

Directed evolution of Lantern enables rapid RNA and protein proximity labeling

Received: 13 June 2025

Accepted: 5 August 2026

Published online: 07 September 2026



Yuxin Fang^{1,2,10}, Ziqi Ren^{1,2,10}, Fu Zheng^{1,2,10}, Ruixiang Wang^{1,2,3,10}, Songrui Zhao³, Yihai Zhang^{3,4}, Wentao Wang^{1,2}, Chen Li³, Luorong Liu-Yang^{1,2}, Chang Lin^{1,2}, Jun Liu^{3,4,5} & Peng Zou^{1,2,3,5,6,7,8,9}✉

In eukaryotic cells, the precise spatial localization of RNAs and proteins is essential for proper cellular function. Genetically encoded photocatalytic proximity labeling techniques have expanded our ability to map subcellular proteomes and transcriptomes, but their temporal resolution remains limited. Here we introduce Lantern, an engineered flavoprotein optimized via directed evolution, which enables sub-minute, spatially resolved labeling of cellular biomolecules. Lantern is targetable to diverse subcellular compartments, including the endoplasmic reticulum (ER), mitochondria and stress granules (SGs), to map local transcriptomes (CAP-seq) and proteomes (CAP-MS). Using Lantern, we observed that *N*⁶-methyladenosine-rich RNAs are recruited to SGs within 10 minutes of stress induction, and ER-proximal RNAs associate with G3BP1 during early SG assembly. Additionally, Lantern was adapted for cell surface tagging (CAP-CELL), enabling spatially resolved cell typing and identifying cell–cell interactions. Collectively, this study establishes Lantern as a powerful tool that offers unprecedented temporal resolution for investigating the dynamic organization of subcellular molecular networks.

Eukaryotic cells are compartmentalized to ensure precise localizations of RNAs and proteins. Rapid redistribution of these molecules underlies processes such as signal transduction¹, viral infection² and cell cycle regulation³. For instance, during stress, cells concentrate specific RNAs and proteins into membraneless organelles termed SGs, which assemble through distinct stages⁴. Defining the temporal sequence of molecular recruitment is essential for understanding cellular stress responses and stress granule (SG) function.

Recent advances in proximity labeling have enabled *in situ* mapping of subcellular proteomes and transcriptomes. Engineered biotin ligases (for example, BioID⁵ and TurboID⁶) covalently label nearby

proteins^{5,6} but show limited efficiency for RNA labeling and lack precise temporal control. Engineered peroxidases (for example, APEX2 (refs. 7–9) and HRP¹⁰) improve temporal resolution by generating short-lived radicals that label proteins and RNAs within tens of nanometers¹¹. Photocatalytic approaches offer even finer temporal control¹². For example, μ Map uses light-activated carbene chemistry¹³, although the need for exogenous photocatalysts can introduce background labeling. Genetically encoded photocatalysis could address these limitations by combining temporal precision with reduced background.

We previously developed a genetically encoded photocatalytic proximity labeling strategy based on miniSOG, which generates

¹College of Chemistry and Molecular Engineering, Peking University, Beijing, China. ²Beijing National Laboratory for Molecular Sciences, Peking University, Beijing, China. ³Academy for Advanced Interdisciplinary Studies, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, China.

⁴School of Life Sciences, Peking University, Beijing, China. ⁵Beijing Advanced Center of RNA Biology (BEACON), Peking University, Beijing, China.

⁶Chinese Institute for Brain Research (CIBR), Beijing, China. ⁷Synthetic and Functional Biomolecules Center, Peking University, Beijing, China.

⁸Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, Peking University, Beijing, China. ⁹PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China. ¹⁰These authors contributed equally: Yuxin Fang, Ziqi Ren, Fu Zheng, Ruixiang Wang.

✉e-mail: zoupeng@pku.edu.cn

reactive oxygen species (ROS) upon blue light illumination to label nearby biomolecules^{14–16}. This approach enables transcriptomic (CAP-seq)¹⁴ and proteomic (RinID¹⁶/PDPL¹⁷) mapping with high spatial specificity across multiple subcellular compartments. However, the requirement for prolonged illumination (15 minutes at 30 mW cm⁻²) limits its ability to capture rapid processes such as SG assembly¹⁸. More recently, LITag, based on a miniSOG variant (LOV*), enables second-scale proteomic labeling but requires substantially higher light intensity (450 mW cm⁻²)¹⁹, and its transcriptome-wide RNA labeling performance remains unclear.

In the present study, we engineered a miniSOG variant, Lantern, via directed evolution to achieve sub-minute, spatially resolved labeling of RNAs and proteins in live cells. Lantern enables chromophore-assisted proximity labeling for transcriptomic (CAP-seq), proteomic (CAP-MS) and cellular spatial organization (CAP-CELL) analyses. Kinetic transcriptomic profiling reveals rapid recruitment of *N*⁶-methyladenosine (m⁶A)-enriched RNAs into SGs within minutes of stress, along with early association of endoplasmic reticulum (ER)-proximal RNAs with G3BP1, suggesting a role for the ER in SG biogenesis. Proteomic analysis further shows dynamic remodeling of the G3BP1 interactome, including time-dependent recruitment of m⁶A reader proteins. Together, our integrated transcriptomic and proteomic analyses support a model in which RNA mobilization precedes protein recruitment during SG assembly.

Results

Directed evolution of Lantern in mammalian cells

To evolve miniSOG toward higher photocatalytic activity, we sought to couple intracellular labeling efficiency to a selectable phenotype. We developed a mammalian cell-based directed evolution platform inspired by the MS2-MCP RNA tethering system²⁰. A library of miniSOG variants was fused to tMCP (a tandem dimer of MS2-binding protein), and the corresponding mRNA contained 24× MS2 stem loops in the 3′ untranslated region (UTR), enabling each transcript to recruit its encoded miniSOG variant and, thereby, link genotype to labeling activity (Fig. 1a). To minimize expression variability, variants were integrated as single copies into a defined genomic locus in human embryonic kidney 293T (HEK293T) cells using Bxb1-mediated recombination^{21,22}. After propargylamine (PA) labeling with blue light, labeled RNAs were extracted, biotinylated via click chemistry, enriched on streptavidin-coated beads and sequenced to quantify variant activity.

We first validated this system by comparing miniSOG with an inactive mutant (G39C)²³. Integration was confirmed by loss of blue fluorescent protein (BFP) signal and genomic sequencing (Supplementary Fig. 1 and Extended Data Fig. 1a,b). As expected, the G39C mutant showed reduced labeling efficiency, reflected by lower sequence recovery (Extended Data Fig. 1a). When cells expressing miniSOG and miniSOG^{G39C} were mixed at a 1:100 ratio, wild-type miniSOG was enriched 15-fold over the mutant, demonstrating robust selection performance (Extended Data Fig. 1c,d).

We next performed site saturation mutagenesis at residues S3, W80 and Q102 (Fig. 1b). The Q102 residue forms a hydrogen bond with the flavin mononucleotide (FMN), potentially hindering singlet oxygen (¹O₂) production by facilitating electron transfer from the protein to triplet FMN²⁴. The S3 residue interacts with Q102 and may indirectly affect the FMN hydrogen bond network²⁵. In addition, the electron-rich W80 residue is hypothesized to quench triplet FMN²⁶. Under selection with blue light irradiation (1–15 minutes), variants enriched glycine at S3 and alanine, arginine or asparagine at Q102 while disfavoring tryptophan at these positions (Extended Data Fig. 1e,f and Supplementary Data 1).

In the subsequent round of evolution under stringent 1-minute irradiation, we identified 12 mutants exceeding a log₂ fold change (log₂FC) cutoff of 2.1 (Fig. 1c and Supplementary Data 1). Testing these variants in the mitochondrial matrix of HEK293T cells

revealed that nine outperformed wild-type miniSOG, with the triple mutant S3G/W80N/Q102A showing the greatest enhancement (Extended Data Fig. 1g). This variant was designated Lantern and used for all subsequent experiments.

We next compared Lantern with previously reported miniSOG variants, normalizing labeling to protein expression levels (Fig. 1d and Extended Data Fig. 2a). Lantern exhibited 2.8-fold and 2.3-fold higher RNA labeling than SOPP2 and SOPP3 (ref. 26), respectively (Extended Data Fig. 2b). Consistently, Lantern also showed enhanced protein labeling than miniSOG and SOPP2/SOPP3 under 1-minute irradiation at 30 mW cm⁻² (Extended Data Fig. 2c).

Spectroscopic analysis showed that Lantern has an absorption peak at 440 nm, which is similar to SOPP3 and blue shifted by 8 nm relative to miniSOG (Extended Data Fig. 2d). Its singlet oxygen quantum yield (0.42, measured by Singlet Oxygen Sensor Green (SOSG)²⁷) exceeded that of SOPP3 (0.36) and miniSOG (0.10), and it displayed improved photostability (Fig. 1e and Extended Data Fig. 2e). Confocal imaging confirmed correct subcellular localization across multiple compartments (Fig. 1f).

Finally, under high-intensity illumination (1 second at 450 mW cm⁻²), Lantern produced substantially stronger labeling than LITag (Extended Data Fig. 2f). Cell viability remained approximately 80% at 1 hour after irradiation (Supplementary Fig. 2), indicating that cell samples should be collected immediately after photoreaction. Together, these results establish Lantern as an efficient genetically encoded photocatalyst for mammalian cells.

Spatially resolved transcriptome labeling by Lantern

To assess the spatial specificity and sensitivity of Lantern, we targeted it to the ER membrane and benchmarked its performance against APEX-seq²⁸ and miniSOG-based CAP-seq¹⁴ (Fig. 2a). The ER membrane (ERM) is enriched in mRNAs encoding secretory proteins²⁹. We generated a HEK293T cell line stably expressing ER membrane-targeted Lantern by fusing it to Sec61β (ERM-Lantern)³⁰. Cells were labeled with 10 mM PA under blue light (1–15 minutes). Confocal imaging, RNA dot blot and reverse transcription quantitative polymerase chain reaction (RT-qPCR) showed substantially stronger biotinylation with ERM-Lantern than ERM-miniSOG after 1 minute of irradiation (Fig. 2b and Extended Data Fig. 3a–e), with 1 minute of Lantern labeling yielding signals similar to 15 minutes of miniSOG labeling (Extended Data Fig. 3b). Although enrichment of secretory transcripts remained stable across labeling times, prolonged irradiation increased background enrichment by approximately four-fold (Extended Data Fig. 3c–e and Supplementary Note 1).

We next compared Lantern-mediated CAP-seq with APEX-seq using an APEX2-Sec61β cell line²⁸. Under a 1-minute labeling window, Lantern produced markedly stronger signals than APEX-seq or miniSOG (Fig. 2c and Supplementary Fig. 3a). RT-qPCR showed 119-fold and 150-fold higher recovery of secretory genes (*SSR2* and *TMX1*) relative to APEX-seq, whereas non-secretory background genes (*FAU* and *MTCO2*) showed smaller increases (30-fold and 26-fold) (Fig. 2d and Supplementary Fig. 3b,c), indicating improved recovery and spatial specificity.

Duplicated next-generation sequencing (NGS) under 1-minute irradiation identified 1,382 ER-proximal RNAs (1,372 mRNAs and 10 non-coding RNAs) by DESeq2 analysis (Fig. 2e, Supplementary Fig. 4a, Supplementary Data 2 and 3 and Supplementary Note 2). Of these, 96.4% (1,323/1,372) were associated with the secretory pathway based on Gene Ontology and Human Protein Atlas annotations. In comparison, APEX-seq identified 1,044 mRNAs with 89.7% specificity²⁸. Lantern and miniSOG¹⁴ showed similar specificity (96.4% versus 96.2%), but Lantern identified substantially more transcripts (1,372 versus 372) with shorter labeling time (1 minute versus 15 minutes) (Fig. 2f).

We next tested a 1-second labeling protocol by increasing light intensity to 450 mW cm⁻², reducing total light dose four-fold.

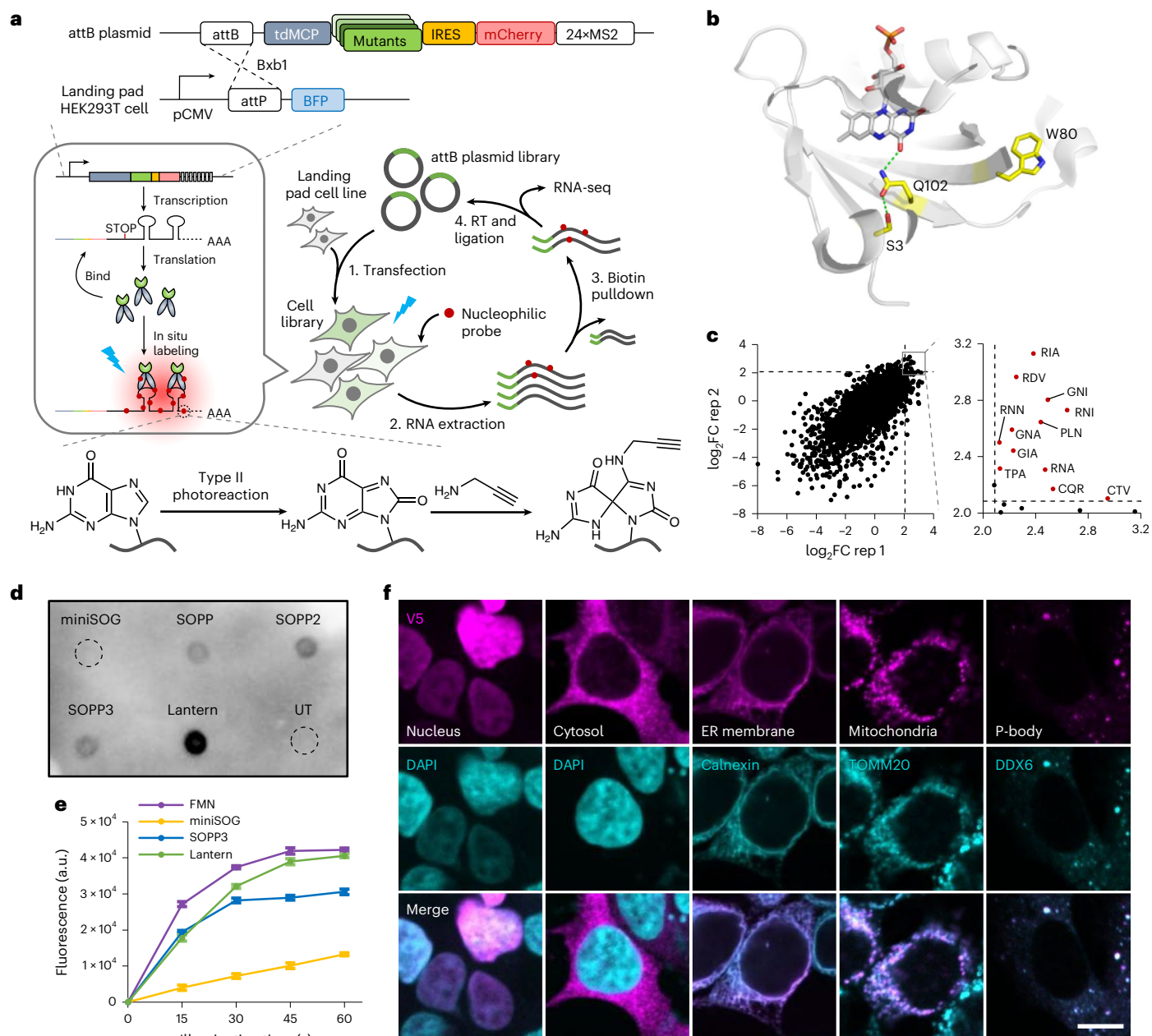


Fig. 1 | Directed evolution of Lantern. a, Schematic of the selection procedure. Each variant was targeted to its own transcript via the MCP-MS2 system and catalyzed PA-mediated RNA labeling upon blue light irradiation. The labeled RNAs were then clicked with biotin handles, enriched with streptavidin-coated beads and sequenced for activity evaluation. **b**, Structure of FMN-bound miniSOG with the mutation sites highlighted in yellow. PDB 6GPU. **c**, Enrichment fold of all mutants in the replicated second-round selection. Right: the 12 most enriched mutants. The three-letter abbreviations denote the amino acid residues at S3, W80 and Q102, respectively. **d**, Representative streptavidin dot blot

analysis of RNA labeling catalyzed by different miniSOG variants. **e**, SOSC oxidation curve upon blue light irradiation. Error bars, mean $\pm$ s.d.; $n = 3$ technical replicates. **f**, Representative immunofluorescence images of HEK293T cells expressing Lantern targeted to different subcellular localizations. Each experiment was repeated independently twice with similar results. Nucleus, H2B-Lantern; cytosol, Lantern-NES; ER membrane, Lantern-Sec61 β ; mitochondria, Mito-Lantern; P-body, Lantern-DDX6. Scale bar, 10 μ m. See Supplementary Information for plasmid details. pCMV, CMV promoter; AAA, poly(A) tail of mRNA; rep, replicate; RT, reverse transcription; UT, untransfected.

