratified by the differential expression status. Our data show that the observed shift in length is independent of differential expression status (**Fig. 5g, h**). These results indicated that stress-induced RNA decay is associated with RNAs leaving the translation pool following translation initiation inhibition upon stress.

As an alternative to ISRIB which modulates translation initiation, we also treated cells with cycloheximide to interrogate the dynamics of ribosome elongation and to prevent ribosome run-off. Cycloheximide blocks elongation and thus traps ribosomes on RNAs along with translation factors in polysomes, thus decreasing SG formation (Jayabalan et al., 2021; Kedersha and Anderson, 2007; Takahashi et al., 2013). We confirmed that treatment with cycloheximide also

reduced the monosome fraction, although to a lower degree than ISRIB (**Sup. Fig. 5c**). Interestingly, cycloheximide treatment also inhibited RNA decay, particularly for the most significantly shortened RNAs largely rescuing their length (**Fig. 5i-k**). We again did not observe any association between the shift in length and differential expression **(Sup. Fig. 5d, e)**. Combined, our results indicate translation initiation inhibition and the exit of mRNAs from the translational pool as a critical step towards stress-induced decay.

Inhibition of stress-granule formation in cells devoid of G3BP1/2 rescues RNA decay

During stress, ribosome run-off and the exit of RNAs from the translation pool is accompanied by the formation of SGs (Protter and Parker, 2016). To test whether stress-induced decaying transcripts associate with SGs, we used publicly available data representing genes enriched in the SG transcriptome (Khong et al., 2017). As previously described (Khong et al., 2017), we found that SG-enriched genes were generally significantly longer, particularly in their coding and 3' UTR sequences, (**Fig. 6a, Sup. Fig. 6a**) and less expressed than SG-depleted ones (**Sup. Fig. 6b**). However, differential gene length analysis showed, interestingly, a significant difference in gene shortening dependent on SG enrichment status. Specifically, SG-enriched genes were found to be shortened at a significantly higher level upon oxidative stress than SG-depleted or non-localized RNAs for both HeLa and U-2 OS cells (**Fig. 6b, c, Sup. Fig. 6c, d**). Consistent with stress-induced gene shortening being associated with SG enrichment, SG-enriched RNAs harbored lower GC content, similar to significantly shortened transcripts **(Fig. 4i, Fig. 6f**). We again found no association with poly(A) tail length as, although SG-enriched RNAs had significantly longer poly(A) tails, their lengths did not significantly change upon stress (**Fig. 6d, e**). To test whether the observed enrichment of shortened genes is specific to the SGs, we also tested for association with processing bodies (P-bodies), membraneless organelles enriched in de-capping factors and exoribonucleases, essential for RNA decay under normal, non-stress conditions (Brothers et al., 2023). We re-analyzed previously published data from HEK293 cells (Matheny et al., 2019), but in contrast to SGs, we found no association between P-body localization and gene lengths (**Sup. Fig. 6e**). We also did not find an association between P-body localization and RNA shortening which however could be a result of comparing different cell lines (**Fig. 6g**). Our data show that shortened RNAs exhibit similar properties to SG-enriched transcripts and are preferentially enriched in the SG transcriptome.

To test the role of stress granule in stress-induced RNA decay, we employed control (WT) U-2 OS and U-2 OS cells genetically ablated for both G3BP1 and G3BP2 (ΔΔG3BP1/2) via CRISPR/Cas9 (gift from Paul Anderson lab) previously developed and characterized in (Kedersha et al., 2016). ΔΔG3BP1/2 cells inhibit SG condensation in response to arsenite and other stressors. Treatment of WT U-2 OS cells with arsenite showed a clear and global shortening of RNA, comparable to our previous results, confirming that stress-induced decay also occurs in U-2 OS cells (**Fig. 6h, Sup. Fig. 6f**). Interestingly, comparison of arsenite-treated and unstressed ΔΔG3BP1/2 cells that lack the ability to form stress granules showed no RNA length shortening (**Fig. 6h, Sup. Fig. 6g**). Further comparison of ΔΔG3BP1/2 against WT U-2 OS showed that inhibition of G3BP1/2-mediated SG formation fully rescues RNA length to the level of unstressed WT cells (**Fig. 6i, Sup. Fig. 6h, i**). These results indicate that G3BP1/2-mediated SG formation is required for stress-induced RNA decay. Alternatively, SGs could be protective of shortened RNAs that otherwise would be rapidly eliminated from cells. Comparison of differentially expressed transcripts upon arsenite treatment in ΔΔG3BP1/2 cells

showed no association with shortening proportion indicating that in the absence of SGs these RNAs are not rapidly eliminated (**Fig. 6j**), consistent with previous findings (Bley et al., 2015). Collectively our results show that G3BP1/2-dependent SG formation is required for stress-induced RNA decay.

Discussion

Cellular response to stress is critical for cell recovery or induction of apoptosis if stress is unresolved. Recent works have delineated many of the biochemical pathways involved in stress responses. However, the full-length state of RNA under cellular stress remains incompletely characterized. In this work, we used TERA-Seq, a protocol that involves the ligation of unique adaptors at the 5' end of RNAs (Ibrahim et al., 2021) to address a key limitation in nanopore direct RNA sequencing – the inability to consistently capture the 5' end of sequenced RNA molecules. By sequencing RNA molecules end-to-end, we show that stress triggers decay at the 5′ end of RNAs. Interestingly our results show that stress-induced 5' end decay leaves "scars" that can also be identified in traditional short-read RNA-Seq as decreased read density at the 5' end of transcripts.

Our findings highlight XRN1 as an essential component for stress-induced decay, in the absence of which RNA length is largely rescued. However, contrary to the traditional decay model, evaluation of the most significantly shortened RNAs did not reveal a dependency on de-adenylation. Instead, these shortened transcripts possess shorter coding sequences and lower GC content. Previous studies in unstressed cells have shown an association between GC content and mRNA stability in P-bodies, with XRN1-dependent decay preferentially acting on GC-rich mRNAs (Courel et al., 2019). P-bodies and SGs are dynamically linked under stress with shared RNA binding proteins and mRNAs shuttling between the two (Kedersha et al., 2005; Moon et al., 2019). The GC content appears to also be a determinant for RNA recruitment to SGs but, as also confirmed in our study, primarily for GC-poor RNAs (Khong et al., 2017). Consistent with this finding, our data show that the SG transcriptome is significantly enriched for stress-induced shortened RNAs. In fact, our results show that the formation of the SGs is required and indispensable for stress-induced decay as genetic ablation of G3BP1/2 rescues RNA length to the level of non-stressed controls. Finally, we find that inhibition of the ISR to re-initiate translation and dissolve SGs also curtails stress-induced RNA decay, further supporting a critical role for SG formation in stress-induced RNA decay.

The primary function of ISR activation during stress is to inhibit translation initiation and preserve energy for cell recovery (Pakos-Zebrucka et al., 2016). While the physiological role of stress-induced RNA decay is currently unknown, it is intriguing to hypothesize that it evolved as a mechanism to further reduce translational load during stress. Degradation of the 5' end of RNAs is expected to eliminate translation initiation sites thus offering an orthogonal strategy to dial down translation. Future studies to test these hypotheses will be required, especially considering the emerging significance of SGs and RNA metabolism in neurodegenerative diseases and aging (Ferrucci et al., 2022; Ripin and Parker, 2022).

Materials and Methods

Cell culture, treatments, and RNA isolation

HeLa (ATCC CCL-2) and U-2 OS (gift from Paul Anderson lab, described in (Kedersha et al., 2016)) cells were cultured at 37ºC, 5% CO_2, 90% humidity in Dulbecco's Modified Eagle Medium (Thermo Fisher Scientific, Cat # 11965-092) supplemented with 10% heat-inactivated fetal bovine serum (GeminiBio #100-106), 2 mM L-Glutamine. The cells were tested for mycoplasma contamination using universal mycoplasma detection kit (ATCC, Cat # 30-1012K). Cells were treated with 500 µM of sodium arsenite (Sigma, Cat # S7400-100G) for 60 minutes and total RNA was isolated using TRIzol reagent (Invitrogen, Cat# 15596-018) following the manufacturer's recommendation. RNA concentration was quantified using a High Sensitivity (HS) RNA qubit assay (Invitrogen, Cat # Q32852) and Nanodrop ND-1000 (ThermoFisher). RNA integrity was assessed on a 2100 Bioanalyzer (Agilent Technologies) or Qubit RNA IQ assay (ThermoFisher).

siRNA transfection

Cells were transfected with control and gene-specific siRNAs at a final concentration of 20 nM using the Neon transfection system 100 µl kit (ThermoFisher, MPK100) according to the manufacturer's protocol. A commercially available siRNA specifically targeting the human XRN1 was used for XRN1 knockdown, and a non-targeting siRNA was used as control (**Sup. Table 2**). Approximately 5 x 10^5 cells were resuspended in 100 µl of siRNA-R buffer mixture, electroporated (1,005 V, 35 ms, 2 pulses) and immediately transferred to a 100 mm dish containing pre-warmed media. Cells were grown for 72 hours before treatment with indicated reagents and final harvesting for RNA and protein assays.

Direct RNA sequencing

Library preparation was performed using the direct RNA sequencing kit (Oxford Nanopore Technologies, SQK-RNA002) as previously described (Ibrahim et al. 2021) with modifications. A minimum of 75 µg of total RNA was used as starting material to purify poly(A) RNAs using Oligo d(T)25 Magnetic Beads (NEB, Cat # s1419S). 50 pmoles of linker (REL5) containing a 5′-Biotin-PC group and a 3'-OH (Sup. Table. 2) were ligated to the 5′ end of RNAs using T4 RNA ligase (NEB, Cat #M0204S) for 3 hours at 37°C, as previously described (Ibrahim et al., 2021). 500-1000 ng of poly(A) RNA was used for library preparation using SQK-RNA002 sequencing kit (Oxford Nanopore Technologies). The final library was quantified using Qubit 1X dsDNA High Sensitivity (HS) assay kit (ThermoFisher # Q33231) and loaded on FLO-MIN106 or FLO-PRO002 flow cells.

MTS assay

MTS assay was performed using CellTiter 96 AQueous One Solution Cell Proliferation Assay (MTS) (Promega, Cat # G3582). HeLa cells (5×10^3 cells per well) were seeded in a 96-well plate and incubated in humidified 5% CO_2 incubator at 37°C for 24 hours. Following arsenite treatment, the cell viability was examined using MTS assay (100 µl of DMEM and 20 µl MTS incubated for indicated time points) for 10 minutes and incubated in a 5% CO_2 incubator at 37 °C protected from light. Finally, the plate was subjected to shaking for 15 min followed by an optical density measurement (OD) at 590 nm using 1420 Multilabel Counter (Perkin Elmer, VICTOR^{3}V). The viable cells in arsenite treated samples were reported as the percentage of viable cells in control (untreated) samples.