RT-qPCR confirmed efficient enrichment of *SSR2* and *TMX1* (Extended Data Fig. 3f), and NGS identified 1,878 mRNAs with 97.3% secretome specificity (Fig. 2f, Supplementary Fig. 4b and Supplementary Data 2). Despite higher overall labeling in the 1-minute condition (Extended Data Fig. 3g,h), the strong overlap between 1-second and 1-minute datasets indicates robust performance, together covering 93% (324/372) of the ERM-miniSOG dataset (Fig. 2g). Comparison with APEX-seq showed 837 shared transcripts, corresponding to 44.6% of the Lantern dataset and 80.2% of APEX-seq (Fig. 2h). By contrast, LITag profiling under identical conditions (LOV*-Sec61 β ,

1 second at 450 mW cm⁻²) yielded minimal enrichment (<0.01% recovery) and identified only eight ER-proximal transcripts (Fig. 2h and Supplementary Fig. 5).

We further compared the 1-second Lantern dataset with ER-associated transcriptomes from fluorescence-activated particle sorting (FAPS)³¹, multiplexed error-robust fluorescence in situ hybridization (MERFISH)³² and fractionation-based approaches³³. Lantern covered 55%, 69% and 77% of transcripts identified by these methods, respectively, while maintaining higher secretome specificity (Supplementary Fig. 6). Together, these results establish

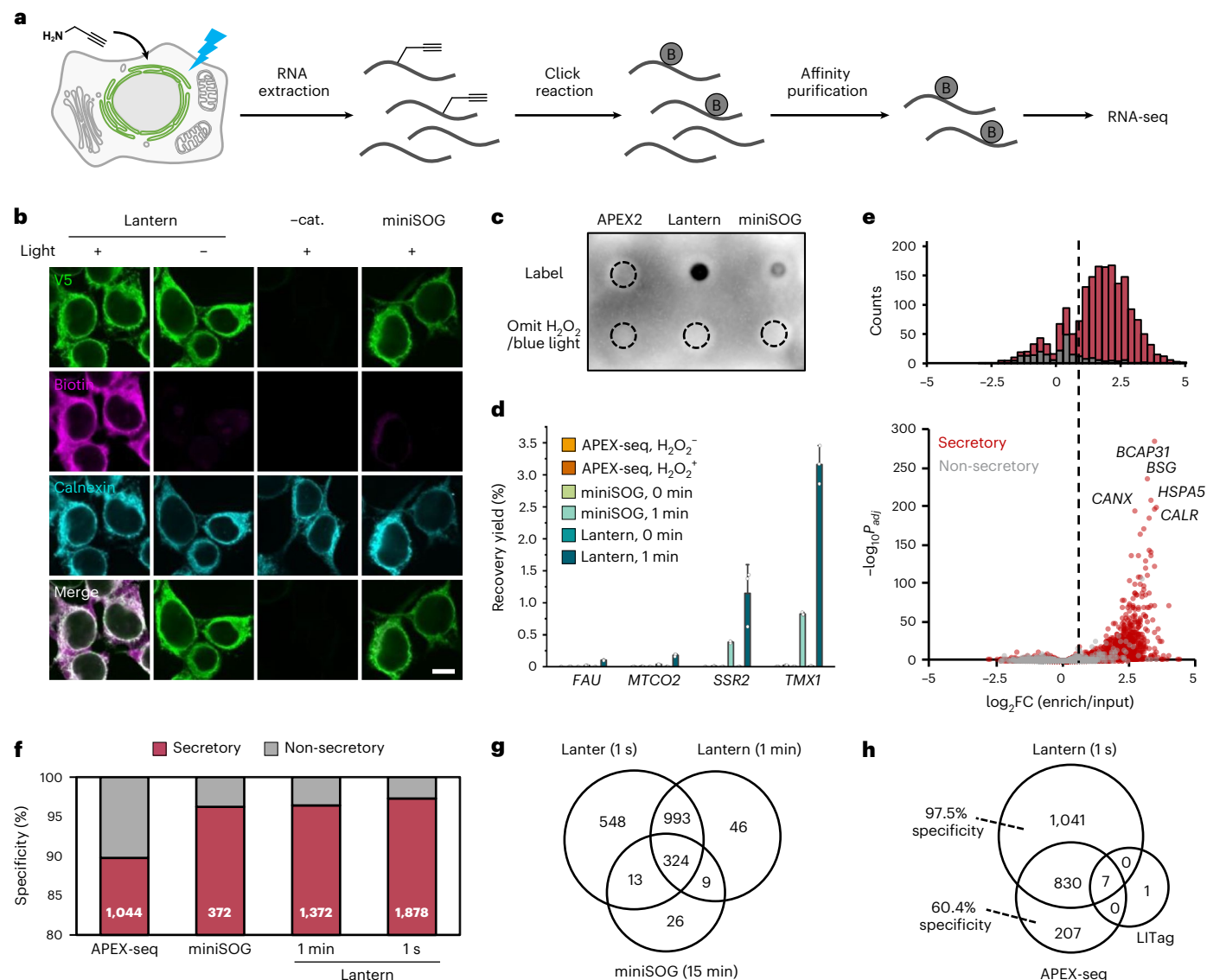


Fig. 2 | Profiling the ERM transcriptome with Lantern-mediated CAP-seq.

a, Schematic of the CAP-seq workflow. ERM-proximal RNAs were labeled with 10 mM PA during 1 minute of irradiation. After click reaction to install biotin handles, labeled RNAs were enriched with streptavidin-coated beads and sequenced. **b**, Representative immunofluorescence images of HEK293T cells labeled with ERM-Lantern. Scale bar, 10 μ m. cat., catalyst. **c**, **d**, Dot blot (**c**) and RT-qPCR (**d**) analyses comparing APEX-seq and CAP-seq labeling. Cells stably expressing APEX2-Sec61 β were incubated with 0.5 mM biotin-phenol for 30 minutes and then labeled with 1 mM H₂O₂ for 1 minute. Cells stably expressing miniSOG-Sec61 β or Lantern-Sec61 β were labeled with 10 mM PA under 1-minute irradiation at 30 mW cm⁻². *SSR2* and *TMX1* are secretory transcripts. *FAU* and *MTCO2* are non-secretory transcripts. Error bars, mean $\pm$ s.d.; $n = 3$ technical replicates. **e**, Volcano plot of the ERM-Lantern dataset. The vertical dashed line indicates the cutoff of 0.925. Analysis was performed with DESeq2 using a

two-sided Wald test, with P values adjusted by the Benjamini-Hochberg method. Top: histogram showing the distribution of secretory (red) and non-secretory mRNAs (gray). **f**, Comparison of the spatial specificity of Lantern CAP-seq with other RNA proximity labeling methods. The number of transcripts captured by each method is indicated in the columns. The 1-minute and 1-second Lantern datasets were generated with 1-minute irradiation at 30 mW cm⁻² (DESeq2 cutoffs: $\log_2FC > 0.925$, $P_{adj} < 0.05$) and 1 second of irradiation at 450 mW cm⁻² (DESeq2 cutoffs: $\log_2FC > 1.014$, $P_{adj} < 0.05$), respectively. **g**, **h**, Venn diagrams comparing Lantern CAP-seq datasets with miniSOG CAP-seq¹⁴ (**g**), APEX-seq²⁸ and LITag¹⁹ (**h**) datasets, with the spatial specificity of each method indicated on the side. The LITag dataset was generated with 1-second irradiation at 450 mW cm⁻² (DESeq2 cutoffs: $\log_2FC > 1$, $P_{adj} < 0.05$). The miniSOG, Lantern and APEX-seq datasets were the same as those used in **f**.

Lantern-mediated CAP-seq as a highly sensitive and specific approach for spatial transcriptome profiling.

Mapping the SG transcriptome during assembly

After establishing Lantern's spatial precision and sensitivity, we applied CAP-seq to map SG-associated transcriptomes during assembly. APEX-seq has been used to map dynamic changes in the SG-proximal transcriptome during heat shock³⁴. However, the kinetics of RNA recruitment to SGs under other stress conditions remain unclear. Recently, reverse transcription-based RBP binding site sequencing

(ARTR-seq) was developed to map G3BP1-interacting transcripts during arsenite-induced SG assembly but relied on a 10-minute paraformaldehyde fixation period, which is incompatible with the rapid dynamics of SG assembly³⁵. Similarly, miniSOG-based CAP-seq requires 15-minute labeling¹⁸. Leveraging Lantern's rapid labeling, we sought to capture SG transcriptome dynamics with improved temporal resolution (Fig. 3a).

We generated a HEK293T cell line stably expressing Lantern fused to the SG scaffold protein G3BP1 (G3BP1-Lantern) through lentiviral transduction. To control for RNAs interacting with diffusive G3BP1 outside of SGs, we also constructed a cell line expressing Lantern

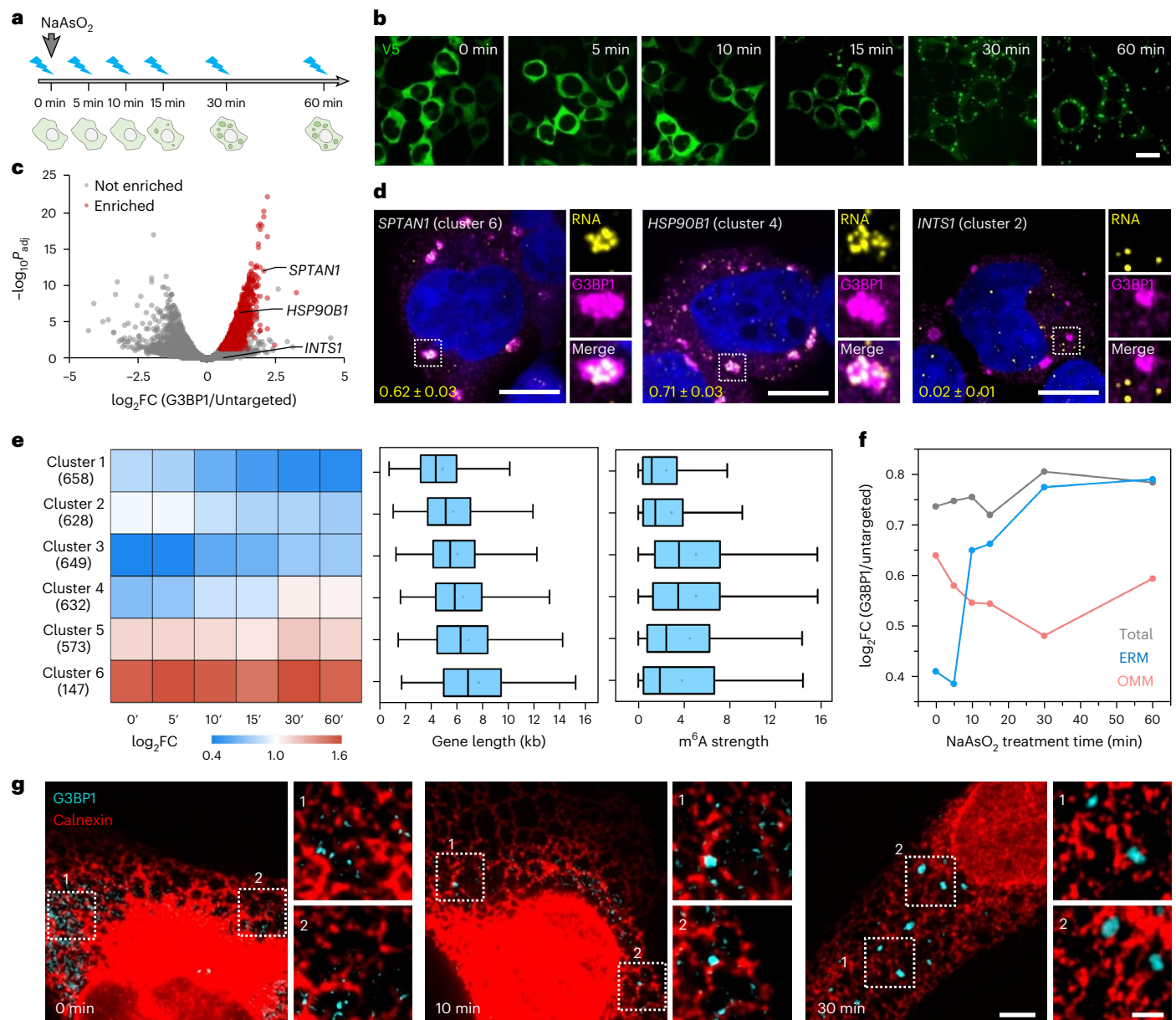


Fig. 3 | Tracking the dynamic G3BP1-interacting transcriptome during SG assembly. **a**, Schematic of temporally resolved SG transcriptome profiling. **b**, Immunofluorescence images of HEK293T cells expressing G3BP1-Lantern after treatment with 0.5 mM NaAsO₂ for 0–60 minutes. The experiment was repeated independently twice with similar results. Scale bar, 20 μ m. **c**, Volcano plot showing enrichment of SG RNAs at 60 min after stress. Red dots indicate G3BP1-proximal RNAs consistently enriched, defined as $\log_2\text{FC} > 0.3$ and $P_{\text{adj}} < 0.05$ in both Enrich versus Input and G3BP1 versus Untargeted analyses. **d**, Representative smFISH images of enriched SG transcripts *SPTAN1* and *HSP90B1* and the non-enriched transcript *INTS1*. Scale bars, 10 μ m. The co-localization ratio of each transcript is indicated at the bottom left as median $\pm$ s.e.m. **e**, RNAs identified by CAP-seq were categorized into six clusters based on their enrichment timecourses (left). Middle: box plot comparing gene lengths. Right:

box plot comparing m⁶A strength, defined as the sum of m⁶A ratios across all m⁶A sites on each transcript. Boxes indicate the first and third quartiles; whiskers indicate the outermost data points within 1.5 $\times$ the interquartile range; empty squares denote the mean; middle lines indicate the median. The n value for each box is indicated (left). Gene lengths were obtained from Ensembl⁵⁷. The m⁶A map was generated from unstressed HEK293T cells following the GLORI2.0 procedure³⁹. **f**, Average $\log_2\text{FC}$ values of ERM RNAs (blue), OMM RNAs (pink) and total RNAs (gray) during the SG assembly. **g**, Representative immunofluorescence images of U-2 OS cells treated with 0.5 mM NaAsO₂ for 0–30 minutes. G3BP1 and calnexin, an ER tubule marker, were visualized by antibody staining. Zoom-in views of the boxed regions are shown on the right. The experiment was repeated independently twice with similar results. Scale bars, 20 μ m for full images and 2 μ m for zoom-in views.

fused to a nuclear export signal (Lantern-NES). Western blot analysis confirmed that G3BP1-Lantern expression was approximately 85% of endogenous levels (Extended Data Fig. 4a). Both eIF2 α phosphorylation and SG morphology confirmed normal SG assembly dynamics (Fig. 3b and Extended Data Fig. 4b,c). To optimize labeling, cells were treated with 0.5 mM NaAsO₂ for 60 minutes and irradiated for 10–60 seconds (30 mW cm⁻²) in the presence of 10 mM PA. NaAsO₂ did not affect background labeling (Supplementary Fig. 7), and

RT-qPCR showed increasing enrichment of G3BP1-associated transcripts (for example, *DYNCH1* and *NORAD*) with longer irradiation (Extended Data Fig. 4d,e). We eventually chose 30 seconds to balance efficiency and temporal resolution.

We performed duplicated CAP-seq at six timepoints (0, 5, 10, 15, 30 and 60 minutes) after stress using Lantern-NES as a control (Extended Data Fig. 5a). G3BP1-proximal RNAs were defined by DESeq2 using both G3BP1 versus Untargeted and Enrich versus Input

comparisons ($\log_2\text{FC} > 0.3$ and adjusted $P(P_{\text{adj}}) < 0.05$). This analysis identified 1,714–2,606 transcripts across timepoints (Supplementary Data 4), with 876 consistently enriched, representing 26.6% of the 3,287 total transcripts (Extended Data Fig. 5b). Cross-correlation analysis of $\log_2\text{FC}$ (G3BP1 versus Untargeted) values showed decreasing similarity over time, indicating high temporal resolution in our datasets (Extended Data Fig. 5c). Compared to previous datasets, Lantern identified 1,425 additional transcripts beyond miniSOG-defined SG lists¹⁸ (Extended Data Fig. 5d) and achieved similar coverage to biochemical purification³⁶. By contrast, ARTR-seq³⁵ and TRIBE-ID³⁷ identified fewer transcripts (768 and 646, respectively), overlapping with Lantern CAP-seq by 51.4% and 30.2% (Extended Data Fig. 5e).

Single-molecule fluorescence in situ hybridization (smFISH) validated SG localization of Lantern-enriched transcripts, with 62–71% overlap for *SPTAN1* and *HSP90B1*, after 60 minutes of stress, whereas a non-enriched transcript (*INTS1*) showed minimal overlap (approximately 2%) (Fig. 3c,d and Extended Data Fig. 5f). Transcripts detected by biochemical purification but not Lantern (*SRRM2* and *KIF1A*) also showed low SG association (6–12%) (Extended Data Fig. 5f,g).

We next clustered the 3,287 transcripts into six groups based on temporal enrichment (Fig. 3e and Extended Data Fig. 6a,b). Clusters 1 and 2, characterized by short transcript lengths and low m⁶A strengths, represented stress-sensitive interactors that dissociated from G3BP1. By contrast, clusters 3 and 4 comprised 1,281 stress-dependent RNAs recruited upon stress, whereas clusters 5 and 6 contained 720 stress-independent interactors that remain associated with G3BP1 under both basal and stressed conditions. These stress-independent RNAs had particularly long coding sequences and may represent components of a stable interaction network centered on SG core proteins (Extended Data Fig. 6c,d). Compared to stress-sensitive interactors (that is, clusters 1 and 2), stress-dependent interactors (that is, clusters 3 and 4) showed 2.60-fold higher median m⁶A strength and 1.21-fold longer median gene length, resulting in 2.15-fold higher m⁶A density (Fig. 3e).

To test whether stress-induced changes in m⁶A affect this analysis, we also profiled m⁶A in HEK293T cells treated with 0.5 mM NaAsO₂ for 1 hour. More than 90% of m⁶A sites detected under stress overlapped with those detected under basal conditions, although global m⁶A levels were significantly reduced after stress treatment (Supplementary Fig. 8). Notably, the relative distribution of m⁶A strength across the six RNA clusters remained largely unchanged when calculated using either basal or stressed-cell m⁶A maps (Extended Data Fig. 6e). Similar patterns were also observed using previously published m⁶A datasets, including GLORI³⁸, GLORI2.0 (ref. 39) and CAM-seq⁴⁰ (Extended Data Fig. 6f and Supplementary Data 5). Gene Ontology analysis showed enrichment of cell division and RNA processing terms in stress-sensitive and stress-independent clusters, whereas stress-dependent clusters were enriched for signaling pathways (Supplementary Fig. 9). Transcripts consistently detected across all timepoints were enriched in cell division and chromatin remodeling functions.

Notably, several ER-associated transcripts (for example, *HSP90B1* and *CANX*) were highly enriched in both ERM and SG datasets. Their localizations in stressed cells were validated by smFISH (Fig. 3d and Extended Data Fig. 5f,g), prompting us to examine potential crosstalk between SGs and the ER membrane. ERM-associated RNAs showed a marked increase in SG enrichment at 10 minutes after stress (Fig. 3f) and were predominantly found in stress-dependent clusters (Extended Data Fig. 6b).

To rule out the possibility of a global increase in $\log_2\text{FC}$ (G3BP1 versus Untargeted) values, we performed the same analysis on total RNAs and outer mitochondrial membrane (OMM)-proximal RNAs. We generated an OMM-Lantern cell line by fusing Lantern to the C-terminal 31 amino acids of mitochondrial antiviral signaling protein (MAVS). CAP-seq identified 1,490 OMM-proximal transcripts, including 1,294

mRNAs, 186 pseudogenes and six long non-coding RNAs (lncRNAs) (Supplementary Data 6). This dataset overlapped with 65.5% and 19.8% of previously reported OMM transcriptomes generated using miniSOG¹⁴ and APEX-seq²⁸, respectively (Supplementary Fig. 10). Among these OMM-proximal transcripts, only 70 (4.7%) were enriched by G3BP1-Lantern at 60 minutes after stress. In contrast to ERM RNAs, the $\log_2\text{FC}$ (G3BP1 versus Untargeted) values of OMM-proximal RNAs gradually decreased during the first 30 minutes after stress (Fig. 3f). Analysis of APEX-seq atlas data further supported preferential enrichment of ER-associated transcripts during stress (Supplementary Fig. 11).

The above data implicate a role for the ER membrane in early SG assembly, consistent with previous reports of ER-condensate interactions^{33,41–43}. To further examine this, we performed immunofluorescence imaging in U-2 osteosarcoma (U-2 OS) cells to assess the spatial relationship between G3BP1 and the ER during SG assembly. At 10 minutes after stress, small G3BP1 puncta were frequently observed in close association with ER tubules, although Manders' co-localization coefficients were not significantly increased compared to unstressed cells (Fig. 3g and Supplementary Fig. 12). As SGs matured, this association decreased by 30 minutes. Together, these results suggest that G3BP1-ER interactions are enhanced during early SG assembly, with a subset of SGs forming on the ER surface.

Profiling SG proteome dynamics by CAP-MS

To complement transcriptomic analysis, we profiled SG-associated proteomes during stress with CAP-MS (Fig. 4a). Using the G3BP1-Lantern HEK293T cell line, we compared two alkynyl probes (10 mM PA and 1 mM 3-ethynylaniline (3-EA)). Streptavidin blot and silver staining showed that 3-EA provided a higher signal-to-noise ratio under 30-second irradiation at 150 mW cm⁻², whereas labeling at 30 mW cm⁻² was less efficient (Supplementary Fig. 13). The 3-EA probe also exhibited higher enrichment efficiency in RNA labeling, despite reduced overall recovery (Supplementary Fig. 14). We, therefore, used 1 mM 3-EA and 150 mW cm⁻² irradiation for subsequent proteomic experiments, with labeling validated by immunofluorescence (Fig. 4b).

For quantitative proteomics, we performed tandem mass tag (TMT)-based CAP-MS at 0, 15, 30 and 60 minutes after stress (Fig. 4c). Because Lantern-NES controls produced stronger labeling (Supplementary Fig. 15), their input was reduced 10-fold to balance enriched protein levels. Duplicate experiments were analyzed using normalized ion intensity ratios (Extended Data Fig. 7a and Methods). At 60 minutes after stress, enriched proteins included canonical SG components such as G3BP1, CAPRIN1 and UBAP2L (Fig. 4d and Extended Data Fig. 7b). Using a cutoff ratio of 1.1, we identified 87 proteins consistently enriched across replicates (Supplementary Data 7).

Comparison with the RNA Granule Database⁴⁴ showed that 85 of 87 proteins (97.7%) were annotated SG components (Tiers 1–3), including 74 Tier 1 markers (Fig. 4e). CAP-MS also showed strong agreement with APEX2 (ref. 45), BioID⁴⁶ and HaloMap⁴⁷ datasets (Fig. 4f). Although fewer proteins were identified compared to HaloMap, CAP-MS uniquely enriched Tier 1 SG components (42.9% versus 7.7%), indicating higher specificity. A similar trend was observed when comparing CAP-MS with a native immunoprecipitation dataset⁴⁸ (Extended Data Fig. 7c).

The SG proteome was dominated by RNA-binding proteins (RBPs) (84/87), including P-body proteins, helicases, m⁶A readers and ubiquitin-related factors (Supplementary Data 7). Five G3BP1 interactors overlapped with the ERM proteome previously defined by RinID¹⁶, including HDLBP, which binds ER-localized mRNAs and promotes their translation through interactions with ribosomes and the signal recognition particle⁴⁹. These findings suggest that ERM-associated RBPs may facilitate the recruitment of secretory transcripts during early SG assembly.

Temporal analysis of G3BP1/Untargeted ratios revealed that most proteins increased over time, with core SG components (for example, G3BP1, CAPRIN1, FMR1 and UBAP2L) ranking highly even before stress,

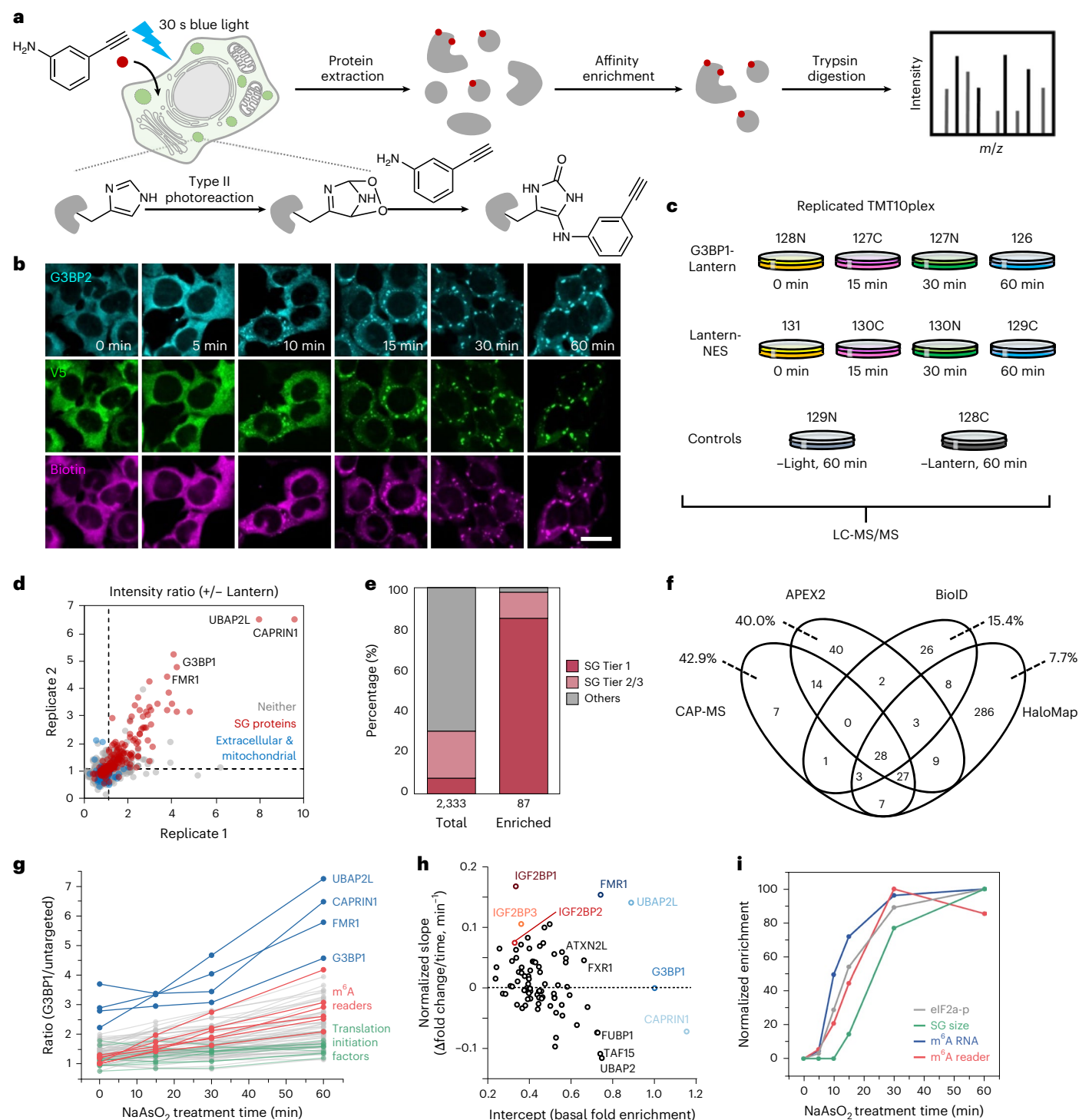


Fig. 4 | Temporally resolved mapping of the SG proteome by CAP-MS during assembly. **a**, Schematic of CAP-MS using G3BP1-Lantern in the presence of 1 mM 3-EA. **b**, Immunofluorescence images of HEK293T cells treated with 0.5 mM NaAsO₂ for 0–60 minutes before CAP-MS labeling. The experiment was repeated independently twice with similar results. Scale bar, 20 μ m. **c**, Schematic of TMT-based quantitative proteomics for profiling the G3BP1-interacting proteome at 0–60 minutes after stress. **d**, Ratio-versus-ratio plots showing enrichment of SG proteins across two replicates. Red, canonical SG proteins, defined as Tier 1 proteins in the RNA Granule Database⁴⁴. Blue, extracellular or mitochondrial matrix proteins, used as false positives. Dashed lines denote the cutoff ratio of 1.1. **e**, SG specificity of proteins enriched at 60 min after stress. SG Tier 1–3

annotations are from the RNA Granule Database⁴⁴. **f**, Venn diagram comparing SG proteome datasets identified by Lantern-mediated CAP-MS, APEX2 (ref. 45), BioID⁴⁶ and HaloMap⁴⁷, with the Tier 1 ratio of each method indicated on the side. Interactors with average spectral counts of 0 under stress were excluded. **g**, Timecourse of CAP-MS enrichment ratios (G3BP1 versus Untargeted) from 0 to 60 min after stress. **h**, Plot showing the intercept, representing basal fold enrichment at $t = 0$, and normalized slope, representing Δ FC per unit time, for the 83 SG proteins identified from 0 to 15 min after stress. **i**, Timecourses of m⁶A RNA (blue) and m⁶A-binding protein (red) recruitment to G3BP1. Changes in eIF2 α phosphorylation (gray) and SG size (green) are shown for comparison.

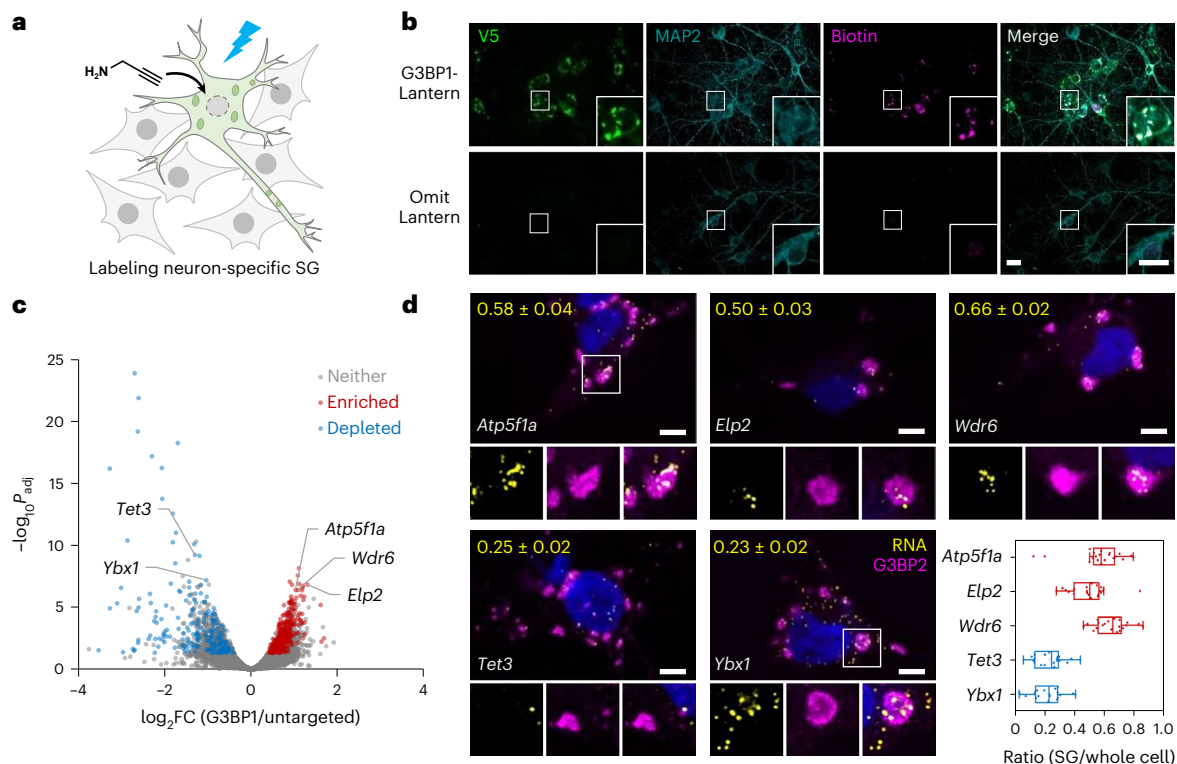


Fig. 5 | Profiling the SG transcriptome in rCNs with Lantern-mediated CAP-seq. a, Schematic of labeling neuron-specific SGs with G3BP1-Lantern. **b**, Immunofluorescence images showing rCNs labeled with Lantern at DIV8. Neurons were treated with 0.25 mM NaAsO₂ for 3 hours. The experiment was repeated independently twice with similar results. Scale bars, 20 μ m. **c**, Volcano plot showing enrichment of SG RNAs in rCNs. Analysis was performed with DESeq2 using a two-sided Wald test, with P values adjusted for multiple comparisons by the Benjamini–Hochberg method. Red and blue dots indicate

enriched and depleted RNAs, respectively. **d**, Representative smFISH images of SG-enriched transcripts (*Atp5f1a*, *Elp2* and *Wdr6*) and SG-depleted transcripts (*Tet3* and *Ybx1*). The co-localization ratio of each transcript is indicated at the top left as median $\pm$ s.e.m. Scale bars, 5 μ m. Bottom right: quantification of the ratio of smFISH signals within SGs versus the whole cell. Boxes indicate the first and third quartiles; whiskers indicate the outermost data points within 1.5 $\times$ the interquartile range; empty squares denote the mean; middle lines indicate the median. $n = 20$ cells.

consistent with pre-existing interaction networks (Fig. 4g). By contrast, some proteins showed stress-dependent recruitment. For instance, IGF2BP1, a cytoplasmic m⁶A reader, displayed minimal enrichment at baseline but increased strongly by 60 minutes (Fig. 4g).