Immunoblotting
Cells were washed with cold PBS and pelleted by centrifugation 5 min at 4 °C 300 *x g*. Whole cell lysates were prepared by adding 1 volume of SDS lysis buffer (60 mM Tris-HCl pH 7.5, 2% SDS, 10% glycerol) supplemented with 1x protease inhibitor cocktail (Roche Diagnostics, Cat # 11836153001) and 1 mM phenylmethylsulfonyl fluoride (PMSF) (Roche, Cat # 10837091001). The lysate was passed 10 times through a 25-gauge needle, heated at 95°C for 20 minutes and cleared by centrifugation at 16,500 *x g* for 10 minutes. Protein concentrations were measured using the Qubit protein assay (Invitrogen, Cat # Q33211). 25 µg of proteins was separated on NuPAGE 4-12% Bis-Tris Gel (Invitrogen, Cat #NP0321BOX) and transferred to PVDF membrane (Millipore, Cat # IPVH00010). After blocking in TBS-T 5% milk for 2 hours, membranes were incubated with primary antibodies overnight at 4°C. Membranes were washed 3 times 5 min in TBS-T and incubated for 2 hours with secondary antibodies. Signals were developed using Chemiluminescence (Azure Biosystems, Cat # AC2204) and acquired on a ChemiDoc MP imaging system (Biorad). For studies of eIF2α phosphorylation (eIF2α-P), the lysis buffer was supplemented with phosphatase inhibitors (Sigma, Cat # Cocktail 2 - P5726 and Cocktail 3-P0044) and 5% BSA was used instead of milk for blocking. Antibodies are listed in Sup. Table 2.

Immunofluorescence
Cells were cultured at 70-80% confluency in Millicell EZ SLIDE 8-well glass (Milipore, Cat # PEZGS0816) in the presence of 50, 100, 250, 500 µM of arsenite +/- 200 nM of ISRIB (Sigma, Cat # SML0843). After the indicated time of treatment, cells were washed 3 times with warm PBS, fixed for 10 min at room temperature (RT) in PBS / 3.7% paraformaldehyde and washed with PBS for 5 min with shaking. After permeabilization in PBS 0.5% Triton X-100 for 10 min at RT, cells were blocked in PBS 0.1% Tween-20 (PBS-T) supplemented with 1% BSA for 1 hour at 37°C with gentle shaking. Primary antibody diluted in 1% BSA were added to the cells and incubated at 4C overnight. Cells were washed in PBS-T three times 5 min with shaking before staining with secondary antibody for 1 hour at RT. Cells were washed four times for 5 min with PBS-T with shaking and nuclei were stained with 4,6-diamidino-2-phenylindole (DAPI) (1:1000 in PBS-T) for 5 min at RT. Slides were mounted with ProLong Glass Antifade Mountant (Invitrogen, Cat # P36982) and images were acquired using a DeltaVision Microscope System (Applied Precision).

Polysome profiling
The polysome profiling was performed as described previously (Panda et al., 2017) with modifications. Briefly, HeLa cells (80-85% confluency) cultured in 100 mm dishes were treated with 500 µM sodium arsenite, 200 nM ISRIB or 25 µg/ml cycloheximide as indicated. DMSO was used as a control. Immediately before harvesting, cells were treated with 100 µg/ml cycloheximide for 10 minutes. After one wash with ice-cold PBS containing 100 µg/ml cycloheximide, 500 µL of polysome extraction buffer (20 mM Tris-HCl pH 7.5, 50 mM NaCl, 50 mM KCl, 5 mM $MgCl_2$, 1mM DTT, 1 X HALT Protease, 1.0% Triton x-100, and 100 µg/ml cycloheximide) was added directly to the plate and lysates were harvested by scraping. The lysate was cleared by centrifugation at 14,000 *x g* for 10 min. The supernatant was loaded onto the 10-50% sucrose gradient followed by high-speed centrifugation (260,800 *x g* for 90 minutes

at 4°C). Using a density gradient fractionation system monitored by UV absorbance detector (A254), 12 fractions were collected.

Nanopore sequencing data processing
Nanopore sequencing data were basecalled using Guppy (v3.4.5). The 5′ adaptors were identified and removed using cutadapt (v2.8). Reads were first aligned against ribosomal sequences obtained from SILVA (Quast et al., 2013). Non-ribosomal reads were subsequently mapped against the human genome hg38 using minimap2 (version 2.17) (Li, 2018) and parameters -a -x splice -k 12 -u b -p 1 --secondary=yes. They were also aligned against the human transcriptome using -a -x map-ont -k 12 -u f -p 1 --secondary=yes.

Poly(A) tail length estimation
The poly(A) tail lengths were extracted from sequenced reads using the nanopolish polya package (Workman et al., 2019). Only the poly(A) tail lengths which passed the software quality control scores and were tagged as "PASS" were used in our analysis.

Differential gene expression and Gene Ontology
Transcript counts were quantified from the transcriptome alignments as the total number of reads aligning on each transcript. Differential expression analysis was performed using DESeq2 (Love et al., 2014). EnhancedVolcano (https://github.com/kevinblighe/EnhancedVolcano) was used for visualization. Gene ontology analysis was performed for significantly changed genes (p-value< 0.05) using "clusterProfiler" (Yu et al., 2012).

Ribosome profiling data
Ribosome profiling sequencing data were downloaded from GEO GSM2100595 (Park et al., 2016). Adaptors were removed using cutadapt (version 2.8) and reads were aligned to the human genome using STAR (2.5.3a). Further processing to calculate counts per transcript was performed with in-house scripts using pysam (v0.15.4) and counts were converted to reads per kilobase of transcript per million (RPKM).

Meta-length calculation
For calculating the bin (meta) lengths, we divided the transcripts into 20 equal bins [0, 19]. Reads were assigned to each of these bins based on the corresponding location of their ends. The binned- (meta-) length is calculated as the difference of the binned 3' end over the binned 5' end.

Transcription start site identification
Transcription start site identification was performed as previously described (Ibrahim et al., 2021). Briefly, all reads from all replicates were combined and the distribution of read 5' ends within the 5' UTR was quantified. The TSS was selected as the position with the highest read density withing the 5' UTR with minimum 5 supporting reads.

NanopLen
NanopLen reads a file of read lengths with library identifiers and gene or transcript identifiers, and a metadata file describing the experimental design. The latter includes the library identifiers and the corresponding condition for each library. The software supports three models: t-test,

Wilcoxon, and linear mixed model (LMM). In the t-test and LMM options, the user can also supply a customized model to use extra variables in the metadata. In this work we have used the LMM to adjust for putative batch effects across libraries.

Given a vector Y of read lengths associated with a gene/transcript, the t-test is functionally equivalent to a linear regression model. The model also supports optional read-specific extra covariates but are not used in this work:

$Y = \beta_0 + \beta_{cond} cond + \epsilon +$ (extra covariates).

Similarly, the LMM models Y using the condition as fixed effect and the library identifier as a random effect.

$Y = \beta_0 + \beta_{cond} cond + lib + \epsilon +$ (extra covariates),

where *lib* is the random effect of each library. By adding the random effect, the analysis is more robust against potential false positives from library variation and prevents deeper libraries from dominating the effect size. A Wald test is used to test for significance of β_{cond}. The Wilcoxon test option is restricted to only testing the condition variable and cannot adjust for additional variables. Each gene/transcript is assigned a statistic corresponding to the test used and a p-value. p-values are adjusted for multiple testing using the Bonferroni-Hochberg method.

To test NanopLen models, transcript length data with known shortening rates were simulated. Each simulation was parameterized with a "true" transcript length, a read count as a proxy for transcript expression, and a shortening proportion. The database of known human genic lengths from Ensembl (release 91) was used and the median was selected as "true" length. Since the simulated variance is in proportion of the true length, using any length will result in similar results so we do not simulate additional true lengths. Expression counts were simulated, ranging from 10 to 200 reads per transcripts to capture the dynamic expression range particularly towards small read depth. Varying shortening percentages were simulated from 50, 60, .., 100% of the true length (equivalent of -1, -0.74, .., 0 log2 fold change) with the 100% corresponding to the null simulation of no-change.

Given the parameters, the simulated lengths were generated based on the mixed model in NanopLen, with the number of sampled lengths being the selected expression counts. The library random effect was simulated as if imitating a fluctuation with standard deviation of 10% of the true length, and the global error having a standard deviation of 20% of the expected length mean. As with the real experiment, six libraries of three control and three condition were simulated, with 1000 genes per scenario. Using the LMM option with the logscale option selected, the resulting log2FCs and p-values were calculated.

NanopLen is available in https://github.com/maragkakislab/nanoplen

IGV read density plots

Transcripts that were identified as significantly shortened at a false discovery rate of 0.05 by NanopLen, had at least 50 supporting reads in both conditions and had a difference of 200 nts in

length were selected. From the list of 32 transcripts, 5 were randomly selected (*ASCC3*, *FASTKD2*, *MTRR, HNRNPM* and *DSG2*) for visualization. As control, 3 transcripts with non-significant, lower than 10 nt difference (*AGPAT2*, *EPHX1*, *TBCA*) were selected. Libraries were randomly downsampled to maximum 50 reads per window and the libraries were overlayed in Affinity Designer. All reads were used, irrespective of adaptor ligation status.

Data availability

Sequencing data have been deposited in the Gene Expression Omnibus (GEO); accession: GSE204785.

Acknowledgements

We would like to thank Dr. Paul Anderson for his gift of the ΔΔG3BP1/2 and WT U-2 OS cells. We thank Dr. David Schlessinger for providing feedback on our manuscript. We thank the NIH Fellows Editorial Board for editorial assistance. This research was supported by the Intramural Research Program of the National Institute on Aging, National Institutes of Health and funded by grant NIH ZIA AG000696 to MM.

Contributions

SM performed wet-lab experiments with assistance from MJP, CB and JLM. SAD performed computational analysis with help from VM, CTL, JMS and AK. CTL and SAD developed NanopLen. SAD, SM and MM interpreted the results and wrote the manuscript.

Declaration of interests

The authors declare no competing interests.

References

Anderson P, Kedersha N. 2008. Stress granules: the Tao of RNA triage. *Trends Biochem Sci* **33**:141–150.

Beckman KB, Ames BN. 1998. The free radical theory of aging matures. *Physiol Rev* **78**:547–581.

Bley N, Lederer M, Pfalz B, Reinke C, Fuchs T, Glaß M, Möller B, Hüttelmaier S. 2015. Stress granules are dispensable for mRNA stabilization during cellular stress. *Nucleic Acids Res* **43**:e26.

Borbolis F, Syntichaki P. 2022. Biological implications of decapping: beyond bulk mRNA decay. *FEBS J* **289**:1457–1475.

Bresson S, Shchepachev V, Spanos C, Turowski TW, Rappsilber J, Tollervey D. 2020. Stress-Induced Translation Inhibition through Rapid Displacement of Scanning Initiation Factors. *Molecular Cell* **0**. doi:10.1016/j.molcel.2020.09.021

Brothers WR, Ali F, Kajjo S, Fabian MR. 2023. The EDC4-XRN1 interaction controls P-body dynamics to link mRNA decapping with decay. *EMBO J* e113933.

Courel M, Clément Y, Bossevain C, Foretek D, Vidal Cruchez O, Yi Z, Bénard M, Benassy M-N, Kress M, Vindry C, Ernoult-Lange M, Antoniewski C, Morillon A, Brest P, Hubstenberger A, Roest Crollius H, Standart N, Weil D. 2019. GC content shapes mRNA storage and decay in human cells. *Elife* **8**. doi:10.7554/eLife.49708

Derisbourg MJ, Hartman MD, Denzel MS. 2021. Perspective: Modulating the integrated stress response to slow aging and ameliorate age-related pathology. *Nat Aging* **1**:760–768.