To resolve early dynamics, we performed additional TMT analysis at 0, 5, 10 and 15 minutes after stress (Supplementary Fig. 16). Following the same analysis workflow described above, we found that most SG proteins (83/87) were already detected by 15 minutes, with a transient dip at 5 minutes followed by gradual enrichment (Extended Data Fig. 7d). To account for increased local concentration of photocatalyst due to SG assembly, enrichment ratios were normalized with respect to G3BP1. This analysis revealed increased interactions between G3BP1 and m⁶A readers beginning at approximately 10 minutes after stress (Extended Data Fig. 7e). Linear regression of normalized ratios identified IGF2BP1 as showing the strongest stress-dependent increase, with low basal interaction but a steep rise upon stress (Fig. 4h).

Motivated by the enrichment of m⁶A-binding proteins in our dataset, we examined the relationship between m⁶A modification and RNA recruitment. Transcripts stratified by m⁶A strength showed that m⁶A-rich RNAs responded most rapidly and reached the highest enrichment at 60 minutes (Extended Data Fig. 7f). Comparing temporal profiles of m⁶A-rich RNA, m⁶A reader, eIF2 α phosphorylation and SG growth revealed that m⁶A RNAs accumulated earliest, followed by m⁶A readers and eIF2 α phosphorylation, whereas SG size increased more slowly (Fig. 4i). This temporal ordering suggests that RNA recruitment precedes protein accumulation and visible SG formation. Together, these results highlight the ability of Lantern to resolve proteomic dynamics with high temporal precision.

Lantern profiles SG transcriptomes in neurons

Building on RNA labeling in HEK293T cells, we extended CAP-seq to primary cortical neurons, a more complex and sensitive system (Fig. 5a). Given the link between SGs and neurodegeneration, we used arsenite-challenged neurons to examine SG composition. Standard treatment (0.5 mM NaAsO₂, 1 hour) failed to induce SGs in all neurons, whereas prolonged exposure (3 hours) reduced viability. A milder condition (0.25 mM NaAsO₂, 3 hours) induced SG formation in most cells (Supplementary Fig. 17). Using G3BP1-miniSOG with 15-minute irradiation resulted in high background labeling in wild-type neurons, likely due to endogenous photocatalysts, as omit-light and omit-probe controls showed minimal signals (Extended Data Fig. 8a,b). These results highlight the importance of minimizing light dosage to reduce non-specific labeling.

We, therefore, replaced miniSOG with Lantern and optimized irradiation conditions. One minute of blue light produced strong biotinylation in neurons expressing G3BP1-Lantern, with minimal background in control cells lacking Lantern (Fig. 5b and Extended Data Fig. 8c,d). This condition was adopted for subsequent experiments. In addition to the wild-type (omit-Lantern) control to account for non-specific labeling, we included an Untargeted Lantern-NES control to capture labeling from diffusing G3BP1-Lantern outside SGs (Extended Data Fig. 8e). DESeq2 analysis comparing the Enrich samples to Input, Wild-type and Untargeted controls identified 1,268, 1,246 and 769 transcripts, respectively, with 321 shared across all conditions (Supplementary Data 8).

We next compared SG transcriptomes identified from rat cortical neurons (rCNs) with those from HEK293T cells. After removing non-homologous transcripts and applying a minimum read count

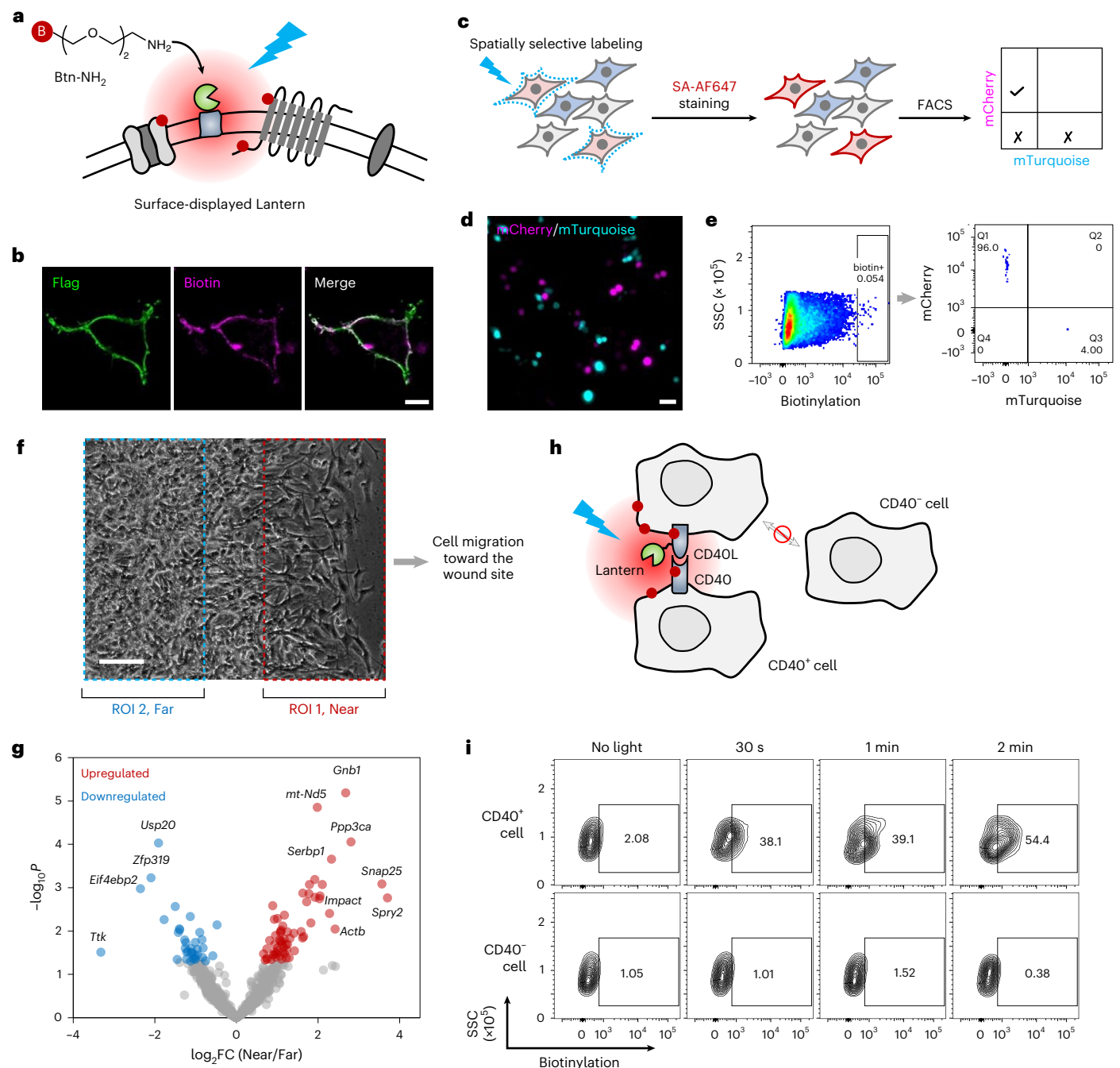


Fig. 6 | Spatially resolved cellular tagging with Lantern. **a**, Schematic illustration of cell surface labeling using Lantern-TM. **b**, Representative immunofluorescence images of HEK293T cells labeled with Lantern-TM. Scale bar, 10 μ m. **c**, Schematic of DMD-controlled cellular tagging and sorting from a mixed cell population. Cells expressing mCherry and Lantern-TM were selectively illuminated with a 445-nm laser in the presence of 2 mM Btn-NH₂. Biotinylated cells were subsequently enriched by flow cytometry and analyzed by RT-qPCR. **d**, Fluorescence images of co-cultured HEK293T cells expressing either mCherry (magenta) or mTurquoise (cyan). Scale bar, 50 μ m. **e**, Representative flow cytometry plot showing selective enrichment of mCherry-expressing cells.

f, Brightfield images of MC38 cells in a wound-healing assay after 24 hours of migration. Scale bar, 100 μ m. **g**, Volcano plot showing differentially expressed genes in migrating cells. Analysis was performed with DESeq2 using a two-sided Wald test, with raw *P* values reported. Red and blue dots indicate RNAs upregulated and downregulated, respectively, in cells closer to the wound site. **h**, Schematic illustrating CAP-CELL labeling for cell-cell interactions. **i**, Flow cytometry plots showing intercellular trans-labeling of CD40-expressing cells (top) versus control cells (bottom). The experiment was repeated independently three times. ROI, region of interest.

cutoff (10), 291 rCN and 1,615 HEK293T transcripts were retained. Among these, 113 transcripts were shared, whereas 178 (61.2%) were neuron specific (Extended Data Fig. 8f), consistent with prior reports of cell-type-specific SG proteomes (64%)⁴⁵.

For validation, three enriched transcripts (that is, *Atp5f1a*, *Elp2* and *Wdr6*; $\log_2FC > 1$) and two depleted transcripts (*Tet3* and *Ybx1*;

$\log_2FC < -1$) were analyzed by smFISH (Fig. 5c). Enriched transcripts co-localized with SGs (overlap 0.5–0.7), whereas depleted transcripts showed substantially lower co-localization (0.23–0.25) (Fig. 5d). Together, these results demonstrate that Lantern enables robust and specific profiling of SG transcriptomes in primary neurons, revealing substantial cell-type-specific RNA composition.

Spatially resolved cellular tagging with Lantern

We investigated whether Lantern-mediated protein labeling could be repurposed for cell typing by tagging cell surface proteins (CAP-CELL). To display Lantern on the cell surface, we fused it to an N-terminal mouse Ig κ -chain leader sequence and a C-terminal transmembrane domain from platelet-derived growth factor receptor (Lantern-TM). In the presence of a biotin-conjugated primary amine probe (BtN-NH₂) and blue light, proteins near Lantern-TM were biotinylated (Fig. 6a,b).

Photocatalytic labeling enables precise spatial control when paired with digital micromirror device (DMD)-assisted illumination (Fig. 6c). We co-expressed Lantern-TM with either red (mCherry) or cyan (mTurquoise2) fluorescent protein in HEK293T cells and mixed the two populations to generate a co-culture model (Fig. 6d). Fluorescence imaging was used to generate DMD masks (Extended Data Fig. 9a) for selective illumination of either population in the presence of 2 mM BtN-NH₂. After illumination, cells were dissociated, stained with streptavidin-conjugated dyes, sorted by fluorescence-activated cell sorting (FACS) and analyzed by RT-qPCR (Fig. 6c). After optimization of irradiation time (2 minutes) (Supplementary Fig. 18), two independent replicates were performed for each cell population. FACS analysis showed that more than 80% of sorted cells expressed the expected fluorescent protein (Fig. 6e and Extended Data Fig. 9b), consistent with enrichment of the corresponding transcripts by RT-qPCR (Extended Data Fig. 9c).

We next applied CAP-CELL to examine spatial transcriptomic differences during cancer cell migration. Mouse Colon 38 (MC38) cancer cells stably expressing Lantern-TM were subjected to a wound migration assay using a culture insert to create a 500- μ m cell-free gap. Optimization showed that 1 minute of 445-nm irradiation was sufficient for robust FACS detection with minimal labeling in omit-light controls (Extended Data Fig. 9d). After 24 hours of migration, DMD-guided CAP-CELL labeling was performed. Two spatial groups were defined as follows: 'Near' cells within approximately 200 μ m of the wound edge and 'Far' cells at least 300 μ m away (Fig. 6f). For each group, 50 highly biotinylated cells were isolated for RNA sequencing in five replicates. Differential expression analysis showed that 'Near' cells were enriched in migration-related and adhesion-related genes, including *Actb*, *Fbln2* and *Tubb5*, and exhibited upregulation of cancer-related signaling genes such as *Snap25*, *Spry2* and *Ppp3ca* (Fig. 6g and Supplementary Data 9). Gene Ontology analysis further highlighted translation regulation, ribosome biogenesis and cell adhesion as active pathways in migrating cells (Supplementary Fig. 19).

During optimization, prolonged illumination increased biotinylation in untransfected cells, suggesting intercellular labeling potentially caused by ROS diffusion (Supplementary Fig. 18). We reasoned that this effect could be tuned to map contact-dependent cell-cell interactions. To test this, we modeled CD40/CD40L interactions by separately transfecting HEK293T cells with CD40L-Lantern (bait) and CD40 (prey), followed by co-culture and BtN-NH₂ (100 μ M) labeling under blue light irradiation (30 mW cm⁻² for up to 15 minutes) (Fig. 6h). Untransfected cells (CD40⁻) were included as negative controls (Supplementary Fig. 20). FACS analysis revealed that 30 seconds of irradiation induced significant *cis*-labeling in CD40L-Lantern cells (Extended Data Fig. 9e). CD40-expressing prey cells also showed substantially higher *trans*-labeling than untransfected cells, reaching up to 28.6-fold enrichment after 30-second illumination (Fig. 6i and Extended Data Fig. 9f,g). Together, these results establish Lantern as a genetically encoded photocatalyst for spatially resolved cell tagging and reporting ligand-dependent cell-cell interactions.

Discussion

In this study, we engineered Lantern, a genetically encoded photocatalyst, through directed evolution. Previous rational design efforts produced miniSOG variants with improved singlet oxygen quantum yields and enhanced photocatalytic proximity labeling performance^{26,50}, but large-scale screening of miniSOG variants, particularly for RNA labeling,

has not been reported. Although phage and yeast display are widely used for directed evolution, variants selected in microbial systems often perform poorly in the mammalian cytoplasm⁵¹. We addressed this challenge by developing a mammalian cell-based selection platform that converts intracellular RNA labeling into a quantitative selection readout, thereby linking photocatalytic activity to the encoding gene sequence. In principle, this RNA self-labeling strategy can be extended to engineer other RNA-modifying enzymes.

Lantern outperforms existing proximity labeling approaches in labeling efficiency and temporal resolution. Compared to miniSOG-based implementations, Lantern achieved greater transcriptome coverage while maintaining high spatial specificity. Its rapid labeling kinetics, down to approximately 1 second, also compare favorably with enzyme-mediated proximity labeling methods, which typically require 1 minute for APEX2 or more than 10 minutes for TurboID. Small-molecule photocatalyst-based methods, including Halo-seq⁵², HaloMap⁴⁷ and μ Map-seq⁵³, provide alternative strategies but rely on exogenous targeting moieties and usually require extensive washing to remove unbound catalysts, increasing experimental complexity and background.

Using Lantern to profile SG assembly, we found that the SG transcriptome is enriched for longer RNAs with higher m⁶A modification levels, consistent with previous studies. Because CAP-seq is biased toward longer RNAs, we normalized this effect using G3BP1-to-NES fold changes. We also identified stress-independent RNA interactors with long coding sequences but low m⁶A levels that associate with G3BP1 before stress and remain enriched during SG assembly. These transcripts may reflect the intrinsic organization of G3BP1 nanodomains within mature SG cores.

We further observed a subset of ERM transcripts that associate with G3BP1 in a stress-dependent manner. Recruitment of ER-localized transcripts into cytosolic granules has been reported during unfolded protein response³³, and previous imaging studies showed ER-associated SGs under multiple stress conditions^{41,43} and SG fission at ER contact sites⁴². Consistently, our imaging revealed small G3BP1 clusters associated with ER tubules as early as 10 minutes after stress (Fig. 3g). Together, these transcriptomic and imaging results indicate that SG-ER interactions begin early in the stress response.

RNA m⁶A modification is closely linked to cytoplasmic membrane-less organelles, including SGs and P-bodies⁵⁴⁻⁵⁶. Upon stress, both m⁶A RNAs and m⁶A reader proteins showed increased enrichment by G3BP1-Lantern. Temporally resolved analysis revealed that m⁶A RNA enrichment rose more steeply than that of reader proteins during the first 5–10 minutes, suggesting earlier recruitment of m⁶A-modified RNAs during SG assembly. However, this difference may partly reflect methodological sensitivity, because CAP-seq labels guanine bases, whereas CAP-MS labels histidine residues. Differences in labeling efficiency could, therefore, contribute to the higher apparent enrichment of m⁶A RNAs relative to m⁶A reader proteins at similar local concentrations.

Despite its rapid and efficient labeling, Lantern currently requires high probe concentrations to compete with endogenous nucleophiles such as water and glutathione. Substituting PA with 3-EA reduced probe concentration by an order of magnitude, from 10 mM to 1 mM, but also lowered labeling efficiency. Future optimization could focus on probes with enhanced nucleophilicity or flavoproteins that operate through nucleophile-independent photo-induced electron transfer mechanisms.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41589-026-02313-y>.

References

- Saxton, R. A. & Sabatini, D. M. mTOR signaling in growth, metabolism, and disease. *Cell* **168**, 960–976 (2017).
- Kamel, W. et al. Alphavirus infection triggers selective cytoplasmic translocation of nuclear RBPs with moonlighting antiviral roles. *Mol. Cell* **84**, 4896–4911 (2024).
- Safieddine, A. et al. Cell-cycle-dependent mRNA localization in P-bodies. *Mol. Cell* **84**, 4191–4208 (2024).
- Wheeler, J. R., Matheny, T., Jain, S., Abrisch, R. & Parker, R. Distinct stages in stress granule assembly and disassembly. *eLife* **5**, e18413 (2016).
- Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801–810 (2012).
- Branon, T. C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887 (2018).
- Martell, J. D. et al. Engineered ascorbate peroxidase as a genetically encoded reporter for electron microscopy. *Nat. Biotechnol.* **30**, 1143–1148 (2012).
- Rhee, H.-W. et al. Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science* **339**, 1328–1331 (2013).
- Lam, S. S. et al. Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* **12**, 51–54 (2014).
- Loh, K. H. et al. Proteomic analysis of unbounded cellular compartments: synaptic clefts. *Cell* **166**, 1295–1307 (2016).
- Yang, Z. et al. Spatial barcoding reveals reaction radii and contact-dependent mechanism of proximity labeling. *Nat. Chem. Biol.* <https://doi.org/10.1038/s41589-025-02086-w> (2025).
- Fang, Y. & Zou, P. Photocatalytic proximity labeling for profiling the subcellular organization of biomolecules. *ChemBioChem* **24**, e202200745 (2023).
- Geri, J. B. et al. Microenvironment mapping via Dexter energy transfer on immune cells. *Science* **367**, 1091–1097 (2020).
- Wang, P. et al. Mapping spatial transcriptome with light-activated proximity-dependent RNA labeling. *Nat. Chem. Biol.* **15**, 1110–1119 (2019).
- Ding, T. et al. Chromophore-assisted proximity labeling of DNA reveals chromosomal organization in living cells. *Angew. Chem. Int. Ed.* **59**, 22933–22937 (2020).
- Zheng, F., Yu, C., Zhou, X. & Zou, P. Genetically encoded photocatalytic protein labeling enables spatially-resolved profiling of intracellular proteome. *Nat. Commun.* **14**, 2978 (2023).
- Zhai, Y. et al. Spatiotemporal-resolved protein networks profiling with photoactivation dependent proximity labeling. *Nat. Commun.* **13**, 4906 (2022).
- Ren, Z., Tang, W., Peng, L. & Zou, P. Profiling stress-triggered RNA condensation with photocatalytic proximity labeling. *Nat. Commun.* **14**, 7390 (2023).
- Hananya, N., Ye, X., Koren, S. & Muir, T. W. A genetically encoded photoproximity labeling approach for mapping protein territories. *Proc. Natl Acad. Sci. USA* **120**, e2219339120 (2023).
- Grünwald, D. & Singer, R. H. In vivo imaging of labelled endogenous β -actin mRNA during nucleocytoplasmic transport. *Nature* **467**, 604–607 (2010).
- Lin, C., Liu, L. & Zou, P. Functional imaging-guided cell selection for evolving genetically encoded fluorescent indicators. *Cell Rep. Methods* **3**, 100544 (2023).
- Matreyek, K. A., Stephany, J. J., Chiasson, M. A., Hasle, N. & Fowler, D. M. An improved platform for functional assessment of large protein libraries in mammalian cells. *Nucleic Acids Res.* **48**, e1 (2019).
- Swartz, T. E. et al. The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. *J. Biol. Chem.* **276**, 36493–36500 (2001).
- Torra, J. et al. Tailing miniSOG: structural bases of the complex photophysics of a flavin-binding singlet oxygen photosensitizing protein. *Sci. Rep.* **9**, 2428 (2019).
- Dietler, J. et al. Signal transduction in light-oxygen-voltage receptors lacking the active-site glutamine. *Nat. Commun.* **13**, 2618 (2022).
- Westberg, M., Bregnhøj, M., Eterodt, M. & Ogilby, P. R. No photon wasted: an efficient and selective singlet oxygen photosensitizing protein. *J. Phys. Chem. B* **121**, 9366–9371 (2017).
- Gollmer, A. et al. Singlet Oxygen Sensor Green®: photochemical behavior in solution and in a mammalian cell. *Photochem. Photobiol.* **87**, 671–679 (2011).
- Fazal, F. M. et al. Atlas of subcellular RNA localization revealed by APEX-seq. *Cell* **178**, 473–490 (2019).
- Reid, D. W. & Nicchitta, C. V. Diversity and selectivity in mRNA translation on the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.* **16**, 221–231 (2015).
- Mavylutov, T., Chen, X., Guo, L. & Yang, J. APEX2- tagging of Sigma 1-receptor indicates subcellular protein topology with cytosolic N-terminus and ER luminal C-terminus. *Protein Cell* **9**, 733–737 (2017).
- Horste, E. L. et al. Subcytoplasmic location of translation controls protein output. *Mol. Cell* **83**, 4509–4523 (2023).
- Xia, C., Fan, J., Emanuel, G., Hao, J. & Zhuang, X. Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proc. Natl Acad. Sci. USA* **116**, 19490–19499 (2019).
- Villanueva, E. et al. System-wide analysis of RNA and protein subcellular localization dynamics. *Nat. Methods* **21**, 60–71 (2023).
- Padrón, A., Iwasaki, S. & Ingolia, N. T. Proximity RNA labeling by APEX-seq reveals the organization of translation initiation complexes and repressive RNA granules. *Mol. Cell* **75**, 875–887 (2019).
- Xiao, Y. et al. Profiling of RNA-binding protein binding sites by in situ reverse transcription-based sequencing. *Nat. Methods* **21**, 247–258 (2024).
- Khong, A. et al. The stress granule transcriptome reveals principles of mRNA accumulation in stress granules. *Mol. Cell* **68**, 808–820 (2017).
- Seo, K. W. & Kleiner, R. E. Profiling dynamic RNA–protein interactions using small-molecule-induced RNA editing. *Nat. Chem. Biol.* **19**, 1361–1371 (2023).
- Liu, C. et al. Absolute quantification of single-base m⁶A methylation in the mammalian transcriptome using GLORI. *Nat. Biotechnol.* **41**, 355–366 (2023).
- Sun, H. et al. Mild and ultrafast GLORI enables absolute quantification of m⁶A methylome from low-input samples. *Nat. Methods* **22**, 1226–1236 (2025).
- Wang, P., Ye, C., Zhao, M., Jiang, B. & He, C. Small-molecule-catalysed deamination enables transcriptome-wide profiling of N⁶-methyladenosine in RNA. *Nat. Chem.* **17**, 1042–1052 (2025).
- Child, J. R., Chen, Q., Reid, D. W., Jagannathan, S. & Nicchitta, C. V. Recruitment of endoplasmic reticulum-targeted and cytosolic mRNAs into membrane-associated stress granules. *RNA* **27**, 1241–1256 (2021).
- Lee, J. E., Cathey, P. I., Wu, H., Parker, R. & Voeltz, G. K. Endoplasmic reticulum contact sites regulate the dynamics of membraneless organelles. *Science* **367**, eaay7108 (2020).
- Nguyen, D. T. M., Koppers, M. & Farias, G. G. Endoplasmic reticulum – condensate interactions in protein synthesis and secretion. *Curr. Opin. Cell Biol.* **88**, 102357 (2024).
- Youn, J.-Y. et al. Properties of stress granule and P-body proteomes. *Mol. Cell* **76**, 286–294 (2019).

45. Markmiller, S. et al. Context-dependent and disease-specific diversity in protein interactions within stress granules. *Cell* **172**, 590–604 (2018).
46. Youn, J.-Y. et al. High-density proximity mapping reveals the subcellular organization of mRNA-associated granules and bodies. *Mol. Cell* **69**, 517–532 (2018).
47. Pan, C., Knutson, S. D., Huth, S. W. & MacMillan, D. W. C. μ Map proximity labeling in living cells reveals stress granule disassembly mechanisms. *Nat. Chem. Biol.* **21**, 490–500 (2024).
48. Hein, M. Y. et al. Global organelle profiling reveals subcellular localization and remodeling at proteome scale. *Cell* **188**, 1137–1155 (2025).
49. Zinnall, U. et al. HDLBP binds ER-targeted mRNAs by multivalent interactions to promote protein synthesis of transmembrane and secreted proteins. *Nat. Commun.* **13**, 2727 (2022).
50. Westberg, M., Holmegaard, L., Pimenta, F. M., Etzerodt, M. & Ogilby, P. R. Rational design of an efficient, genetically encodable, protein-encased singlet oxygen photosensitizer. *J. Am. Chem. Soc.* **137**, 1632–1642 (2015).
51. Lambert, G. G. et al. CaBLAM: a high-contrast bioluminescent Ca^{2+} indicator derived from an engineered *Oplophorus gracilirostris* luciferase. *Nat. Methods* **23**, 205–215 (2025).
52. Engel, K. L. et al. Analysis of subcellular transcriptomes by RNA proximity labeling with Halo-seq. *Nucleic Acids Res* **50**, e24 (2022).
53. Knutson, S. D. et al. Parallel proteomic and transcriptomic microenvironment mapping (μ Map) of nuclear condensates in living cells. *J. Am. Chem. Soc.* **147**, 488–497 (2024).
54. Ries, R. J. et al. m^6A enhances the phase separation potential of mRNA. *Nature* **571**, 424–428 (2019).
55. Fu, Y. & Zhuang, X. m^6A -binding YTHDF proteins promote stress granule formation. *Nat. Chem. Biol.* **16**, 955–963 (2020).
56. Shan, T. et al. m^6A modification negatively regulates translation by switching mRNA from polysome to P-body via IGF2BP3. *Mol. Cell* **83**, 4494–4508 (2023).
57. Aken, B. L. et al. The Ensembl gene annotation system. *Database (Oxford)* **2016**, baw093 (2016).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026

Methods

Mammalian cell line culture

HEK293T, U-2 OS and MC38 cells were cultured in DMEM supplemented with 10% (v/v) FBS and maintained at 37 °C with 5% CO₂.

Stable cell lines were generated by lentiviral transduction. To produce lentivirus, HEK293T cells were cultured in six-well plates and transfected at approximately 90% confluence using Lipofectamine 3000 reagent. For each well, a 200- μ l transfection mixture was prepared by combining the lentiviral vector pLX304 (1 μ g) with the packaging plasmids dR8.91 (1 μ g) and pVSV-G (0.7 μ g), together with 6 μ l of P3000 and 6 μ l of Lipofectamine 3000. The transfection mixture was incubated at room temperature for 15 minutes before being added to the cells. During transfection, the culture medium was replaced with 1 ml of serum-free DMEM. At 6–8 hours after transfection, the transfection medium was replaced with 2 ml of complete culture medium. At 48 hours after transfection, the culture medium containing lentivirus was collected and filtered through a 0.45- μ m filter.

The landing pad cell line was constructed as previously described²². HEK293T cells were infected at approximately 0.1 multiplicity of infection (MOI) when cell confluence reached 60%. BFP was used as a marker for successful insertion of the attP site and Bxb1 recombinase²². After 48 hours, BFP⁺ cells were sorted by flow cytometry and seeded at one cell per well in 96-well plates. Monoclonal cells were cultured for at least 2 weeks and then validated by transfection with attB plasmids. The selected clone was expanded and used for mutant library construction. The cell culture medium was supplemented with 2.5 μ g ml⁻¹ blasticidin.

To construct cell lines stably expressing Lantern, 1 ml of virus suspension was mixed with 1 ml of fresh culture medium and added to HEK293T cells cultured in six-well plates at approximately 30% confluence. At 48 hours after lentiviral transduction, the culture medium was replaced with fresh complete medium. For selection, HEK293T cell lines were cultured in medium containing 5 μ g ml⁻¹ blasticidin for at least 7 days. The percentage of Lantern⁺ cells was measured by immunofluorescence.

To control G3BP1 overexpression in the G3BP1-Lantern cell line, cells with moderate green fluorescence from Lantern were sorted by flow cytometry. G3BP1 expression levels in each monoclonal cell line were assessed by western blot using an anti-G3BP1 antibody. A clone in which G3BP1-Lantern expression was 85% of endogenous G3BP1 expression was selected for transcriptomic and proteomic experiments.

Biological replicates were defined as cells from independent passages, and two biological replicates were analyzed for each transcriptomic and proteomic experiment.

Primary culture of rCNs

Neonatal Sprague Dawley rats within 24 hours after birth were used for cortical neuron dissection. Dissected cortical tissue was cut into approximately 1-mm³ blocks, washed three times with 1 $\times$ HBSS supplemented with 10 mM HEPES and digested with 0.25% trypsin-EDTA at 37 °C for 15 minutes. The digested tissue blocks were washed three times with HBSS containing 10 mM HEPES, thoroughly dispersed in 15 ml of glia plating medium (GPM; DMEM supplemented with 10% (v/v) FBS, penicillin and streptomycin) and filtered through a 70- μ m filter.

Cells were plated onto 0.12-mm glass coverslips or petri dishes and cultured in GPM for 3–6 hours before switching to neural cell culture medium (Neurobasal medium supplemented with 2% (v/v) B-27, 1% (v/v) GlutaMAX, penicillin and streptomycin). Glass coverslips were pre-incubated overnight at 37 °C with 20 μ g ml⁻¹ poly-D-lysine (PDL), followed by overnight incubation with 10 μ g ml⁻¹ mouse laminin. Petri dishes were pre-incubated overnight at 37 °C with 50–100 μ g ml⁻¹ PDL. The day of rat dissection was designated days in vitro 0 (DIV0), and half of the spent culture medium was replaced with fresh neural cell culture medium every 3 days thereafter.

Biological replicates were defined as cells derived from independent rats, and at least two biological replicates were analyzed for each transcriptomic experiment.

Directed evolution of Lantern

In designing the mutant library, we considered the modest recombination efficiency of Bxb1 (<10%) and, therefore, limited mutagenesis to no more than four target sites. With approximately 10⁷ landing pad cells as the starting population, this design ensured at least five-fold library redundancy. The mutant library was constructed by degenerate codon mutagenesis. In brief, mutations were introduced by PCR using primers containing NNK triplets at the target sites, with the miniSOG sequence as the template. Mutation-containing fragments were ligated by overlap extension PCR. Vector fragments were prepared by PCR using the attB-tdMCP-miniSOG^{G39C}-IRES-mCherry-24 $\times$ MS2 plasmid as the template. The insert fragment (363 bp, 300 ng) and vector fragment (7.16 kb, 1.6 μ g) were ligated using DNA Assembly Mix Plus (LabLead) in a 100- μ l reaction. The ligation products were purified with a DNA Clean & Concentrator Kit (Zymo Research) and eluted with 20 μ l of nuclease-free water.

The DNA ligation products were transformed into *Escherichia coli* Stbl3 cells by electroporation. To prepare electrocompetent cells, bacteria were inoculated into 100 ml of Super Optimal Broth (SOB) medium containing 20 g l⁻¹ tryptone, 5 g l⁻¹ yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂ and 10 mM MgSO₄ and incubated at 37 °C until the optical density at 600 nm (OD₆₀₀) reached approximately 1. The bacterial culture was cooled on ice for 15–30 minutes and centrifuged at 2,000g at 4 °C for 10 minutes. The supernatant was discarded. Cells were washed with 100 ml of ice-cold sterile water and pelleted by centrifugation at 2,000g at 4 °C for 10 minutes. Cells were then washed with 50 ml of ice-cold 10% (v/v) glycerol and pelleted under the same conditions. After another wash with 5 ml of ice-cold 10% glycerol, cells were pelleted again and resuspended in 250 μ l of ice-cold 10% glycerol. The cell suspension was gently mixed with 20 μ l of DNA ligation products and incubated on ice for 15 minutes. Thereafter, each 55 μ l of the cell suspension was transferred to a pre-cooled electroporation cuvette (Bio-Rad; 0.1-cm gap) and electroporated using a MicroPulser Electroporator (Bio-Rad) with one pulse in EC1 mode. Bacteria were incubated in 10 ml of SOC medium (SOB supplemented with 20 mM glucose) at 37 °C for 1 hour with shaking at 110 r.p.m. Cells were then pelleted, resuspended in 1 ml of SOB medium and plated onto LB plates supplemented with 70 μ g ml⁻¹ ampicillin. After overnight incubation at 37 °C, bacterial colonies were counted and collected for plasmid extraction.

To construct the HEK293T mutant library, landing pad cells were seeded in a 10-cm dish. When confluence reached 80%, the cells were transfected with 13.5 μ g of purified attB plasmids encoding the mutant library using Lipofectamine 3000. At 3–4 days after transfection, cells that were mCherry⁺ and BFP⁺ were selected by FACS and expanded for proximity labeling activity selection.