Ferrucci L, Wilson DM 3rd, Donegà S, Gorospe M. 2022. The energy-splicing resilience axis hypothesis of aging. *Nat Aging* **2**:182–185.

Galluzzi L, Yamazaki T, Kroemer G. 2018. Linking cellular stress responses to systemic homeostasis. *Nat Rev Mol Cell Biol* **19**:731–745.

Gowrishankar G, Winzen R, Dittrich-Breiholz O, Redich N, Kracht M, Holtmann H. 2006. Inhibition of mRNA deadenylation and degradation by different types of cell stress. *Biol Chem* **387**:323–327.

Haigis MC, Yankner BA. 2010. The aging stress response. *Mol Cell* **40**:333–344.

Harding HP, Zhang Y, Zeng H, Novoa I, Lu PD, Calfon M, Sadri N, Yun C, Popko B, Paules R, Stojdl DF, Bell JC, Hettmann T, Leiden JM, Ron D. 2003. An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. *Mol Cell* **11**:619–633.

Hilgers V, Teixeira D, Parker R. 2006. Translation-independent inhibition of mRNA deadenylation during stress in Saccharomyces cerevisiae. *RNA* **12**:1835–1845.

Hinnebusch AG. 1994. Translational control of GCN4: an in vivo barometer of initiation-factor activity. *Trends Biochem Sci* **19**:409–414.

Horvathova I, Voigt F, Kotrys AV, Zhan Y, Artus-Revel CG, Eglinger J, Stadler MB, Giorgetti L, Chao JA. 2017. The Dynamics of mRNA Turnover Revealed by Single-Molecule Imaging in Single Cells. *Mol Cell* **68**:615-625.e9.

Ibrahim F, Maragkakis M, Alexiou P, Mourelatos Z. 2018. Ribothrypsis, a novel process of canonical mRNA decay, mediates ribosome-phased mRNA endonucleolysis. *Nat Struct Mol Biol* **25**:302–310.

Ibrahim F, Oppelt J, Maragkakis M, Mourelatos Z. 2021. TERA-Seq: true end-to-end sequencing of native RNA molecules for transcriptome characterization. *Nucleic Acids Res* **49**:e115.

Jayabalan AK, Adivarahan S, Koppula A, Abraham R, Batish M, Zenklusen D, Griffin DE, Leung AKL. 2021. Stress granule formation, disassembly, and composition are regulated by alphavirus ADP-ribosylhydrolase activity. *Proc Natl Acad Sci U S A* **118**:e2021719118.

Kedersha N, Anderson P. 2007. Mammalian stress granules and processing bodies. *Methods Enzymol* **431**:61–81.

Kedersha N, Panas MD, Achorn CA, Lyons S, Tisdale S, Hickman T, Thomas M, Lieberman J, McInerney GM, Ivanov P, Anderson P. 2016. G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. *J Cell Biol* **212**:845–860.

Kedersha N, Stoecklin G, Ayodele M, Yacono P, Lykke-Andersen J, Fritzler MJ, Scheuner D, Kaufman RJ, Golan DE, Anderson P. 2005. Stress granules and processing bodies are dynamically linked sites of mRNP remodeling. *J Cell Biol* **169**:871–884.

Kharel P, Fay M, Manasova EV, Anderson PJ, Kurkin AV, Guo JU, Ivanov P. 2023. Stress promotes RNA G-quadruplex folding in human cells. *Nat Commun* **14**:205.

Khong A, Matheny T, Jain S, Mitchell SF, Wheeler JR, Parker R. 2017. The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. *Mol Cell* **68**:808-820.e5.

Łabno A, Tomecki R, Dziembowski A. 2016. Cytoplasmic RNA decay pathways - Enzymes and mechanisms. *Biochim Biophys Acta* **1863**:3125–3147.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**:3094–3100.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:550.

Maier KC, Gressel S, Cramer P, Schwalb B. 2020. Native molecule sequencing by nano-ID reveals synthesis and stability of RNA isoforms. *Genome Res* **30**:1332–1344.

Maragkakis M, Malla S, Hatzoglou M, Trifunovic A, Glick AB, Finkel T, Longo VD, Kaushik S, Muñoz-Cánoves, Pura LGJ, Naidoo N, Booth LN, Payea MJ, Herman AB, de Cabo R, Wilson DM, Ferrucci L, Gorospe M. 2022. Biology of Stress Responses in Aging. *Aging Biology*.

Marcelo A, Koppenol R, de Almeida LP, Matos CA, Nóbrega C. 2021. Stress granules, RNA-binding proteins and polyglutamine diseases: too much aggregation? *Cell Death Dis* **12**:592.

Matheny T, Rao BS, Parker R. 2019. Transcriptome-Wide Comparison of Stress Granules and P-Bodies Reveals that Translation Plays a Major Role in RNA Partitioning. *Mol Cell Biol* **39**. doi:10.1128/MCB.00313-19

Moon SL, Morisaki T, Khong A, Lyon K, Parker R, Stasevich TJ. 2019. Multicolour single-molecule tracking of mRNA interactions with RNP granules. *Nat Cell Biol* **21**:162–168.

Mugridge JS, Coller J, Gross JD. 2018. Structural and molecular mechanisms for the control of eukaryotic 5'-3' mRNA decay. *Nat Struct Mol Biol* **25**:1077–1085.

Pakos-Zebrucka K, Koryga I, Mnich K, Ljujic M, Samali A, Gorman AM. 2016. The integrated stress response. *EMBO Rep* **17**:1374–1395.

Panas MD, Ivanov P, Anderson P. 2016. Mechanistic insights into mammalian stress granule dynamics. *J Cell Biol* **215**:313–323.

Panda AC, Martindale JL, Gorospe M. 2017. Polysome Fractionation to Analyze mRNA Distribution Profiles. *Bio Protoc* **7**. doi:10.21769/BioProtoc.2126

Park J-E, Yi H, Kim Y, Chang H, Kim VN. 2016. Regulation of poly(A) tail and translation during the somatic cell cycle. *Mol Cell* **62**:462–471.
Passmore LA, Coller J. 2021. Roles of mRNA poly(A) tails in regulation of eukaryotic gene expression. *Nat Rev Mol Cell Biol*. doi:10.1038/s41580-021-00417-y
Payea MJ, Dar SA, Malla S, Maragkakis M. 2023. Ribonucleic Acid-Mediated Control of Protein Translation Under Stress. *Antioxid Redox Signal* **39**:374–389.
Pelechano V, Wei W, Steinmetz LM. 2015. Widespread Co-translational RNA Decay Reveals Ribosome Dynamics. *Cell* **161**:1400–1412.
Protter DSW, Parker R. 2016. Principles and Properties of Stress Granules. *Trends Cell Biol* **26**:668–679.
Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**:D590-6.
Rabouw HH, Langereis MA, Anand AA, Visser LJ, de Groot RJ, Walter P, van Kuppeveld FJM. 2019. Small molecule ISRIB suppresses the integrated stress response within a defined window of activation. *Proc Natl Acad Sci U S A* **116**:2097–2102.
Ripin N, Parker R. 2022. Are stress granules the RNA analogs of misfolded protein aggregates? *RNA* **28**:67–75.
Sidrauski C, Acosta-Alvear D, Khoutorsky A, Vedantham P, Hearn BR, Li H, Gamache K, Gallagher CM, Ang KK-H, Wilson C, Okreglak V, Ashkenazi A, Hann B, Nader K, Arkin MR, Renslo AR, Sonenberg N, Walter P. 2013. Pharmacological brake-release of mRNA translation enhances cognitive memory. *Elife* **2**:e00498.
Sidrauski C, McGeachy AM, Ingolia NT, Walter P. 2015. The small molecule ISRIB reverses the effects of eIF2α phosphorylation on translation and stress granule assembly. *Elife* **4**. doi:10.7554/eLife.05033
Stadtman ER, Berlett BS. 1997. Reactive oxygen-mediated protein oxidation in aging and disease. *Chem Res Toxicol* **10**:485–494.
Takahashi M, Higuchi M, Matsuki H, Yoshita M, Ohsawa T, Oie M, Fujii M. 2013. Stress granules inhibit apoptosis by reducing reactive oxygen species production. *Mol Cell Biol* **33**:815–829.
Tudek A, Krawczyk PS, Mroczek S, Tomecki R, Turtola M, Matylla-Kulińska K, Jensen TH, Dziembowski A. 2021. Global view on the metabolism of RNA poly(A) tails in yeast Saccharomyces cerevisiae. *Nat Commun* **12**:4951.
Watkins CP, Zhang W, Wylder AC, Katanski CD, Pan T. 2022. A multiplex platform for small RNA sequencing elucidates multifaceted tRNA stress response and translational regulation. *Nat Commun* **13**:2491.
Woo YM, Kwak Y, Namkoong S, Kristjánsdóttir K, Lee SH, Lee JH, Kwak H. 2018. TED-seq identifies the dynamics of poly(A) length during ER stress. *Cell Rep* **24**:3630-3641.e7.
Workman RE, Tang AD, Tang PS, Jain M, Tyson JR, Razaghi R, Zuzarte PC, Gilpatrick T, Payne A, Quick J, Sadowski N, Holmes N, de Jesus JG, Jones KL, Soulette CM, Snutch TP, Loman N, Paten B, Loose M, Simpson JT, Olsen HE, Brooks AN, Akeson M, Timp W. 2019. Nanopore native RNA sequencing of a human poly(A) transcriptome. *Nat Methods* **16**:1297–1305.
Yamagishi R, Hosoda N, Hoshino S-I. 2014. Arsenite inhibits mRNA deadenylation through proteolytic degradation of Tob and Pan3. *Biochem Biophys Res Commun* **455**:323–331.

Yan LL, Simms CL, McLoughlin F, Vierstra RD, Zaher HS. 2019. Oxidation and alkylation stresses activate ribosome-quality control. *Nat Commun* **10**:5611.
Yu G, Wang L-G, Han Y, He Q-Y. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**:284–287.