For the first round of selection, the cell library was seeded in 10-cm dishes. When confluence reached 90%, cells were incubated with 5 mM PA in HBSS at 37 °C for 5 minutes and irradiated with 30 mW cm⁻² blue light for 1, 5 or 15 minutes. Cells were then washed once with PBS and lysed with 4 ml of TRIzol reagent. Total RNA was extracted, clicked with N₃-PEG₃-biotin and purified with an RNA Clean & Concentrator Kit (Zymo Research). Purified RNA (180 μ g) was enriched with 18 μ l of streptavidin beads following the standard CAP-seq procedure described below. Enriched RNA was dissolved in 10 μ l of nuclease-free water. Input RNA (800 ng) and 7 μ l of enriched RNA were reverse transcribed with ProtoScript II Reverse Transcriptase using 10 pmol of gene-specific primer targeting the IRES sequence downstream of the miniSOG gene in a 20- μ l reaction. The reverse transcription product (5 μ l) was used as the template in a 50- μ l PCR reaction and amplified for 35 cycles with Phanta Max Super-Fidelity DNA Polymerase (Vazyme).

PCR products were analyzed by agarose gel electrophoresis, and the 363-bp band was purified using an E.Z.N.A. Gel Extraction Kit (Omega). For each sample, 30 ng of purified RT–PCR product was used for NGS library construction with a VAHTS Universal DNA Library Prep Kit (Vazyme) and sequenced on an Illumina NovaSeq 6000 platform in PE150 mode.

For the second round of selection, 150 ng of RT–PCR product from RNAs enriched from cells labeled with 1-minute irradiation was ligated with 800 ng of vector fragment using DNA Assembly Mix Plus (LabLead) in a 50- μ l reaction. The ligation product was purified and transformed into *E. coli* Stbl3 cells by electroporation for plasmid amplification. To construct the mutant library in HEK293T cells, landing pad cells were seeded in a 6-cm dish and transfected with 5 μ g of purified plasmid when confluence reached 80%. At 3 days after transfection, cells that were mCherry⁺ and BFP⁺ were selected by FACS and expanded for proximity labeling activity selection. The cell library was seeded in 10-cm dishes and labeled with 5 mM PA and 1-minute irradiation at 30 mW cm⁻². Experiments were conducted in duplicate. For each replicate, 180 μ g of purified RNA was enriched with 18 μ l of streptavidin beads following the standard CAP-seq RNA affinity enrichment procedure. RT–PCR analysis was conducted as described above. RT–PCR products (15 ng) were used for NGS library construction and sequenced on an Illumina NovaSeq 6000 platform in PE150 mode.

Recombinant protein expression

E. coli BL21(DE3) cells were transformed with pET21a plasmids encoding miniSOG, SOPP3 or Lantern. Bacteria were grown in 500 ml of LB medium supplemented with 70 μ g ml⁻¹ ampicillin at 37 °C. When the OD₆₀₀ reached 0.5, protein expression was induced with 0.5 mM IPTG, and bacteria were further cultured at 18 °C for 20 hours. Cells were collected by centrifugation at 4,000g for 30 minutes, resuspended in 20 ml of binding buffer (50 mM Tris-HCl, 300 mM NaCl, pH 7.5) and lysed by ultrasonication on ice. The supernatant was collected after centrifugation at 20,000g for 30 minutes at 4 °C and incubated with 1.5 ml of Ni-NTA agarose beads. The slurry was mixed on a rotator for 30 minutes at 4 °C and loaded onto a column. Excess liquid was removed by gravity flow. The Ni-NTA beads were washed with 20 ml of wash buffer (50 mM Tris-HCl, 300 mM NaCl, pH 7.5, 30 mM imidazole), and recombinant protein was eluted with 10 ml of elution buffer (50 mM Tris-HCl, 300 mM NaCl, pH 7.5, 200 mM imidazole). Purified protein was dialyzed overnight against PBS containing 1 mM dithiothreitol (DTT), concentrated by ultrafiltration, aliquoted and stored at -80 °C.

Spectroscopic measurements

Absorbance spectra of miniSOG, SOPP3 and Lantern were measured using a UV-3600i Plus UV-VIS-NIR spectrophotometer (Shimadzu). The absorbance of blank PBS was subtracted, and spectra were scaled to a range of 0–1 by minimum–maximum normalization. Fluorescence emission spectra were measured using an F-7000 fluorescence spectrophotometer (Hitachi) with excitation at 440 nm and were similarly scaled to a range of 0–1 by minimum–maximum normalization.

Singlet oxygen quantum yield measurements

SOSG was used for ¹O₂ detection. To prepare a 5 mM stock solution, 100 μ g of SOSG powder was dissolved in 25 μ l of methanol and then diluted with PBS to 3 μ M. A 30 μ M FMN solution was prepared in PBS, and its absorbance at 440 nm was measured using a NanoDrop 2000c (Thermo Fisher Scientific). Protein solutions of miniSOG, SOPP3 and Lantern were diluted in PBS to match the absorbance of 30 μ M FMN at 440 nm. Equal volumes of SOSG and photocatalyst solutions were mixed and irradiated with blue light (440 nm, 75 mW cm⁻²) for 0, 15, 30, 45 or 60 seconds. The ¹O₂ production was analyzed by measuring the fluorescence of SOSG oxidation products (excitation 504 nm, emission 535 nm). Absorbance at 440 nm was also recorded to assess photocatalyst photobleaching.

Immunofluorescence

Cells were seeded onto glass coverslips in 24-well plates at a density of approximately 70,000 cells per well. Glass coverslips were pre-treated overnight at 37 °C with Corning Matrigel matrix diluted 1:100 in DMEM and washed once with PBS before use. After 24 hours, cells were treated with 0.5 mM NaAsO₂ or subjected to the photocatalytic proximity labeling procedure. Cells were then washed once with PBS and fixed with 4% (m/v) paraformaldehyde in PBS for 15 minutes at room temperature, followed by permeabilization with 0.1% (v/v) Triton X-100 in PBST (PBS containing 0.05% (v/v) Tween 20) for 15 minutes at 4 °C.

For imaging of labeling signals, 150 μ l of click reaction mixture was added to each well, containing 50 μ M N₃-PEG₃-biotin, 1 mM CuSO₄, 2 mM THPTA and 1 mg ml⁻¹ sodium ascorbate and incubated at room temperature for 30 minutes. After the click reaction, cells were rinsed twice with PBS and blocked with 3% BSA in PBST for 1 hour at room temperature. Cells were then incubated with primary antibodies for 1 hour at room temperature. After three washes with PBST (10 minutes each), cells were incubated with secondary antibodies, streptavidin-Alexa Fluor 637 (1:1,000) and DAPI (1:1,000) for 30 minutes at room temperature.

After three additional washes with PBST (10 minutes each), immunofluorescence images were acquired using an inverted fluorescence microscope (Nikon Ti-E) equipped with a spinning disk confocal unit (Yokogawa, CSU-X1) and a scientific complementary metal-oxide semiconductor camera (Hamamatsu, ORCA-Flash 4.0 V2). The imaging system was controlled with customized software written in LabVIEW version 15.0 (National Instruments).

RNA dot blot analysis

After CAP-seq labeling and affinity purification, enriched RNA was redissolved in 10 μ l of ultrapure water. Equal volumes of purified biotinylated RNA were loaded onto an Immobilon-Ny transfer membrane and crosslinked to the membrane using an ultraviolet crosslinker. The membrane was blocked with 3% BSA in PBST at room temperature for 1 hour and incubated with 0.5 μ g ml⁻¹ streptavidin-HRP in PBST at room temperature for 1 hour. The membrane was washed three times with PBST for 10 minutes each, incubated with Clarity Max Western ECL Substrate and imaged using a ChemiDoc MP Imaging System (Bio-Rad).

RT–qPCR analysis of enriched RNA

Total RNA samples before and after enrichment were reverse transcribed with ProtoScript II Reverse Transcriptase using random primers in a 10- μ l reaction, according to the manufacturer's protocol. The resulting cDNA was aliquoted into four tubes, one for each target gene, and used as the template for qPCR. Templates were mixed with PowerUp SYBR Green Master Mix and primers and quantified using an ABI StepOne Plus system.

The total RNA mass of the pre-enrichment sample used for RT–qPCR was denoted as M_{input} (in μ g). For each biotin affinity purification reaction, M_{total} (in μ g) of total RNA was subjected to streptavidin bead enrichment, and the eluted RNA was resuspended in ddH₂O to a final volume of V_{total} (in μ l). For RT–qPCR analysis, V_{enrich} (in μ l) of the post-enrichment sample was used. The theoretical mass of enriched RNA (assuming 100% recovered from affinity purification), M_{enrich} (in μ g), was calculated as follows:

$$M_{\text{enrich}} = \frac{V_{\text{enrich}}}{V_{\text{total}}} \times M_{\text{total}}$$

C_i values were averaged from three replicate measurements. Negative controls, including omit-Lantern, omit-light or the non-target gene *MTCO2*, were used to calculate enrichment fold using the $\Delta\Delta C_i$ method:

$2^{-(\Delta C_t^{\text{label}} - \Delta C_t^{\text{control}})}$, where $\Delta C_t = C_t^{\text{enrich}} - C_t^{\text{input}}$. Recovery yield was calculated using the following formula:

$$\text{Recovery yield\%} = 2^{-\Delta C_t} \times \frac{M_{\text{input}}}{M_{\text{enrich}}} \times 100\%$$

Single-molecule FISH

Oligonucleotide sequences used for smFISH experiments are provided as Supplementary Data 10. Primary probes targeting specific mRNAs were designed using Oligostan⁵⁸, with the FLAP X complementary sequence appended to the 3' end of each probe. Probes for each target mRNA were dissolved in Tris-EDTA buffer to 100 μM , mixed in equal volumes and diluted in Tris-EDTA buffer to a final concentration of 20 μM . Secondary probes containing Alexa Fluor 647 conjugated to both the 3' and 5' termini of the FLAP X sequence were synthesized by Invitrogen and stored in Tris-EDTA buffer as 100 μM stock solutions.

Cells stably expressing G3BP1-Lantern were cultured to approximately 50% confluence on glass coverslips. Coverslips were pre-coated overnight at 37 °C with Corning Matrigel matrix diluted 1:100 in DMEM. After treatment with 0.5 mM NaAsO₂ for 60 minutes, cells were washed once with PBSM (PBS containing 1 mM MgCl₂), fixed with 3.2% (w/v) paraformaldehyde in PBSM for 10 minutes at room temperature and washed with cold PBSM containing 10 mM glycine. Cells were then washed once with PBS and permeabilized on ice for 20 minutes with PBSM containing 0.1% Triton X-100 and 2 mM vanadyl ribonucleoside complex (VRC). After washing with PBSM, cells were incubated in prehybridization-30 buffer (30% formamide in 2× SSC) at room temperature for 15 minutes.

For primary probe hybridization, cells were stained overnight at 37 °C in hybridization buffer containing 10% (v/v) dextran sulfate, 30% formamide, 2× SSC, 2 mM VRC, 10 $\mu\text{g ml}^{-1}$ salmon sperm DNA, 10 $\mu\text{g ml}^{-1}$ *E. coli* tRNA, 10 $\mu\text{g ml}^{-1}$ BSA and 110 nM primary probe mix. The next day, cells were washed twice with prehybridization-30 buffer at 37 °C for 20 minutes each. Cells were then post-fixed with 1% (w/v) paraformaldehyde in PBSM for 5 minutes at room temperature, rinsed twice with 2× SSC and incubated in prehybridization-10 buffer (10% formamide in 2× SSC) at 37 °C for 10 minutes.

For secondary probe hybridization, cells were stained at 37 °C for 3 hours in hybridization buffer containing 10% (v/v) dextran sulfate, 10% formamide, 2× SSC, 2 mM VRC, 10 $\mu\text{g ml}^{-1}$ salmon sperm DNA, 10 $\mu\text{g ml}^{-1}$ *E. coli* tRNA, 10 $\mu\text{g ml}^{-1}$ BSA and 10 nM secondary probe. Stained cells were washed twice with prehybridization-10 buffer at 37 °C for 10 minutes each and then washed once with 2× SSC. Finally, cells were counterstained with DAPI diluted 1:1,000 in 2× SSC for 10 minutes at room temperature. Coverslips were mounted on glass slides using Fluoromount-G Anti-Fade.

For smFISH combined with immunofluorescence, cells were incubated with anti-V5 antibody for HEK293T cells (1:500) or anti-G3BP2 antibody for rCNs (1:200) in PBSM containing 2 mM VRC for 1 hour at room temperature. Cells were washed three times with PBSM for 5 minutes each at room temperature with gentle shaking, followed by incubation with goat anti-rabbit Alexa Fluor 488 (1:1,000) and DAPI (1:1,000) in PBSM containing 2 mM VRC for 30 minutes. Cells were then washed three times with PBSM and mounted on glass slides using Fluoromount-G Anti-Fade.

For smFISH imaging, confocal image stacks were acquired using a ×60 oil immersion objective with ×1.5 magnification and a z-step size of 0.4 μm . The scan range was set to 6 μm for HEK293T cells and 6–10 μm for rCNs, depending on the depth of the SG signal. Quantification of smFISH images was performed using a custom MATLAB script (R2021a, version 9.10.0.1602886). In brief, cell boundaries were manually outlined from maximum intensity projections of the V5 or G3BP2 image stacks. SG⁺ regions were defined as pixels with intensities above a manually determined threshold in maximum intensity projection images of the V5 or G3BP2 channel. RNA puncta were identified from

FISH images as pixels with intensities at least two-fold higher than the average intensity. RNA puncta smaller than 12 or 15 pixels, corresponding to 0.063 μm^2 or 0.078 μm^2 , depending on the target mRNA, were excluded from further analysis. RNA co-localization with SGs was quantified as the ratio of RNA puncta within SG⁺ regions to the total number of RNA puncta in each cell. RNA enrichment in SGs was also calculated using the following formula:

$$\text{Level of co-localization} = \frac{\text{pixels (RNA inside SG)} / \text{SG area}}{\text{pixels (RNA outside SG)} / (\text{whole cell} - \text{SG area})}$$

Gel electrophoresis and western blot analysis

Samples were loaded onto a 4–20% gradient SDS-PAGE gel and analyzed by electrophoresis. For Coomassie Brilliant Blue staining, the gel was incubated overnight at room temperature with 0.25% (m/v) Coomassie Brilliant Blue G250 in 10% (v/v) acetic acid and 20% (v/v) isopropanol and then destained in 10% (v/v) acetic acid and 30% (v/v) methanol.

For western blot analysis, proteins were transferred to a PVDF membrane at 230 mA for 1 hour. The membrane was blocked with 5% BSA in TBST at room temperature for 60 minutes and incubated with streptavidin-HRP (1:5,000) in TBST at room temperature for 1 hour. For anti-V5 western blotting, the membrane was incubated with mouse anti-V5 primary antibody (1:5,000) at room temperature for 60 minutes, followed by HRP-conjugated anti-mouse secondary antibody (1:5,000) for 30 minutes. The membrane was washed three times with TBST after each incubation step, 5 minutes per wash. The membrane was incubated with Clarity Western ECL Substrate and imaged using a ChemiDoc Imaging System (Bio-Rad).

Lantern-mediated RNA labeling (CAP-seq) in live cells

For dot blot or RT-qPCR analysis, cells were seeded in six-well plates. When confluence reached 90%, cells were washed once with PBS, incubated with 1 ml of pre-warmed 10 mM PA in HBSS at 37 °C for 5 minutes and irradiated with blue light at 30 mW cm⁻². Irradiation time was set according to experimental requirements. Cells were then washed once with PBS and immediately lysed with 1 ml of TRIzol reagent.

For NGS analysis, cells were seeded in 10-cm dishes. When confluence reached 90%, cells were washed once with PBS, incubated with 4 ml of pre-warmed 10 mM PA in HBSS at 37 °C for 5 minutes and irradiated with blue light at 30 mW cm⁻² for 30 seconds (SG transcriptome profiling) or 1 minute (ERM transcriptome profiling). Cells were then washed once with PBS and immediately lysed with 4 ml of TRIzol reagent. To induce SGs, cells were treated with 0.5 mM NaAsO₂ in complete medium at 37 °C for 5–60 minutes. At 5 minutes before the end of stress treatment, the medium was replaced with HBSS containing 10 mM PA and 0.5 mM NaAsO₂. Second-scale labeling experiments were performed using cells cultured in 35-mm dishes, following the same procedure used for labeling in six-well plates, except that an intense light source of 450 mW cm⁻² was used.

Rat cortical neurons were cultured in 6-cm dishes for dot blot and RT-qPCR analysis and in 10-cm dishes for NGS. For viral infection, lentivirus was diluted in culture medium according to a predetermined MOI at DIV3. At 20 hours after infection, the medium was replaced with fresh neuronal cell culture medium, and cells were cultured until DIV8 for labeling. For labeling of unstimulated cells, neurons were incubated with 10 mM PA in Tyrode's buffer at 37 °C for 5 minutes, followed by blue light LED irradiation at 30 mW cm⁻². For CAP-seq labeling under arsenite stress, cells were treated with complete neuronal cell culture medium containing 0.25 mM NaAsO₂ at 37 °C for 1 or 3 hours. At 5 minutes before the end of stress treatment, the medium was replaced with Tyrode's buffer containing 10 mM PA and 0.25 mM NaAsO₂. All lentivirus used for neuronal infection was prepared by OBiO Technology.

Affinity purification of biotinylated RNAs

Total RNA was extracted and purified with TRIzol reagent according to the manufacturer's protocol. In brief, each 1 ml of homogenized sample was mixed with 200 μ l of chloroform and centrifuged at 12,000g at 4 °C for 15 minutes. The upper aqueous phase was collected, mixed with an equal volume of isopropanol and incubated at –20 °C for at least 20 minutes. RNA was pelleted by centrifugation at 12,000g at 4 °C for 15 minutes, washed with 1 ml of 75% ethanol and dissolved in 50 μ l of RNase-free water. Purified RNA was treated with 0.5–1 μ l of DNase I at 37 °C for 30 minutes and then incubated with click reagents. Each 50 μ l of RNA sample (approximately 50 μ g) was diluted to a 62.5- μ l click reaction containing 0.1 mM N_3 -PEG₃-biotin, 2 mM THPTA, 0.5 mM CuSO₄ and 5 mM sodium ascorbate. After incubation at room temperature for 10 minutes with vigorous shaking, RNA was purified with an RNA Clean & Concentrator Kit (Zymo Research) and eluted with 50 μ l of pre-warmed nuclease-free water.

For dot blot and RT–qPCR analysis, 10–30 μ g of purified RNA was enriched with 10 μ l of Dynabeads MyOne Streptavidin C1 beads. For each NGS library of HEK293T samples, 150 μ g of purified RNA from 7.5 million cells was enriched with 40 μ l of C1 beads. For each NGS library of neuronal samples, 50 μ g of purified RNA from one 10-cm dish (seeded with 8 million cells on DIV0) was enriched with 20 μ l of C1 beads.

C1 beads were washed three times with 200 μ l of B&W buffer (5 mM Tris, pH 7.5, 1 M NaCl, 0.5 mM EDTA, 0.1% (v/v) Tween 20), followed by one wash with 200 μ l of solution A (0.1 M NaOH, 0.05 M NaCl in nuclease-free water) and one wash with 200 μ l of solution B (0.1 M NaCl in nuclease-free water). The beads were then incubated with 200 μ l of blocking buffer (1 mg ml^{–1} BSA, 1 mg ml^{–1} yeast tRNA in B&W buffer) at 25 °C for 2 hours with gentle rotation. Pre-blocked beads were washed three times with 200 μ l of B&W buffer and resuspended in 2 $\times$ B&W buffer (10 mM Tris, pH 7.5, 2 M NaCl, 1 mM EDTA, 0.2% (v/v) Tween 20).

Equal volumes of bead suspension and RNA solution were mixed and incubated at room temperature for 45 minutes with vigorous shaking. The supernatant was discarded, and the beads were washed three times with 200 μ l of B&W buffer, twice with 200 μ l of urea buffer (4 M urea, 0.1% (w/v) SDS in PBS) and twice with 200 μ l of PBS. The beads were finally resuspended in 50 μ l of elution buffer (95% formamide, 10 mM EDTA, 1.5 mM biotin), heated at 50 °C for 5 minutes and then heated at 90 °C for 5 minutes. The eluted RNA was mixed with 1 ml of TRIzol reagent and purified using the same protocol described above. To promote precipitation of dissolved RNA, 20 μ g of glycogen was added to the isopropanol precipitation system, followed by incubation at –20 °C overnight before pelleting by centrifugation. The pellet was washed with 1 ml of 75% (v/v) ethanol and dissolved in 10 μ l of pre-warmed nuclease-free water.

cDNA library construction for NGS

RNA concentration was measured using a Qubit RNA HS Assay Kit. For ERM transcriptome profiling, 100 ng of RNA from pre-enrichment and post-enrichment samples was used for NGS library construction. For SG transcriptome profiling, rRNA was removed using a Ribo-off rRNA Depletion Kit according to the manufacturer's protocol. Purified RNA was dissolved in 5 μ l of nuclease-free water.

NGS libraries were constructed using the NEBNext Ultra II RNA Library Prep Kit for Illumina. RNA was fragmented to approximately 300 nucleotides before first-strand synthesis. The resulting DNA libraries were sequenced on an Illumina NovaSeq 6000 platform to generate 150-bp paired-end reads. For ERM and neuronal SG transcriptome profiling, total RNA libraries were sequenced to a depth of 50 million paired-end reads per library. For OMM and HEK293T SG transcriptome profiling, rRNA-depleted libraries were sequenced to a depth of 10 million paired-end reads per library.

NGS data analysis

For analysis of mutant library selection results, adapters were first removed from raw sequencing reads using Trimmomatic (version 0.36).

Reads were then mapped to the miniSOG DNA sequence using Bowtie2 (version 2.3.4.3), and reads with mapping quality (MAPQ) $\leq$ 30 were discarded. Codons mapped to the targeted mutation sites were extracted and converted to the corresponding amino acids. For each read pair, amino acids at the three mutation sites were combined and counted as one mutant genotype. Only read pairs successfully mapped to all three mutation sites were retained for counting. Genotypes with counts $\geq$ 10 in the input library were retained, and enrichment folds were calculated using the following formula:

$$\log_2 FC = \log_2 \left(\frac{\text{Counts}^{\text{Enrich}} \times \text{TotalCounts}^{\text{Input}}}{\text{TotalCounts}^{\text{Enrich}} \times \text{Counts}^{\text{Input}}} \right)$$

For analysis of subcellular transcriptome data, sequencing reads were first processed with Cutadapt (version 1.18) to remove adapters. Reads from HEK293T samples were then mapped to the human genome assembly GRCh38 (hg38) using HISAT2 (version 2.1.0). Using gene annotation v87 downloaded from Ensembl, reads mapped to each gene were counted with HTSeq (version 0.7.2) using the option '--stranded no'. Reads from rCN samples were mapped to the rat genome assembly Rnor_6.0 downloaded from Ensembl and counted using Ensembl gene annotation v104. Differential expression analysis was then performed using DESeq2 (version 1.16.1).

To define ERM-proximal RNAs, protein-coding genes with $P_{\text{adj}} < 0.05$ were used for receiver operating characteristic (ROC) analysis. True-positive RNAs were defined by combining the GOCC-secretome and HPA-secretome lists from previous CAP-seq work¹⁴. False-positive RNAs were defined as genes included in the MitoCarta version 2.0 database but not in the GOCC-secretome or HPA-secretome lists. In DESeq2 analysis, both Enrich-versus-Input and Enrich-versus-Control comparisons were performed. The Enrich-versus-Input comparison generated more significantly enriched genes ($P_{\text{adj}} < 0.05$), possibly due to poor sequencing quality of enriched RNA from unlabeled cells. Cutoffs of $\log_2 FC$ (Enrich versus Input) > 0.925 and > 1.014 were determined for the 1-minute and 1-second labeling datasets, respectively.

To define OMM-proximal RNAs, DESeq2 analysis was performed to compare (1) Enrich and Input samples from labeled Lantern-MAVS cells and (2) Enrich samples from labeled Lantern-MAVS cells and omit-light control cells. Genes enriched with $\log_2 FC > 1$ and $P_{\text{adj}} < 0.05$ in both comparisons were defined as OMM-proximal RNAs.

To define G3BP1-associated RNAs in HEK293T cells, DESeq2 analysis was performed to compare Enrich and Input samples from G3BP1-Lantern cells as well as Enrich samples from G3BP1-Lantern and Lantern-NES cells. Genes enriched with $\log_2 FC > 0.3$ and $P_{\text{adj}} < 0.05$ in both comparisons were defined as G3BP1-associated RNAs.

To define G3BP1-associated RNAs in rCNs, DESeq2 analysis was performed to compare (1) Enrich and Input samples from neurons infected with G3BP1-Lantern; (2) Enrich samples from neurons infected with G3BP1-Lantern and neurons infected with Lantern-NES; and (3) Enrich samples from neurons infected with G3BP1-Lantern and wild-type neurons. Genes enriched with $\log_2 FC > 0.3$ and $P_{\text{adj}} < 0.05$ in all three comparisons were defined as G3BP1-associated RNAs.

For RNA feature analysis, transcript lengths of the human transcriptome were downloaded from BioMart Ensembl. Based on Ensembl gene annotation v87, the longest stable transcripts were selected for analysis. The 5' UTR, coding sequence and 3' UTR lengths of rat transcripts were downloaded from BioMart Ensembl, and the longest stable transcripts were selected based on Ensembl gene annotation v104. RNA feature analysis was performed using Origin 2023b.

For analysis of spatial transcriptome data during cancer cell migration, sequencing reads were first processed with Cutadapt (version 1.18) to remove adapters. Reads were then mapped to the mouse genome assembly GRCm39 using HISAT2 (version 2.1.0). Using gene annotation v107 downloaded from Ensembl, reads mapped to each gene were counted with HTSeq (version 0.7.2) using the option '--stranded no'.

Differential expression analysis was then performed using DESeq2 (version 1.16.1). Transcripts with average counts <10 were excluded. Transcripts with $\log_2FC$ (Near/Far) > 0 and $P < 0.05$ were defined as upregulated genes, whereas transcripts with $\log_2FC$ (Near/Far) < 0 and $P < 0.05$ were defined as downregulated genes.

Lantern-mediated protein (CAP-MS) labeling in live cells

For comparison of different miniSOG variants, HEK293T cells were cultured in six-well plates. When cell confluence reached 80%, the culture medium was replaced with 1 ml of DMEM containing 1% (v/v) FBS, and cells were transfected with 2 μ g of plasmid and 3 μ l of polyethylenimine (PEI) in 200 μ l of Opti-MEM reduced-serum medium. After incubation at 37 °C for 4–6 hours, the transfection mixture was replaced with 2 ml of complete medium. At 16–20 hours after transfection, cells were washed once with 1 ml of PBS, incubated with 1 ml of HBSS containing 10 mM PA at 37 °C for 5 minutes and irradiated with blue light at 30 mW cm⁻² for 1 minute.

For comparison of LITag and Lantern-mediated CAP-MS, HEK293T cells were cultured in 35-mm dishes and transfected as described above. For LITag labeling, cells were washed once with 1 ml of PBS, incubated with 1 ml of Live Cell Imaging Solution containing 0.5 mM biotin-phenol at 37 °C for 30 minutes and irradiated with blue light at 30 mW cm⁻² for 1 minute or 450 mW cm⁻² for 1 second. For CAP-MS labeling, cells were incubated with 1 ml of complete medium containing 1 mM 3-EA at 37 °C for 1 hour, washed once with 1 ml of PBS, immersed in 1 ml of HBSS containing 1 mM 3-EA and irradiated with blue light at 30 mW cm⁻² for 1 minute or 450 mW cm⁻² for 1 second.

For optimization of labeling conditions for SG proteome identification, HEK293T cells stably expressing G3BP1-Lantern, Lantern-NES or G3BP1-miniSOG were cultured in six-well plates. When cell confluence reached approximately 90%, cells were incubated at 37 °C either with 1 ml of HBSS containing 10 mM PA for 5 minutes or with 1 ml of complete medium containing 1 mM 3-EA for 1 hour. Before irradiation, complete medium containing 3-EA was replaced with 1 ml of HBSS containing 1 mM 3-EA.

After irradiation, cells were scraped off, resuspended in 1 ml of PBS, centrifuged at 300g for 3 minutes at 4 °C and lysed with 200 μ l of RIPA lysis buffer containing 25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1% (v/v) NP-40, 1% (m/v) sodium deoxycholate, 2% (m/v) SDS and 1× Protease Inhibitor Cocktail (Roche), with sonication on ice. Each 100 μ l of cell lysate was mixed with 800 μ l of pre-chilled methanol for at least 30 minutes and incubated at –80 °C overnight.

Samples were then centrifuged at 4,000g for 10 minutes at 4 °C. Protein pellets were dissolved in 100 μ l of 0.5% (m/v) SDS. Protein concentration was measured using the Pierce BCA Protein Assay Kit, and samples were diluted to equal concentrations. Subsequently, 50 μ l of click reaction mixture was added to each 100 μ l of protein solution, yielding final concentrations of 300 μ M N₃-PEG₃-biotin, 2 mM CuSO₄, 4 mM BTAA and 1.5 mg ml⁻¹ sodium ascorbate. The click reaction was performed at room temperature for 1 hour with shaking. Protein samples were mixed with 5× protein loading buffer, boiled at 95 °C for 10 minutes and analyzed by gel electrophoresis and western blotting.

Mass spectrometry sample preparation

Cells stably expressing G3BP1-Lantern or Lantern-NES were cultured in 10-cm dishes in complete medium. At approximately 90% confluence, cells were incubated with 1 mM 3-EA in complete medium for 60 minutes at 37 °C. Oxidative stress was induced by adding 0.5 mM NaAsO₂ to the culture medium for 5, 10, 15, 30 or 60 minutes. Cells were then washed once with 1 ml of PBS, incubated in 4 ml of HBSS containing 1 mM 3-EA and 0.5 mM NaAsO₂ and irradiated with blue light at 150 mW cm⁻² for 30 seconds. Cells were collected and lysed on ice in 600 μ l of RIPA lysis buffer with sonication. Proteins in the lysate were precipitated by adding 4.8 ml of pre-chilled methanol and incubating at –80 °C overnight.

After centrifugation at 4,500g for 10 minutes at 4 °C, protein pellets were washed twice with cold methanol, dissolved in 600 μ l of 0.5% (m/v) SDS and reacted with biotin-conjugated azide in a 900- μ l click reaction containing 100 μ M N₃-PEG₃-biotin, 0.67 mM CuSO₄, 1.34 mM BTAA and 0.5 mg ml⁻¹ sodium ascorbate. The click reaction was performed at room temperature for 1 hour with vigorous shaking. Proteins were then precipitated again by adding 7.2 ml of cold methanol and incubating at –80 °C overnight. Protein pellets were washed twice with cold methanol (–80 °C) and dissolved in 800 μ l of 0.5% (m/v) SDS. Protein concentration was measured using the Pierce BCA Protein Assay Kit.

For enrichment, 4 mg of protein from G3BP1-Lantern cells and wild-type HEK293T cells, or 0.4 mg of protein from Lantern-NES cells, was used. For each enrichment, 30 μ l of streptavidin agarose beads were washed once with 1 ml of PBS and incubated with the protein solution for 3 hours at room temperature with gentle rotation. Beads were then centrifuged at 3,000g for 2 minutes, the supernatant was removed and the beads were washed with 1 ml of 0.5% SDS in PBS for 10 minutes with gentle rotation, followed by six washes with 1 ml of PBS. The beads were resuspended in 500 μ l of PBS containing 6 M urea and 10 mM DTT and incubated at 60 °C for 15 minutes. Then, 25 μ l of 400 mM iodoacetamide was added, followed by incubation at 25 °C for 30 minutes in the dark. The beads were washed twice with 1 ml of 100 mM triethylammonium bicarbonate (TEAB) buffer and resuspended in 200 μ l of 100 mM TEAB. Sequencing-grade trypsin (1 μ g) was added for protein digestion at 37 °C for 16 hours with shaking at 1,200 rpm. Released peptides in the supernatant were collected by centrifugation at 15,000g for 10 minutes.

TMT 10-plex Mass Tag Labeling Kits and Reagents were used for quantitative proteomics. Peptide samples were first dried by rotary evaporation (1,500g, 45 °C, 4–6 hours) and desalted with Pierce C18 tips (100 μ l; Thermo Fisher Scientific, 87784). Desalted samples were dried again as described above and labeled with different TMT reagents, with 0.4 mg of TMT label reagent used for each reaction. Labeled samples were combined, dried and stored at –80 °C.

For profiling SG proteome during 60-minute stress, the TMT labels were assigned as follows:

- 126: G3BP1-Lantern cells treated with NaAsO₂ for 60 minutes;
- 127N: G3BP1-Lantern cells treated with NaAsO₂ for 30 minutes;
- 127C: G3BP1-Lantern cells treated with NaAsO₂ for 15 minutes;
- 128N: G3BP1-Lantern cells not treated with NaAsO₂;
- 128C: omitting Lantern;
- 129N: omitting light;
- 129C: Lantern-NES cells treated with NaAsO₂ for 60 minutes;
- 130N: Lantern-NES cells treated with NaAsO₂ for 30 minutes;
- 130C: Lantern-NES cells treated with NaAsO₂ for 15 minutes;
- 131: Lantern-NES cells not treated with NaAsO₂.

For profiling SG proteome during 15-minute stress, the TMT labels were assigned as follows:

- 126: G3BP1-Lantern cells treated with NaAsO₂ for 15 minutes;
- 127N: G3BP1-Lantern cells treated with NaAsO₂ for 10 minutes;
- 127C: G3BP1-Lantern cells treated with NaAsO₂ for 5 minutes;
- 128N: G3BP1-Lantern cells not treated with NaAsO₂;
- 128C: omitting Lantern;
- 129N: omitting light;
- 129C: Lantern-NES cells treated with NaAsO₂ for 15 minutes;
- 130N: Lantern-NES cells treated with NaAsO₂ for 10 minutes;
- 130C: Lantern-NES cells treated with NaAsO₂ for 5 minutes;
- 131: Lantern-NES cells not treated with NaAsO₂.