Figures

Figure 1

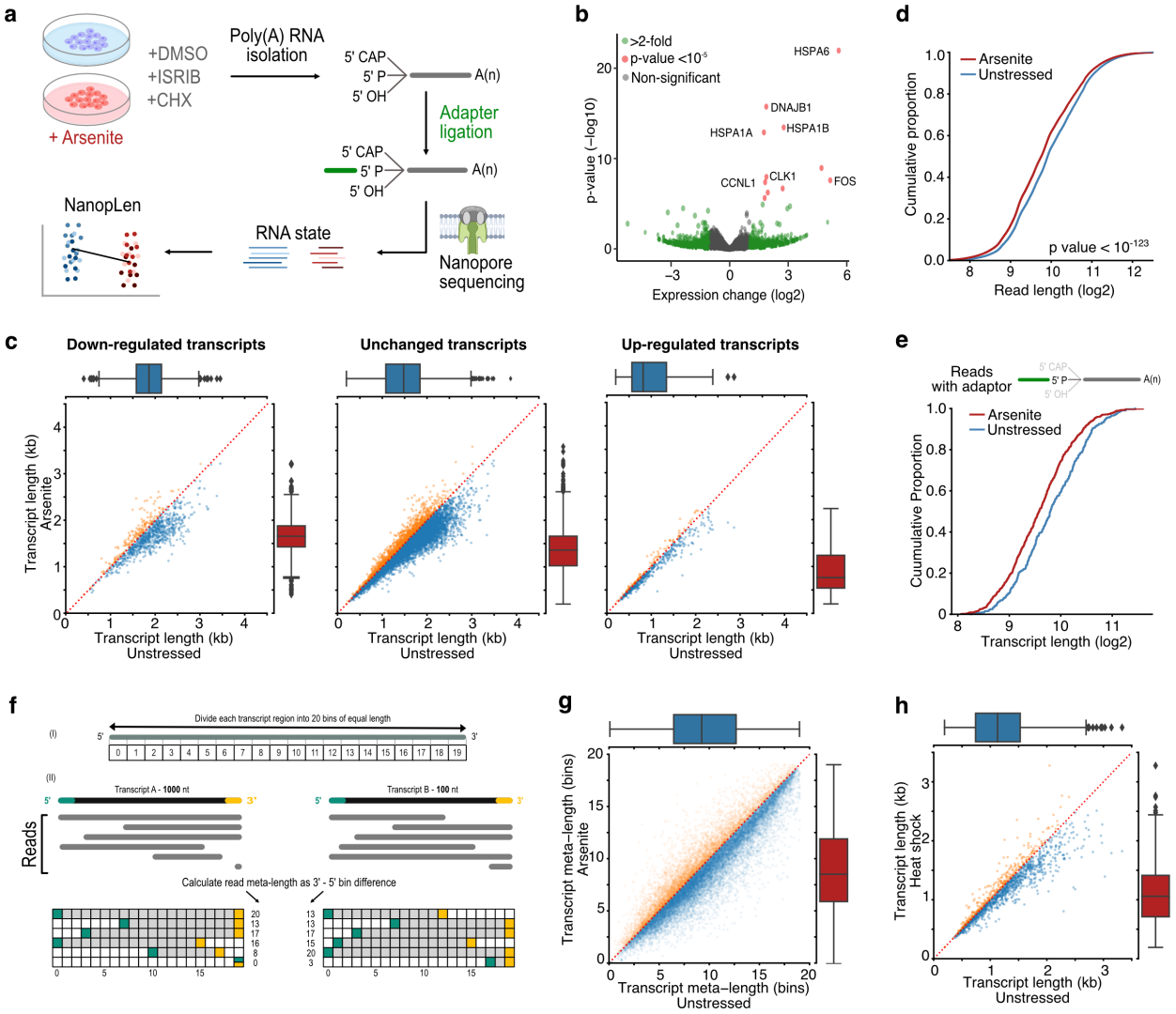


Fig 1: RNA shortening upon cellular stress. a) Schematic of experimental design **b)** Volcano plot for the differential expression of arsenite-treated and unstressed cells. Green color indicates genes with more than 2-fold difference and red indicates statistical significance higher than 10^{-5}. **c)** Scatter plots of average transcript length for arsenite-treated and unstressed cells stratified by their differential expression change. Down-regulated: (-Inf, -0.5), Unchanged: (-0.5, 0.5), Up-regulated (0.5, Inf) fold-change. Only transcripts with at least 5 aligned reads are shown. Red dotted indicates the y=x line. Color indicates transcripts below (blue) and above (orange) the diagonal. **d)** Cumulative distribution of read length for arsenite-treated and unstressed cells. **e)** Cumulative distribution of transcript length for arsenite-treated and unstressed cells using only reads with adaptor ligated at the 5' end. **f)** Schematic of read meta-length calculation. Each annotated transcript is divided into 20 equally sized bins. Each read is then assigned meta-coordinates depending on the bin in which its 5' and 3' templated ends align. The read meta-length is calculated as the difference of the meta-coordinates and presented as a percentage of full length. **g)** Scatter plot of average transcript meta-length as a percentage of full-length for

arsenite-treated and unstressed cells. Coloring is the same as (d) **h)** Scatter plot of average transcript length for heat shock and unstressed cells.

Figure 2

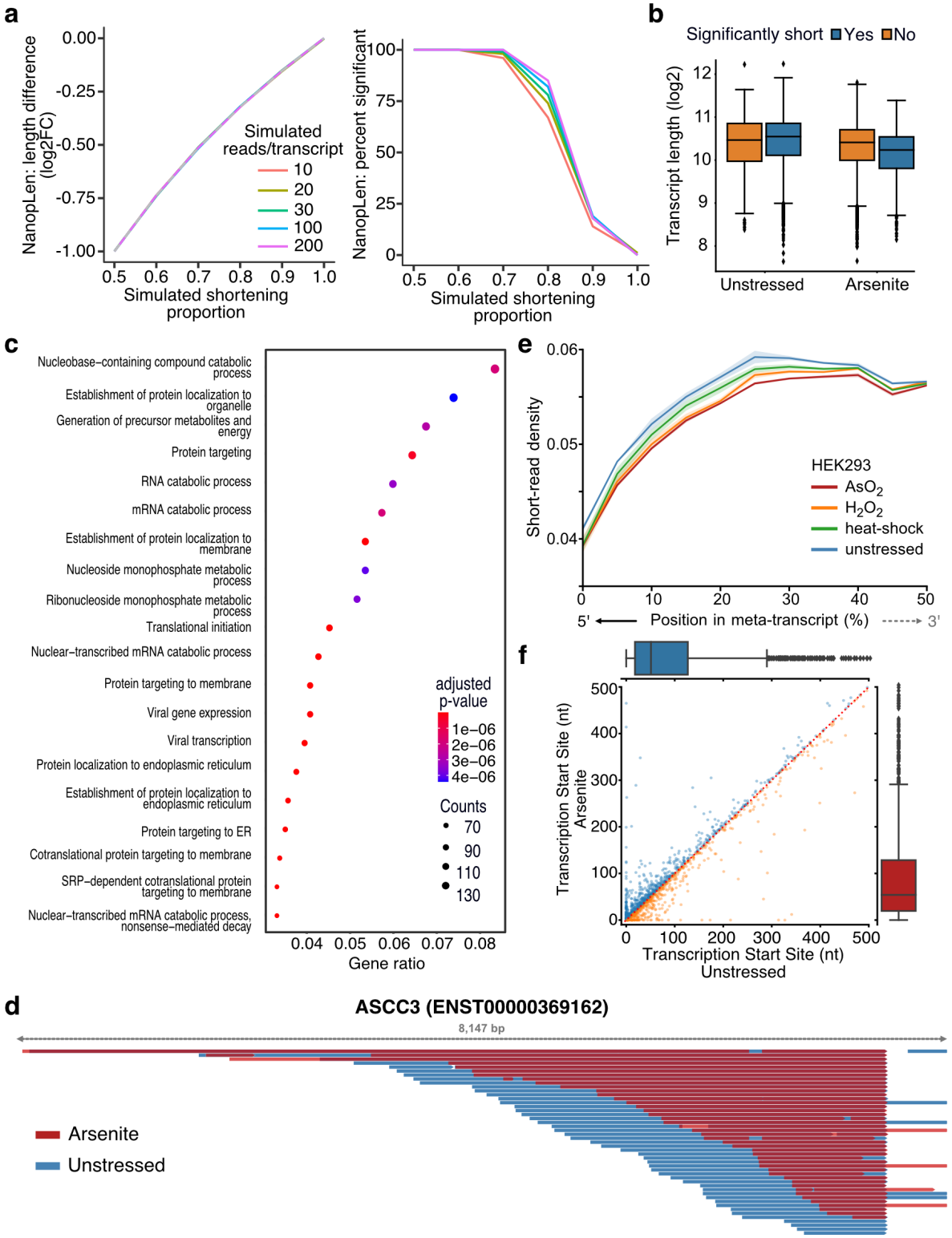


Fig. 2: Characterization of stress-induced shortened RNAs identified by NanopLen: a) Left: Line plot of average length difference estimate for simulated data of varying read depths. True value is depicted by dashed grey line. Right: Percentage of simulated genes that are detected as

significantly different in length. The null simulation of no shortening corresponds to simulated shortening proportion 1.0. **b)** Box plot of average transcript length for statistically significant and non-significantly shortened transcripts identified through differential length analysis in arsenite-treated and unstressed cells using NanopLen. **c)** Gene Ontology analysis for biological processes of significantly shortened transcripts. **d)** IGV screenshot of ASCC3 aligned reads for arsenite-treated (red) and unstressed cells (blue). Libraries were randomly downsampled to maximum 50 reads per window and the libraries were overlayed. All reads were used, irrespective of adaptor ligation status. **e)** Short-read RNA-Seq density at the 5' half of transcripts for unstressed, AsO_2, H_2O_2 and heat shock -treated cells. Shade indicates standard error of the mean for replicates. **f)** Scatter plot of Transcription Start Site position for arsenite-treated and unstressed cells

Figure 3

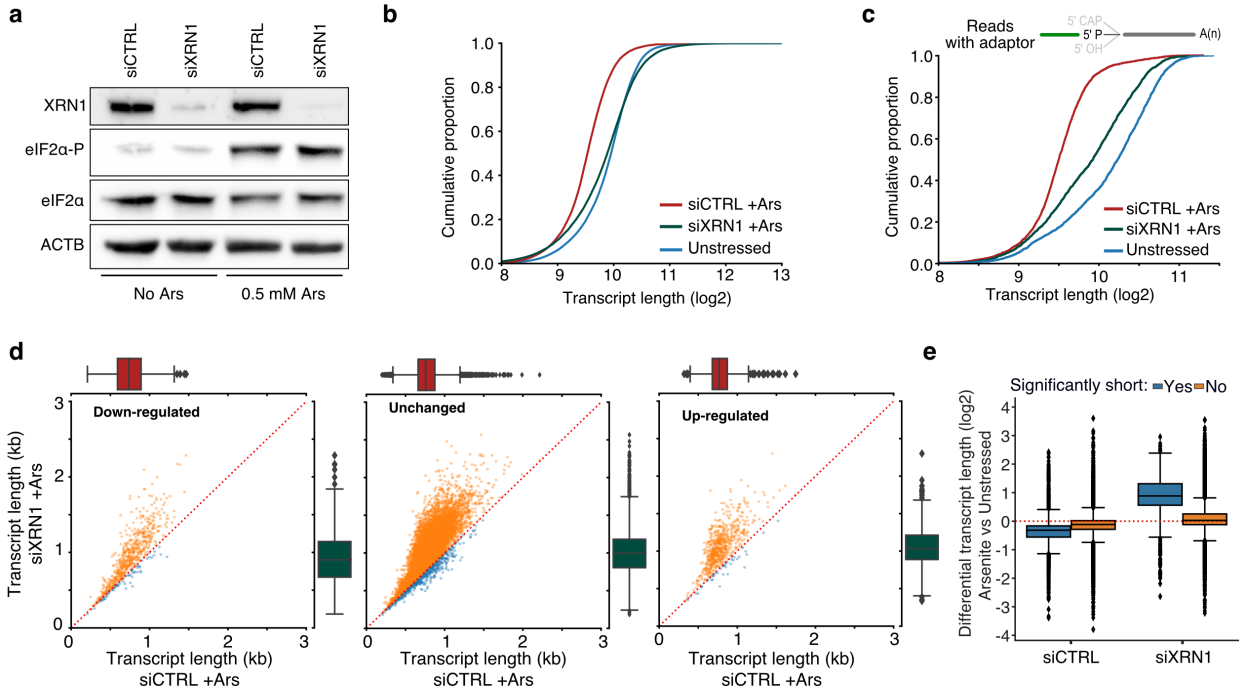


Fig. 3: The effect of XRN1 knockdown on RNA shortening. a) Immunoblot for XRN1, eIF2α-P and eIF2α for cells transfected with a mock (siCTRL) or *XRN1*-targeting siRNA. ACTB is used as control. **b)** Cumulative distribution plot of transcript length for XRN1 knockdown (siXRN1) and control (siCTRL) in arsenite-treated cells. **c)** Same as (b) but only reads with ligated 5' end adaptor are used. **d**) Scatter plots of average transcript length for arsenite-treated cells transfected with mock and *XRN1*-targeting siRNA stratified by their differential expression change. Down-regulated: (-Inf, -0.5), Unchanged: (-0.5, 0.5), Up-regulated (0.5, Inf) fold-change. Only transcripts with at least 5 aligned reads are shown. Red dotted indicates the y=x line. Color indicates transcripts below (blue) and above (orange) the diagonal. **e)** Box plots of differential transcript length in arsenite-treated versus unstressed cells for significantly and non-significantly shortened transcripts in XRN1 knockdown and control.

Figure 4

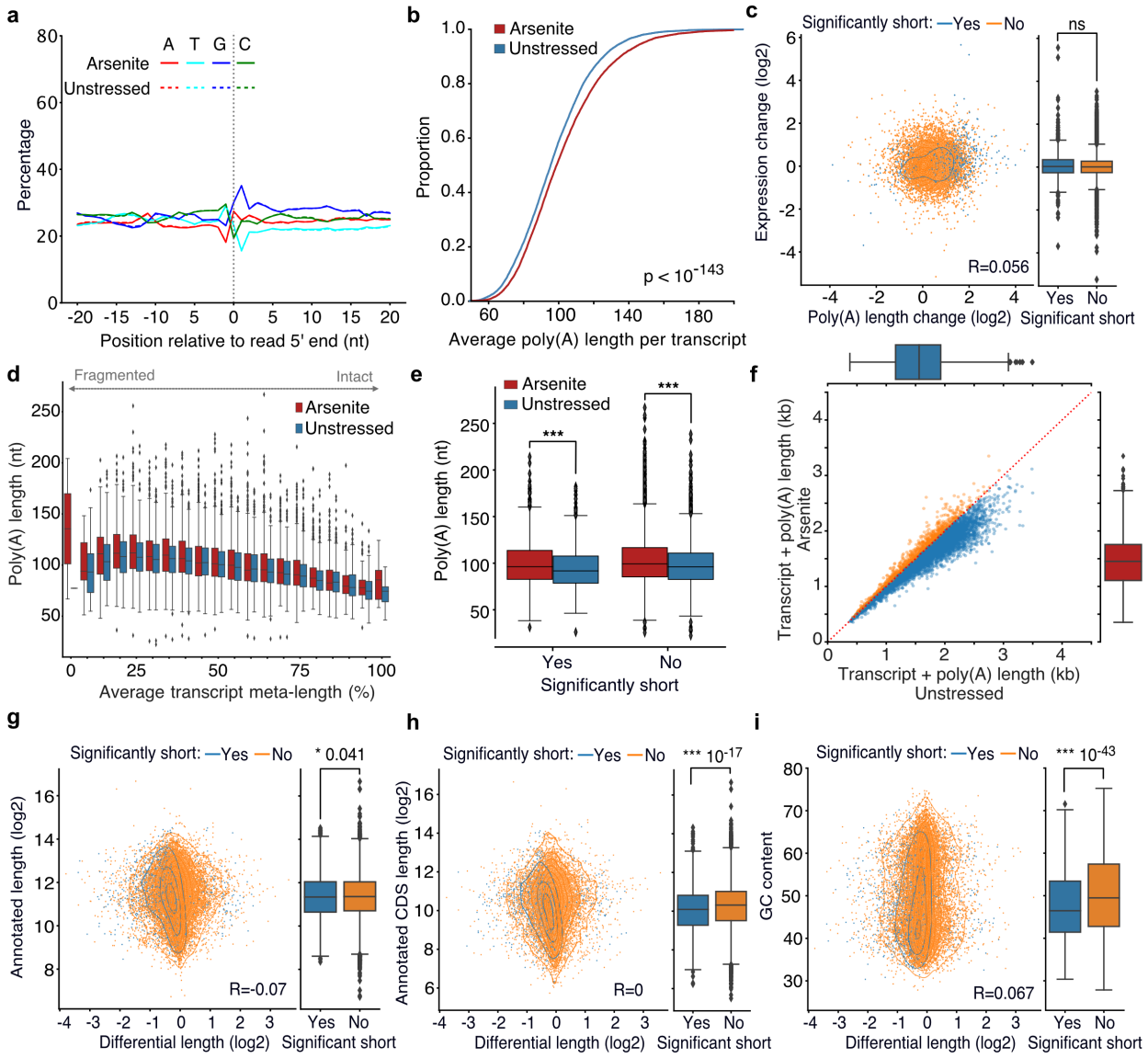


Fig. 4: Association of poly(A) tail length and cis-regulatory elements with RNA shortening. **a)** Nucleotide composition around the 5′ end of reads in arsenite-treated and unstressed cells. All reads were used, irrespective of adaptor ligation status. **b)** Cumulative distribution of the average poly(A) tail length per transcript for arsenite-treated and unstressed cells. **c)** Scatter plot of transcript expression against poly(A) tail length differential change in arsenite-treated and unstressed cells. **d)** Box plots of average transcript poly(A) tail length against average transcript-meta length for arsenite-treated and unstressed cells. **e)** Box plots of average transcript poly(A) tail length for significantly and non-significantly shortened transcripts in arsenite-treated and unstressed cells. **f)** Scatter plot of combined average transcript and poly(A) tail length for arsenite-treated and unstressed cells. Only transcripts with at least 5 aligned reads were used. Red dotted indicates the y=x line. Color indicates transcripts below (blue) and above (orange) the diagonal. **g-i)** Scatter plot of annotated transcript length (g), annotated coding sequence (CDS) length (h) and GC content (i) against transcript differential length in arsenite-treated and

unstressed cells. The box plots on the right side summarize the y-axis variable for significant and non-significantly shortened transcripts. The Pearson's correlation coefficient and the Mann-Whitney-U test p-value are shown.

Figure 5

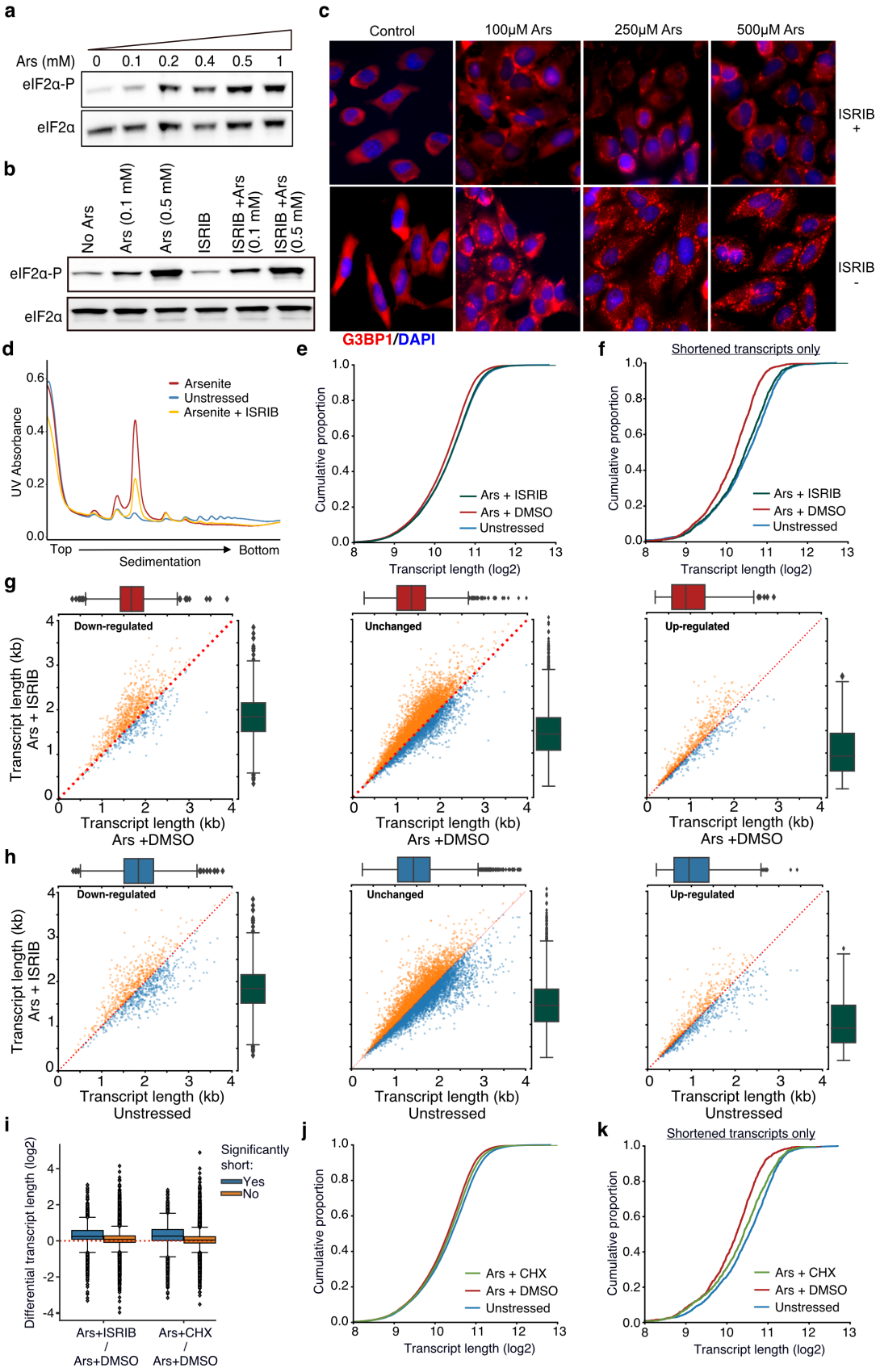

a
Ars (mM) 0 0.1 0.2 0.4 0.5 1
eIF2α-P
eIF2α
b
No Ars
Ars (0.1 mM)
Ars (0.5 mM)
ISRIB
ISRIB +Ars (0.1 mM)
ISRIB +Ars (0.5 mM)
eIF2α-P
eIF2α
c
Control
100μM Ars
250μM Ars
500μM Ars
ISRIB +
ISRIB -
G3BP1/DAPI
d
Arsenite
Unstressed
Arsenite + ISRIB
UV Absorbance
Top
Sedimentation
Bottom
e
Cumulative proportion
Ars + ISRIB
Ars + DMSO
Unstressed
Transcript length (log2)
f
Shortened transcripts only
Cumulative proportion
Ars + ISRIB
Ars + DMSO
Unstressed
Transcript length (log2)
g
Down-regulated
Unchanged
Up-regulated
Transcript length (kb) Ars + ISRIB
Transcript length (kb) Ars +DMSO
h
Down-regulated
Unchanged
Up-regulated
Transcript length (kb) Ars + ISRIB
Transcript length (kb) Unstressed
i
Differential transcript length (log2)
Significantly short:
Yes
No
Ars+ISRIB / Ars+DMSO
Ars+CHX / Ars+DMSO
j
Cumulative proportion
Ars + CHX
Ars + DMSO
Unstressed
Transcript length (log2)
k
Shortened transcripts only
Cumulative proportion
Ars + CHX
Ars + DMSO
Unstressed
Transcript length (log2)

Fig. 5. Translation and RNA shortening. a, b) Immunoblot for eIF2α and eIF2α-P upon increasing concentration of arsenite (a) and cell treatment with 200 nM of ISRIB at different arsenite concentrations (b). **c)** SGs visualized by immunofluorescence of HeLa cells treated with indicated concentration of arsenite in the presence or absence of 200 nM ISRIB. A secondary goat anti-rabbit IgG H&L (Alexa Fluor 594) against G3BP1 (SG marker) and DAPI were used for visualization. **d)** Ribosome sedimentation curve following cell treatment with ISRIB (200 nM) for arsenite-treated and unstressed cells. **e)** Cumulative density plot of transcript length for arsenite-treated cells in the presence or absence of ISRIB. **f)** Same as (e) for significantly shortened transcripts only. **g)** Scatter plots of average transcript length for arsenite-treated cells with and without ISRIB stratified by their differential expression change. Down-regulated: (-Inf, -0.5), Unchanged: (-0.5, 0.5), Up-regulated (0.5, Inf) fold-change. Only transcripts with at least 5 aligned reads are used. Red dotted indicates the y=x line. Color indicates transcripts below (blue) and above (orange) the diagonal. **h)** Same as (g) for arsenite and ISRIB treated cells against control. **i)** Box plots of differential transcript length for comparisons indicated on the x-axis. **i-k)** Same as (e) and (f) for cycloheximide (CHX) instead of ISRIB.

Figure 6

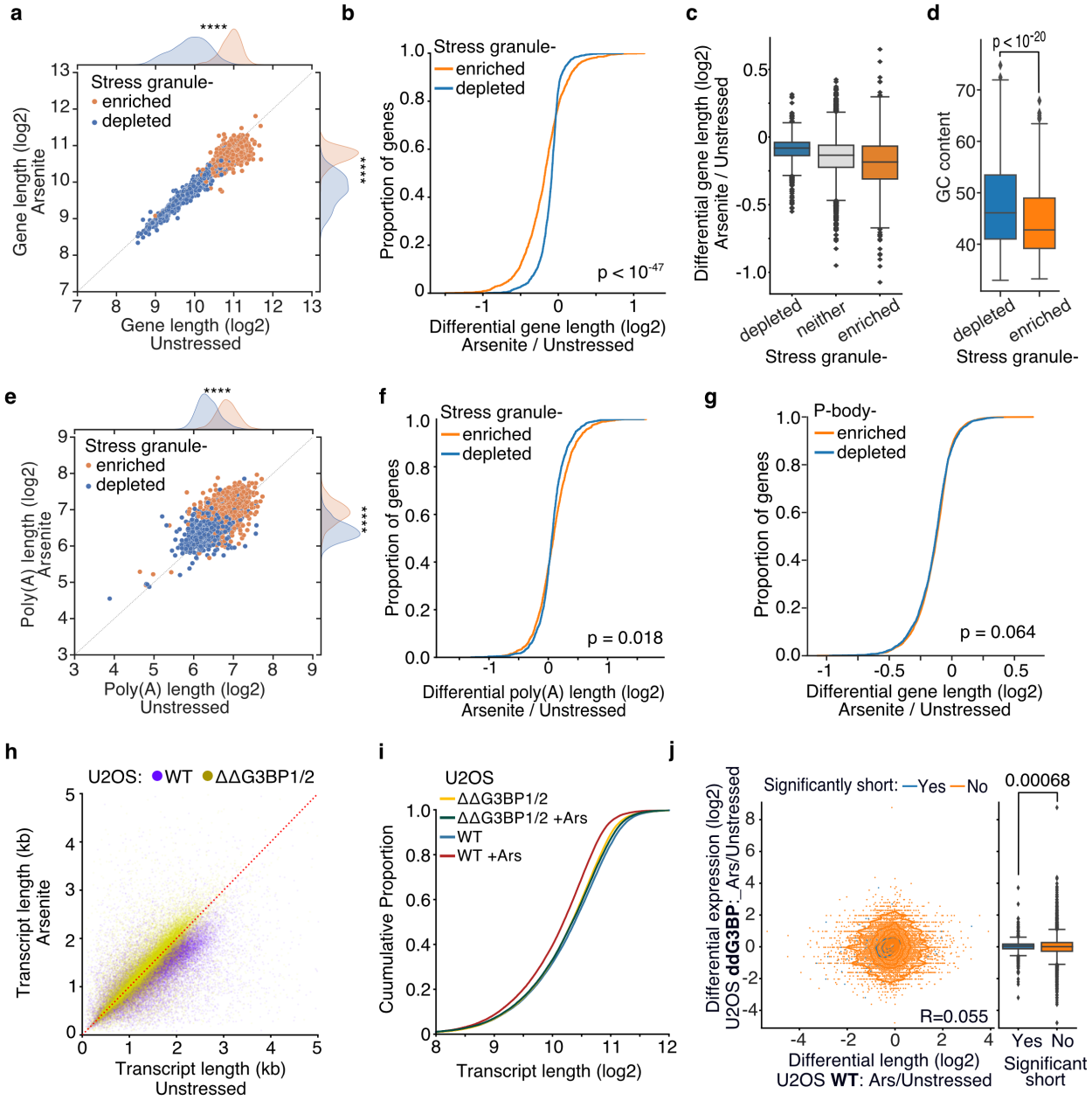
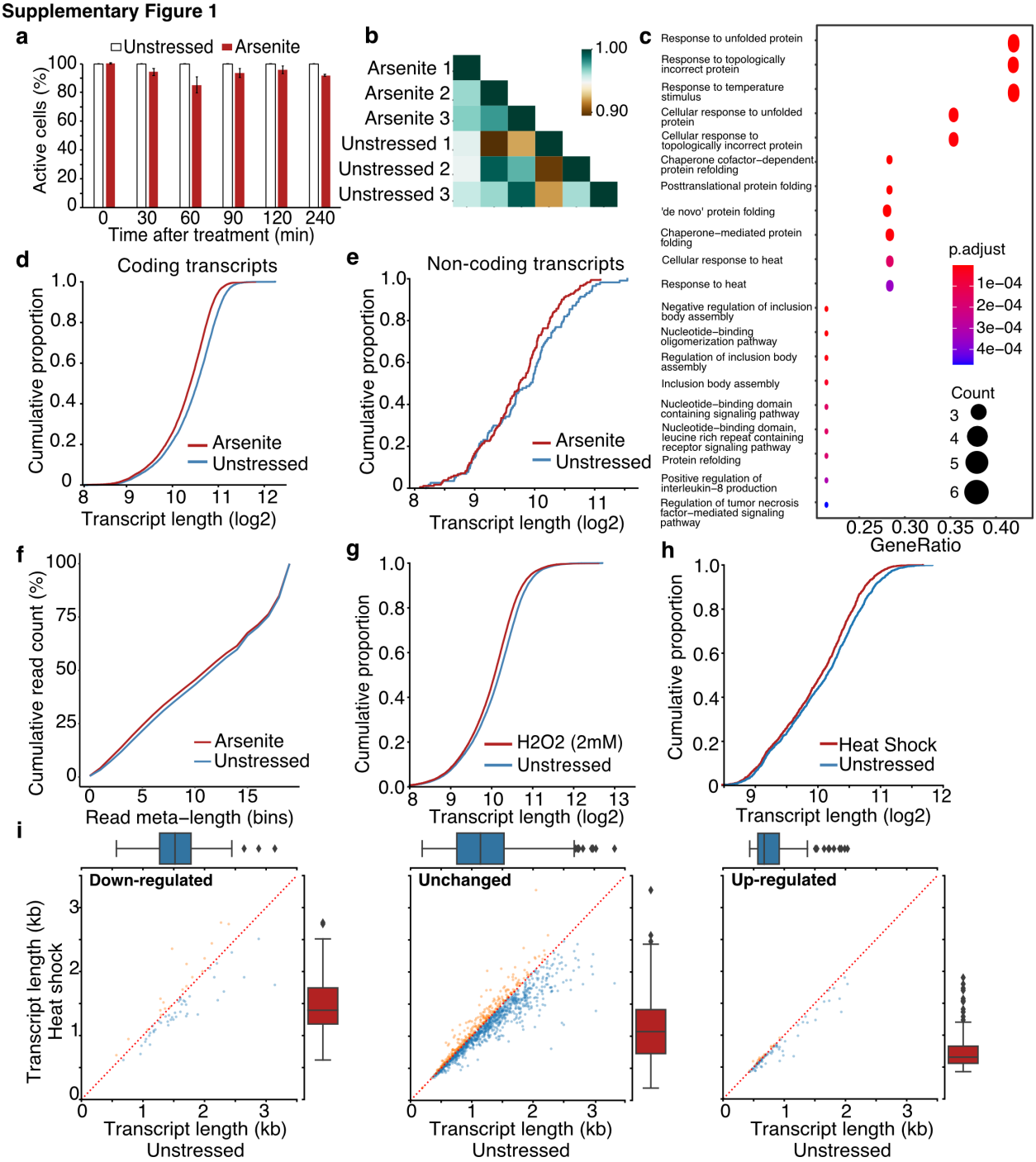


Fig. 6: Inhibition of stress-granule formation in cells devoid of G3BP1/2 rescues RNA decay **a)** Scatter plot of average gene length for arsenite-treated and unstressed cells stratified by gene SG localization. **** indicates p-value < 10^{-172}. **b-c)** Cumulative distribution and box plots of average gene length difference in arsenite-treated and unstressed cells stratified by gene SG localization. **d)** Box plot of GC content percentage for SG enriched and depleted gene transcripts. **e)** Same as (a) for average poly(A) length per gene. **f)** Cumulative distribution of poly(A) tail length for arsenite-treated and unstressed cells stratified by gene SG localization. **g)** Cumulative distribution plot of average gene length difference in arsenite-treated and unstressed cells stratified by gene P-bodies localization. **h)** Scatter plot of average transcript length for arsenite-treated and unstressed U-2 OS and ΔΔG3BP1/2 U-2 OS cells. Only transcripts with at least 5 aligned reads are used. Red dotted indicates the y=x line. **i)** Cumulative distribution plot

of average transcript length difference in arsenite-treated and unstressed U-2 OS and ΔΔG3BP1/2 U-2 OS cells. **j)** Scatter plot and box plot of differential transcript expression in arsenite-treated and unstressed ΔΔG3BP1/2 U-2 OS cells against differential shortening in arsenite-treated and unstressed U-2 OS cells, stratified by shortening significance.

Supplementary Figures

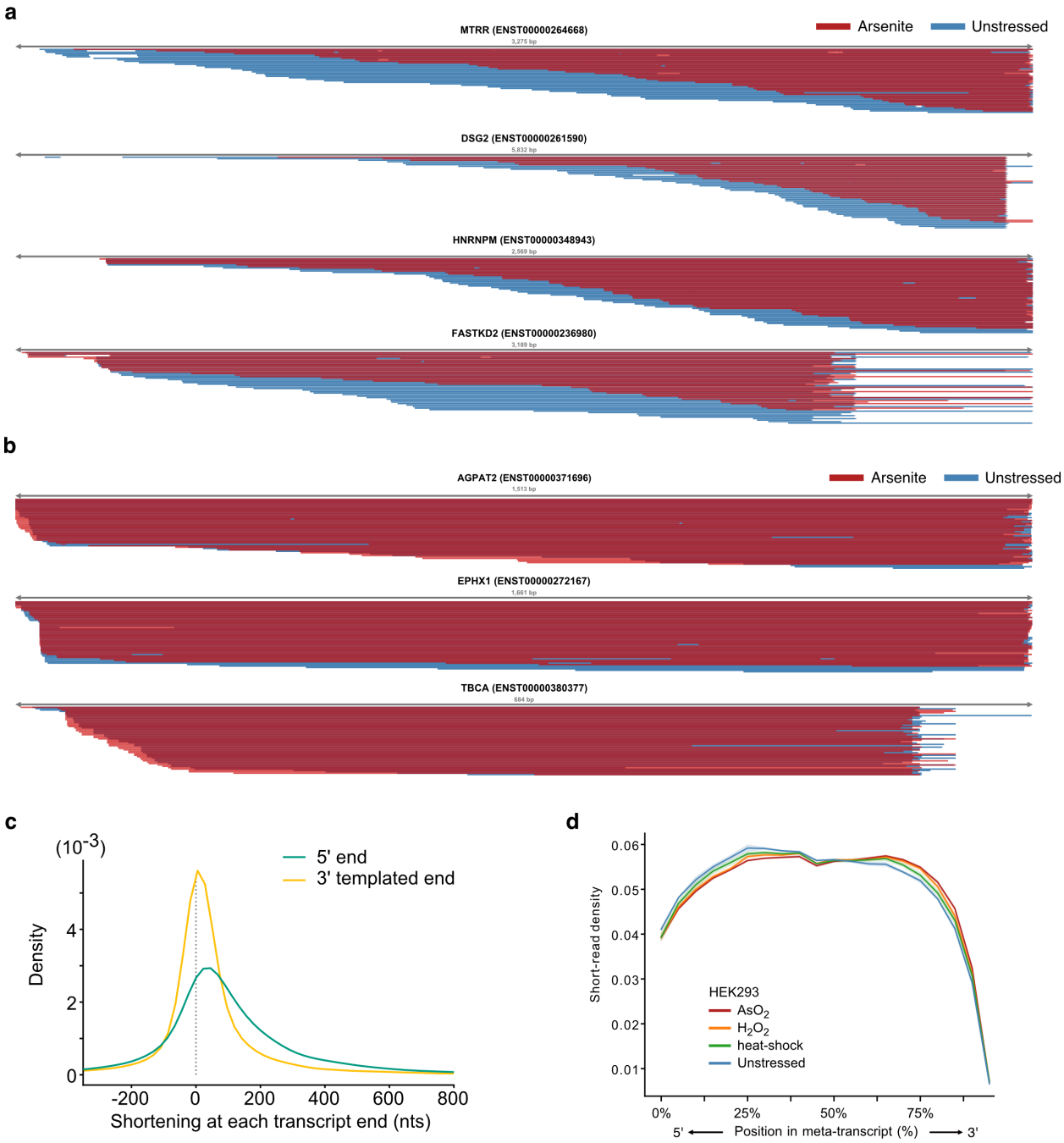
Supplementary Figure 1



Sup. Fig 1: RNA shortening upon cellular stress. a) MTS cell proliferation assay showing the percentage of active cells (y-axis) upon arsenite treatment compared to unstressed at different time points (x-axis). Error bars represent std for biological replicates (n = 2). **b)** Correlogram of transcript expression for unstressed and arsenite-treated cells. **c)** Gene ontology for biological processes for differentially expressed genes in arsenite-treated versus unstressed cells. **d-e)** Cumulative distribution plot of average transcript length for coding (d) and non-coding (e)

transcripts. Transcripts with less than 5 reads have been removed. **f)** Cumulative distribution of read meta-length for arsenite-treated and unstressed cells. **g)** Cumulative distribution plot of average transcript length for H2O2-treated cells and unstressed. **h)** Cumulative distribution plot of average transcript length for heat shock cells and unstressed. **i)** Scatter plots of average transcript length for heat shock against unstressed cells stratified by differential expression change. Down-regulated: (-Inf, -0.5), Unchanged: (-0.5, 0.5), Up-regulated (0.5, Inf) fold-change. Only transcripts with at least 5 aligned reads are shown. Red dotted indicates the y=x line. Color indicates transcripts below (blue) and above (orange) the diagonal.

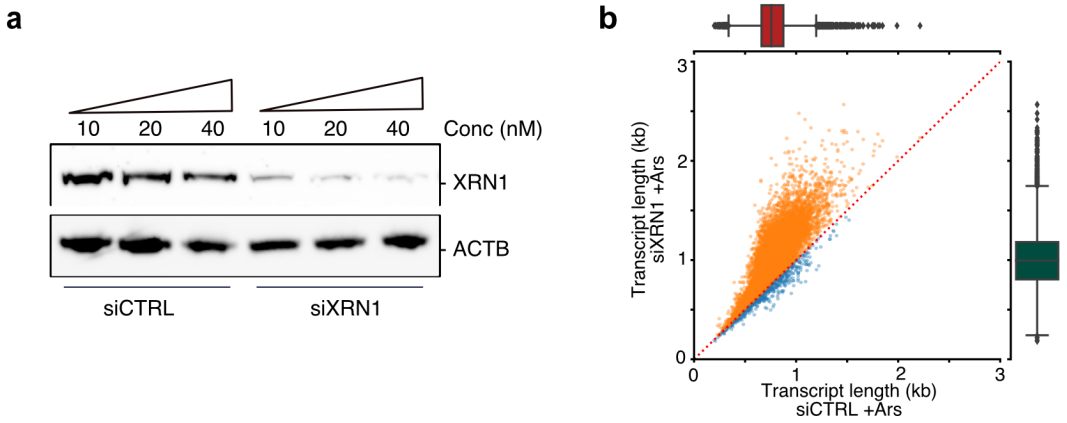
Supplementary Figure 2



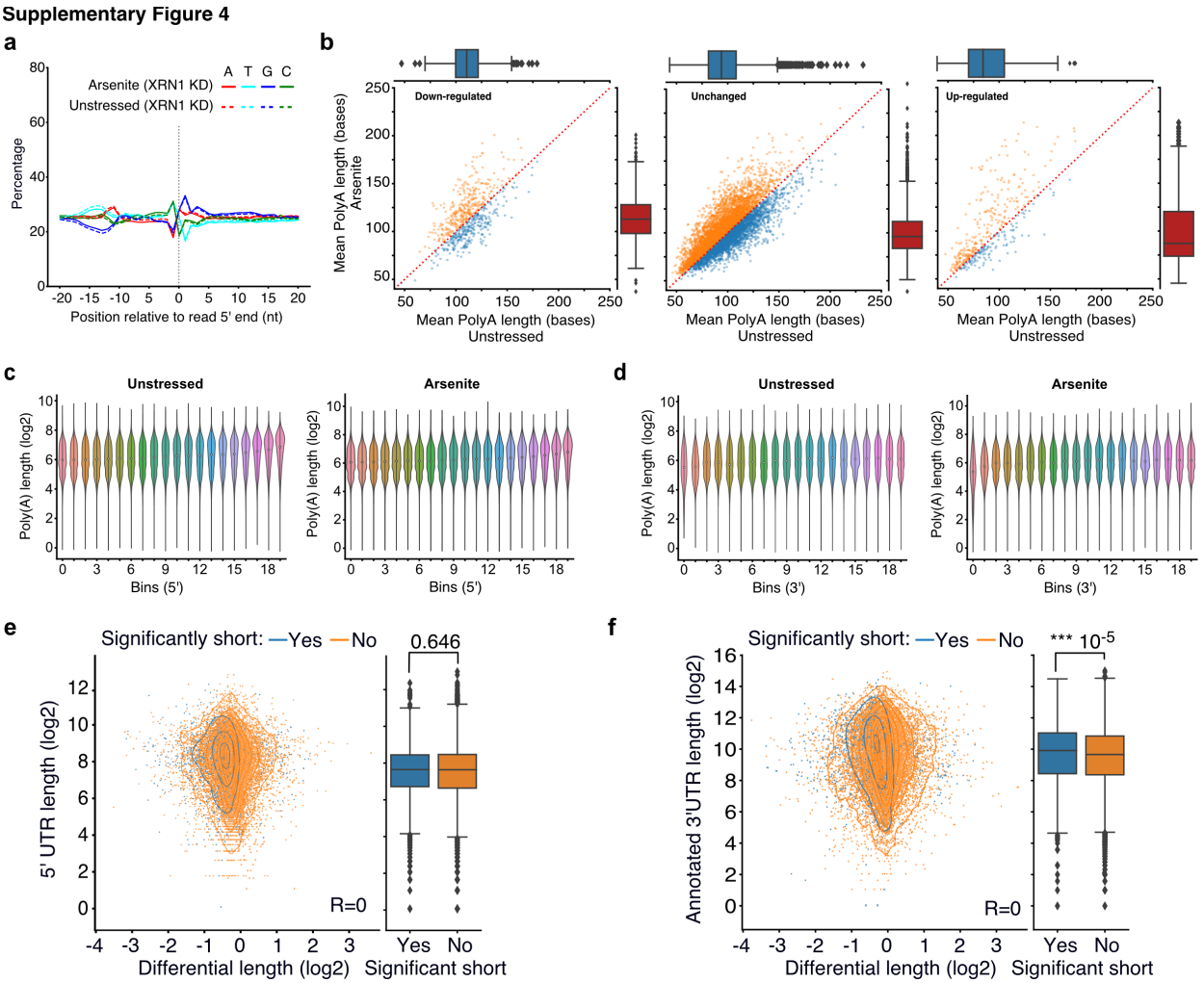
Sup. Fig. 2: Characterization of stress-induced shortened RNAs identified by NanopLen. a) IGV screenshot of aligned reads in arsenite-treated (red) and unstressed cells (blue) for significantly shortened transcripts *FASTKD2*, *HNRNPM*, *DSG2* and *MTRR*. Libraries were randomly downsampled to maximum 50 reads per window and the libraries were overlayed. All reads were used, irrespective of adaptor ligation status. **b)** Same as (a) but for non-significantly shortened transcripts *AGPAT2*, *EPHX1* and *TBCA*. **c)** Density plot of transcript shortening at the 5' and 3' templated end. Positive values on X-axis indicate that the 5' or 3' of transcripts in arsenite-treated cells are correspondingly downstream or upstream of their unstressed

counterparts. **d)** Meta-plot of short-read RNA-Seq density on all transcripts for unstressed, AsO_2, H_2O_2 and heat shock -treated cells. Shade indicates standard error of the mean for replicates.

Supplementary Figure 3

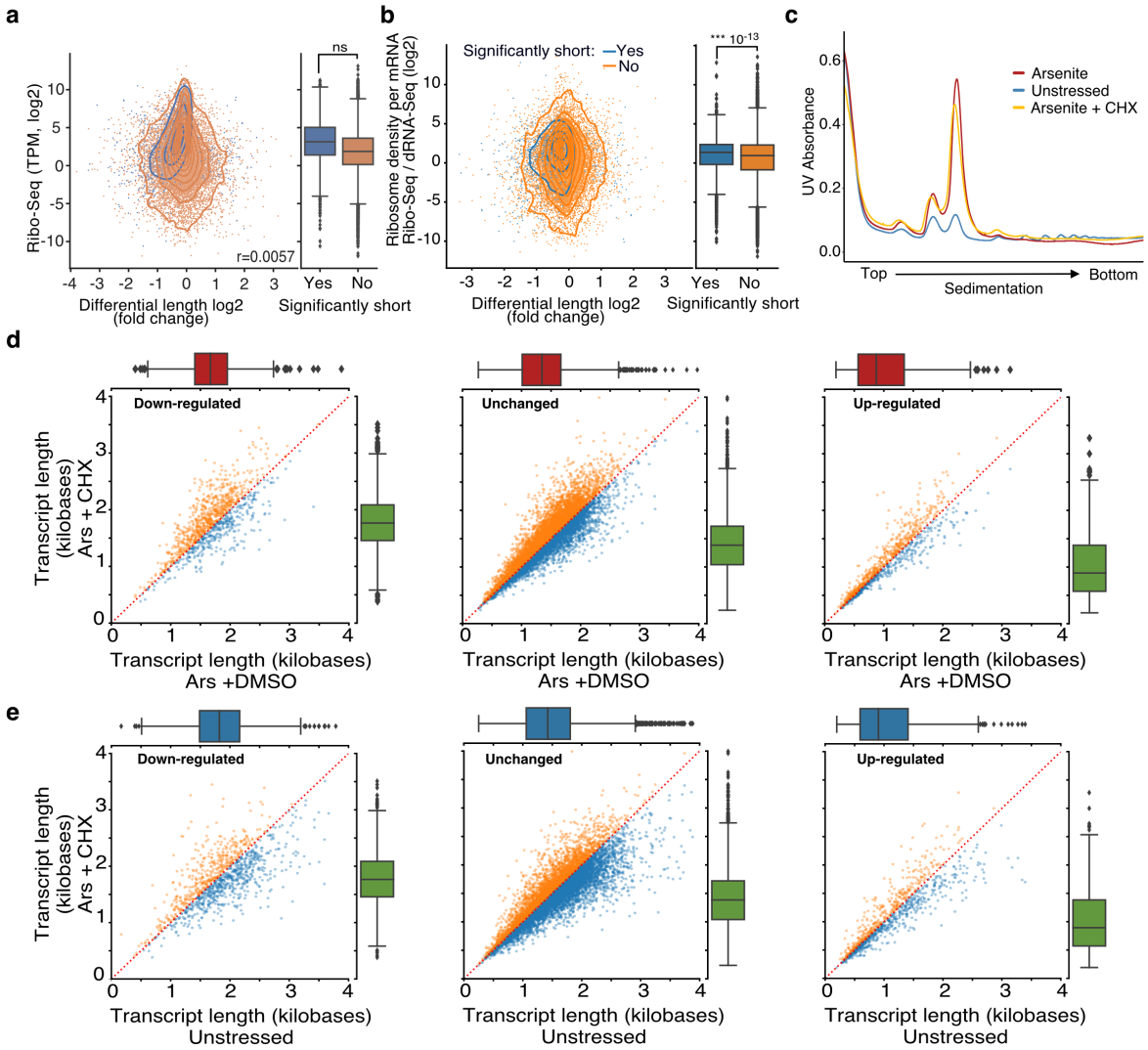


Sup. Fig. 3. The effect of XRN1 knockdown on RNA shortening. a) Immunoblot for XRN1 for cells transfected with mock and *XRN1*-targeting siRNA at varying concentrations. ACTB is used as control. **b)** Scatter plot of average transcript length for XRN1 knockdown and control in arsenite-treated cells. Only transcripts with at least 5 aligned reads are shown. Red dotted indicates the y=x line. Color indicates transcripts below (blue) and above (orange) the diagonal.

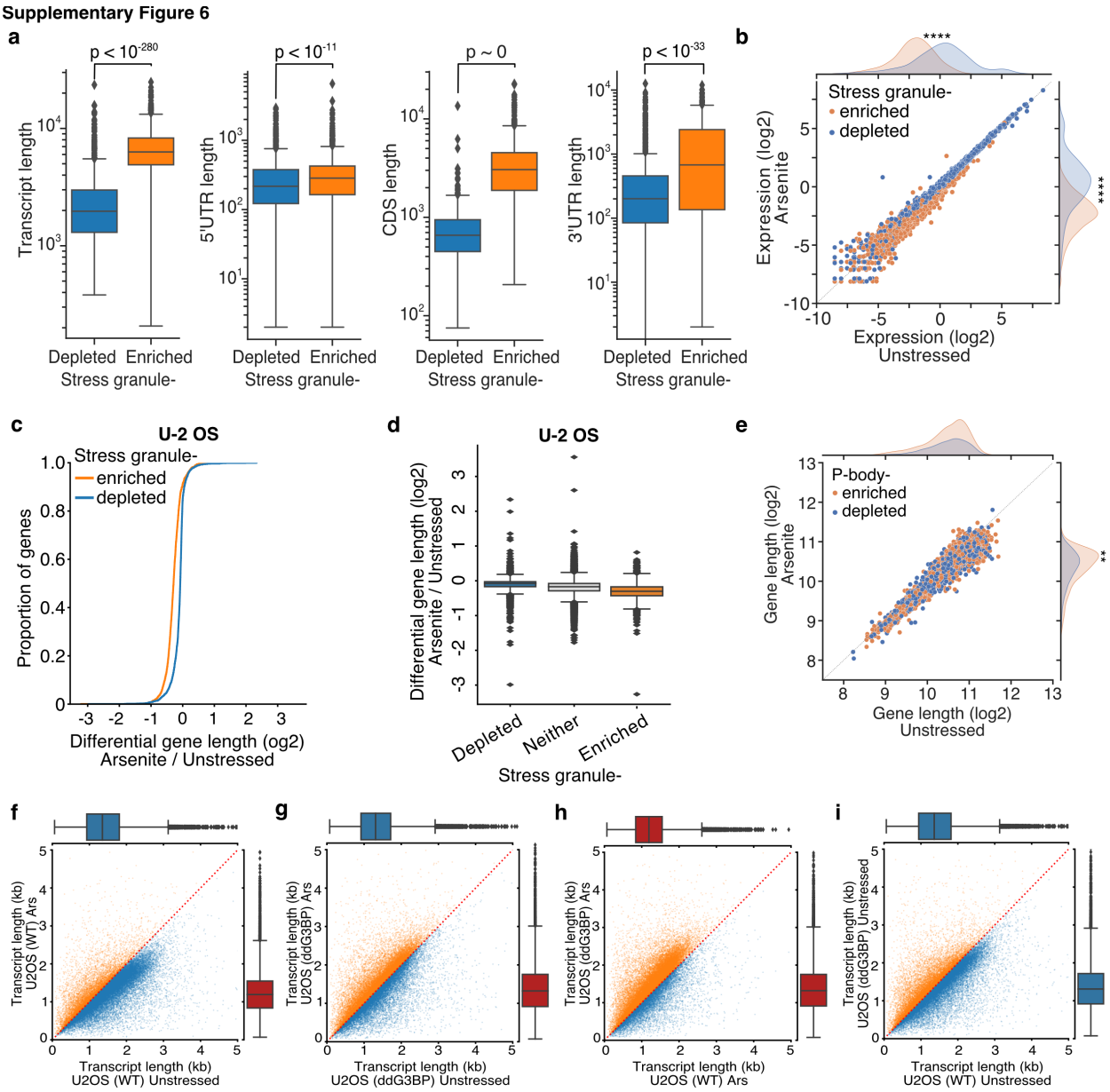


Sup. Fig. 4: Association of poly(A) tail length and cis-regulatory elements with RNA shortening. a) Nucleotide composition around the 5′ end of reads in arsenite-treated and unstressed cells upon XRN1 silencing. All reads were used, irrespective of adaptor ligation status. **b)** Scatter plots of average poly(A) tail length for arsenite-treated and unstressed cells stratified by differential expression. Down-regulated: (-Inf, -0.5), Unchanged: (-0.5, 0.5), Up-regulated (0.5, Inf) fold-change. Only transcripts with at least 5 aligned reads are shown. Red dotted indicates the y=x line. Color indicates transcripts below (blue) and above (orange) the diagonal. **c-d)** Violin plots of average transcript poly(A) length and 5′ or 3′ meta-coordinates for arsenite-treated and unstressed cells. **e-f)** Scatter plot for 5′ (e) and 3'UTR (f) length of annotated transcripts against transcript differential length in arsenite-treated and unstressed cells. The box plots on the right side summarize the y-axis variable for significant and non-significantly shortened transcripts. The Pearson's correlation coefficient and the Mann-Whitney-U test p-value are shown.

Supplementary Figure 5



Sup. Fig. 5. Translation and RNA shortening. a-b) Scatter plot of ribosome profiling footprint (a) and translational efficiency (b) levels against differential length in arsenite-treated and unstressed cells. Transcripts are stratified by significance of shortening **c)** Ribosome sedimentation curve following cell treatment with cycloheximide (CHX) (25 µg/ml) for arsenite-treated and unstressed cells. **d, e)** Scatter plots of average transcript length for arsenite plus cycloheximide treated cells against arsenite plus DMSO (d) and unstressed cells (e) stratified by their differential expression change. Down-regulated: (-Inf, -0.5), Unchanged: (-0.5, 0.5), Up-regulated (0.5, Inf) fold-change. Only transcripts with at least 5 aligned reads are shown. Red dotted indicates the y=x line. Color indicates transcripts below (blue) and above (orange) the diagonal.



Sup. Fig 6: Inhibition of stress-granule formation in cells devoid of G3BP1/2 rescues RNA decay. a) Box plots of annotated total, 5′ UTR, coding sequence and 3′ UTR length for SG enriched and depleted transcripts. **b)** Scatter plot of upper 90th quantile normalized gene expression for arsenite-treated and unstressed cells stratified by gene SG localization. **** indicates p-value < 10^{-172}. **c)** Cumulative density plot of transcript differential length for SG enriched and depleted gene transcripts in U-2 OS cells. Only transcripts of coding genes are used. **d)** Box plots of differential gene length in arsenite-treated and unstressed cells stratified by SG enrichment in U-2 OS cels. **e)** Scatter plot of average gene length for arsenite-treated and unstressed cells stratified by gene P-bodies localization. ** indicates p-value < 0.01. **f-i)** Scatter plot of average transcript length for arsenite-treated and unstressed U-2 OS and ΔΔG3BP1/2 cells. Only transcripts with at least 5 aligned reads are used. Red dotted indicates the y=x line. Color indicates transcripts below (blue) and above (orange) the diagonal.