The peptide sample was then fractionated using a Pierce High pH Reverse Phase Peptide Fractionation Kit. Fractions were combined as '1 + 5', '2 + 6', '3 + 7' and '4 + 8'. Combined samples were dried again by

rotary evaporation (1,500g, 45 °C, 4–6 hours) and analyzed by liquid chromatography–tandem mass spectrometry (LC–MS/MS).

LC–MS/MS

Peptides were separated using a loading column (100 $\mu\text{m} \times 2\text{ cm}$) and a C18 analytical capillary column (75 $\mu\text{m} \times 15\text{ cm}$) packed in-house with Luna 3- μm C18(2) bulk packing material (Phenomenex). Mobile phases A (water with 0.1% formic acid) and B (80% acetonitrile with 0.1% formic acid) were delivered using a Dionex UltiMate 3000 RPLC nano system (Thermo Fisher Scientific). The LC gradient was held at 2% B for the first 8 minutes, increased from 2% to 10% B from 8 to 9 minutes, from 10% to 44% B from 9 to 183 minutes and from 44% to 99% B from 183 to 188 minutes, followed by 99% B from 188 to 199 minutes.

For samples analyzed on an Orbitrap Fusion Lumos Tribrid mass spectrometer, precursors were ionized using an EASY-Spray ionization source (Thermo Fisher Scientific) at +2.0 kV relative to ground, and the inlet capillary temperature was maintained at 320 °C. Survey scans of peptide precursors were collected in the Orbitrap from 350 to 1,600 Th with an AGC target of 400,000, a maximum injection time of 50 ms, a radiofrequency lens setting of 30% and a resolution of 120,000 at 200 m/z . Mono-isotopic precursor selection was enabled for peptide isotopic distributions. Precursors with charge states of $z = 2-7$ were selected for data-dependent MS/MS scans at a resolution of 60,000 over a 3-second cycle time. Dynamic exclusion was set to 15 seconds with a ± 10 -ppm window around the precursor mono-isotope.

For higher-energy collisional dissociation (HCD) scans, automated scan range determination was enabled. Precursor ions were selected with the quadrupole using an isolation window of 1.6 Th. Product ions were collected in the Orbitrap with a first mass of 110 Th, an AGC target of 50,000, a maximum injection time of 118 ms, HCD energy of 38% and a resolution of 60,000.

Mass spectrometry data analysis

All raw data were processed using MaxQuant (version 1.6.1.0). Data were searched against the UniProt UP000005640 proteome database, containing 79,052 human proteins. Half-tryptic termini and up to one missed trypsin cleavage were allowed. Carbamidomethylation of cysteine (+57.0215 Da) was set as a fixed modification, whereas oxidation of methionine (+15.9949 Da) and N-terminal acetylation (+42.0106 Da) were set as variable modifications. TMT-based database searching was performed using the built-in 10-plex TMT workflow in MaxQuant, with all isotopic correction factors set to zero. The mass tolerance for TMT reporter ions in MS/MS spectra was set to 0.003 Da.

Each biological replicate was analyzed separately. The overall TMT labeling efficiency was evaluated by searching for TMT labels as fixed peptide modifications on the N terminus and lysine side chain in the MS data. We observed high TMT labeling efficiencies across all four datasets: 93.5% and 98.8% for the two replicates of 0–60-minute experiments; 92.6% and 98.7% for the two replicates of 0–15-minute experiments. Contaminants and proteins identified as reverse hits were removed. Proteins with fewer than two unique peptides or with TMT reporter ion intensity = 0 were also removed. For each detected protein, TMT reporter ion intensity was normalized to the median intensity of all detected proteins. Ratios of normalized ion intensities were then calculated for G3BP1-Lantern samples treated with 60-minute stress (126) relative to omit-Lantern (128C), omit-light (129N) and Lantern-NES (129C) controls. To define SG proteins, cutoff ratios for $\pm$ Lantern, $\pm$ light and G3BP1/NES were set to 1.1, and only proteins commonly enriched in both replicates were retained as SG proteins.

Selective irradiation with DMD

For DMD-controlled selective irradiation, experiments were performed on a custom-built upright fluorescence microscope equipped with a $\times 25$, 1.10 numerical aperture water immersion objective lens (Nikon,

N25x-APO-MP), scientific CMOS cameras (Hamamatsu ORCA-Flash 4.0 V2), laser lines (Coherent OBIS; 405 nm, 488 nm and 561 nm) and a DMD (Texas Instruments, DLP6910 and DLP6500). The DMD control system was provided by Fldiscovery (F9120 DDR 0.65 1080P). Masks loaded onto the DMD were generated using ImageJ/Fiji (version 2.1.0) and MATLAB (version 23.1).

DMD-guided CAP-CELL labeling in co-cultures

HEK293T cells were separately transfected with Lantern-TM co-expressing either mTurquoise or mCherry. After transfection, mTurquoise-expressing and mCherry-expressing cells were mixed at a 1:1 ratio, seeded onto pre-coated 14-mm coverslips and cultured overnight in complete medium at 37 °C. Cells were then immersed in Live Cell Imaging Solution (Invitrogen) containing 2 mM Btn-NH₂. Before labeling with a DMD-controlled 445-nm laser (5 mW cm⁻²), mTurquoise or mCherry expression patterns were imaged using an upright fluorescence microscope. DMD masks were generated using manually defined thresholds. Each field of view was labeled with 445-nm light for 2 minutes, and 10 fields of view were labeled for each sample.

Cells were then washed once with PBS, dissociated with 0.05% trypsin-EDTA and collected by centrifugation at 700g for 2 minutes. Cells were stained with streptavidin-Alexa Fluor 647 at 1:500 dilution and Calcein-AM at 1:400 dilution for 30 minutes at room temperature. After two washes with PBS, cells were analyzed using a FACSaria III (BD Biosciences). Cells with strong Calcein-AM and Alexa Fluor 647 fluorescence were sorted into PCR tubes at 50 cells per tube. Each tube contained 2 μl of lysis buffer composed of 0.5% (v/v) Triton X-100, 1 U of SUPERaseIn RNase Inhibitor and 0.1 U of DNase I.

After sorting, PCR tubes were incubated at 20 °C for 10 minutes for DNA digestion. Oligo dT (0.1 μl , 100 μM), dNTPs (1 μl , 10 mM each) and EDTA (0.08 μl , 25 mM) were then added, and the mixture was incubated at 65 °C for 10 minutes, followed by 72 °C for 3 minutes to inactivate DNase and denature RNA secondary structures. SuperScript IV reverse transcriptase (0.35 μl ; Invitrogen, 200 U μl^{-1}), SUPERaseIn RNase Inhibitor (0.25 μl ; Invitrogen, 40 U μl^{-1}), 5 $\times$ SSIV buffer (2 μl), 0.1 M DTT (0.5 μl), 5 M betaine (2 μl) and ultrapure water (0.6 μl) were then added. Reverse transcription was performed at 50 °C for 50 minutes, followed by incubation at 80 °C for 10 minutes to inactivate the reverse transcriptase.

After reverse transcription, 2 μl of reaction solution was pre-amplified for 20 cycles using Phanta Max Super-Fidelity DNA polymerase (Vazyme) with gene-specific primers targeting *mTurquoise*, *mCherry* and human *GAPDH*. The PCR product (1 μl) was diluted 16-fold, mixed with PowerUp SYBR Green Master Mix and primers and quantified using an ABI StepOne Plus system. C_t values were averaged from four replicate measurements. The enrichment fold of the *mTurquoise* or *mCherry* gene was calculated as $2^{-\Delta C_t}$, where $\Delta C_t = C_t^{mTurquoise/mCherry} - C_t^{GAPDH}$.

CAP-CELL labeling of migrating cancer cells

MC38 cells stably expressing Lantern-TM were seeded onto 14-mm coverslips pre-coated overnight at 37 °C with Corning Matrigel matrix diluted 1:100 in DMEM. A culture insert (ibidi) was placed at the center of each coverslip to create a 500- μm cell-free gap. After overnight culture at 37 °C, the insert was removed, and cells were cultured for another 24 hours. For DMD-controlled CAP-CELL labeling, cells were immersed in Live Cell Imaging Solution (Invitrogen) containing 2 mM Btn-NH₂. An approximately 200- μm band near the wound border ('Near') and a band at least 300 μm away from the border ('Far') were separately labeled by 445-nm laser irradiation for 1 minute at 85 mW cm⁻². Thereafter, cells were dissociated, stained with streptavidin-Alexa Fluor 647, sorted by FACS, lysed and reverse transcribed as described above.

NGS libraries were constructed using the SHERRY2 protocol⁵⁹. In brief, 10 µl of tagmentation mix, containing 0.05 µl of TTE mix V50, 4 µl of 5× TD buffer, 0.2 µl of SUPERaseIn RNase Inhibitor (Invitrogen, 40 U µl⁻¹), 4 µl of 40% PEG8000 and 0.5 µl of 10 mM ATP, was added to 10 µl of reverse transcription product. The reaction was incubated at 55 °C for 30 minutes for tagmentation of RNA–cDNA hybrids. The reaction was then supplemented with 0.05 µl of Bst 3.0 DNA polymerase (New England Biolabs (NEB)), 20.5 µl of Q5 High-Fidelity 2× Master Mix (NEB) and 0.45 µl of ultrapure water, incubated at 72 °C for 15 minutes and then at 80 °C for 5 minutes. Next, 6 µl of primer mix, containing 1.5 µl of N50X primer, 1.5 µl of N70X primer from the TruePrep Index Kit V2 (Vazyme) and 3 µl of Q5 High-Fidelity 2× Master Mix, was added. The reaction mixture was amplified for 21 cycles and purified with 35 µl of VAHTS cleanup beads. The resulting DNA libraries were sequenced on an Illumina NovaSeq 6000 platform to generate 150-bp paired-end reads.

CAP-CELL labeling of CD40/CD40L-mediated cell–cell interactions

HEK293T cells were separately transfected with tdTomato-P2A-CD40L-Lantern or CD40-EGFP. Transfected cells were dissociated with 0.05% trypsin-EDTA for 1 minute, collected and counted. CD40L-Lantern cells (4×10^5) were mixed with CD40-EGFP cells (2×10^5) and incubated with 100 µM Biotin-NH₂ probe in 1 ml of PBS containing 10 mg ml⁻¹ BSA at 37 °C for 15 minutes. After blue light irradiation (30 mW cm⁻²) in a six-well plate, cells were collected by centrifugation at 700g for 2 minutes and washed three times with PBS. Cells were then stained with streptavidin-Alexa Fluor 647 (1:500) at room temperature for 30 minutes. After two washes, cell mixtures were analyzed using an LSRFortessa Cell Analyzer (BD Biosciences).

APEX2 labeling in live cells

HEK293T cells stably expressing APEX2-Sec61β were generated by lentiviral transduction using the same protocol as that used to establish Lantern-expressing cell lines. For dot blot and RT–qPCR analysis, cells were seeded in six-well plates and grown to approximately 90% confluence. Cells were incubated with 0.5 mM biotin-phenol in 1 ml of fresh culture medium at 37 °C for 30 minutes. Labeling was initiated by adding 10 µl of 100 mM H₂O₂ (final concentration 1 mM) and allowed to proceed for 1 minute at room temperature. The reaction medium was then removed, and cells were washed twice with 1 ml of PBS containing 5 mM Trolox, 10 mM sodium ascorbate and 10 mM sodium azide to quench residual labeling activity. Cells were lysed with 1 ml of TRIzol reagent. RNA extraction, enrichment, dot blotting and RT–qPCR analysis were performed as described for the Lantern-based experiments.

m⁶A profiling with GLORI2.0

mRNA was isolated from total RNA using the Dynabeads mRNA Purification Kit (Invitrogen) according to the manufacturer's instructions. Purified mRNA was fragmented in a 20-µl reaction containing 2 µl of 10× RNA fragmentation buffer (NEB, E6150S), incubated at 85 °C for 5 minutes and immediately placed on ice for 2 minutes. Subsequently, 2 µl of 10× stop buffer (NEB, E6150S) was added. The final solution was purified using an RNA Clean & Concentrator column (Zymo Research, R1016) according to the manufacturer's instructions.

The purified fragmented mRNA was added to a 50-µl glyoxal/nitrite reaction containing 13 µl of glyoxal solution (Sigma-Aldrich, 50649), 10 µl of DMSO, 10 µl of 5 M NaNO₂ (Sigma-Aldrich, S2252), 5 µl of saturated H₃BO₃ solution (Sigma-Aldrich, B0394), 4 µl of 500 mM MES (pH 6.0; Sigma-Aldrich, 76039) and nuclease-free water. The reaction was incubated at 50 °C for 10 minutes and immediately placed on ice for 2 minutes to stop the reaction. The modified RNA was then purified by ethanol precipitation.

The resulting RNA pellet was dissolved in 50 µl of deprotection buffer containing 0.5 M triethylammonium acetate (TEAA)

(pH 8.6) and 47.5% formamide. The mixture was incubated at 95 °C for 10 minutes without shaking and immediately placed on ice for 2 minutes. After deprotection, the RNA was purified again by ethanol precipitation.

The final RNA pellet was dissolved in 8.5 µl of nuclease-free water. The RNA library was constructed using the SMARTer Stranded Total RNA-Seq Kit v2 - Pico Input Mammalian (TaKaRa Bio, 634413) according to the manufacturer's instructions, except that the reverse transcription primer was replaced with a ternary primer.

GLORI-seq data analysis

Raw R2 reads were processed using Trim Galore (version 0.6.6, http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and SeqKit (version 0.13.2, <https://github.com/shenwei356/seqkit>) for adapter trimming and quality control. m⁶A sites were called using GLORI-tools (<https://github.com/liucongcas/GLORI-tools/>)³⁹. For each mRNA, an m⁶A strength score was calculated as the sum of m⁶A ratios across all identified m⁶A sites. Based on these scores, transcripts were classified into three m⁶A levels: high (≥ 3), medium (0.5 to < 3) and low (< 0.5). The enrichment ratio for each mRNA was calculated by dividing its normalized read count in G3BP1-Lantern cells by that in Lantern-NES cells. Transcripts with an average read count < 10 in Lantern-NES cells at 60 minutes after stress were excluded from the analysis.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The transcriptomic data generated in this study have been deposited in the Gene Expression Omnibus database under accession code [GSE292673](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE292673). The proteomic data have been deposited in the ProteomeXchange database under accession code [IPX0014765000](https://www.ebi.ac.uk/psd/entry/IPX0014765000). Lantern-expressing plasmids are available through Addgene: mito-Lantern ([250922](https://www.addgene.org/250922) and [250920](https://www.addgene.org/250920)), Lantern-HaloTag-NES ([250919](https://www.addgene.org/250919)), G3BP1-Lantern ([250918](https://www.addgene.org/250918)), Lantern-MAVS ([250917](https://www.addgene.org/250917)) and Lantern-Sec61β ([250916](https://www.addgene.org/250916)). Gel-based western blot data are provided in the Source Data file. Source data are provided with this paper.

Code availability

Python script for the statistics of mutation abundance in the mutant library from NGS results is provided as Supplementary Code 1.

References

58. Tsanov, N. et al. smiFISH and FISH-quant – a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res.* **44**, e165 (2016).
59. Di, L. et al. Rapid and sensitive single-cell RNA sequencing with SHERRY2. *BMC Biol.* **20**, 213 (2022).

Acknowledgements

We thank X. Zhou for assistance with mass spectrometry data analysis, G. Wang for generating the APEX2-Sec61β cell line and all laboratory members for helpful discussions. CD40/CD40L-related plasmids were generously provided by P. R. Chen's laboratory at Peking University. We thank the National Center for Protein Sciences at Peking University for assistance with Fragment Analyzer 12 analysis. Mass spectrometry measurements were performed at the Analytical Instrumentation Center of Peking University, with help from W. Zhou.

Author contributions

P.Z. and Y.F. conceptualized the project. Y.F., Z.R., F.Z., R.W. and P.Z. designed experiments. Y.F. and C. Lin performed directed evolution

experiments and Lantern characterizations. Y.F., Z.R. and F.Z. performed SG experiments. Y.F. and Y.Z. performed RNA methylation analysis. Z.R. and S.Z. performed primary neuron experiments. W.W. performed OMM transcriptome profiling experiments. Y.F., R.W., C. Li and L.L.-Y. performed selective cellular tagging experiments. Y.F., Z.R., F.Z., R.W., Y.Z., J.L. and P.Z. analyzed data. Y.F. and P.Z. wrote the paper, with input from all authors.

Funding

This work was supported by the Ministry of Science and Technology (2022YFA1304700), the National Natural Science Foundation of China (32088101) and Beijing National Laboratory for Molecular Sciences (BNLMS-CXXM-202403).

Competing interests

The authors declare no competing interests.

Additional information

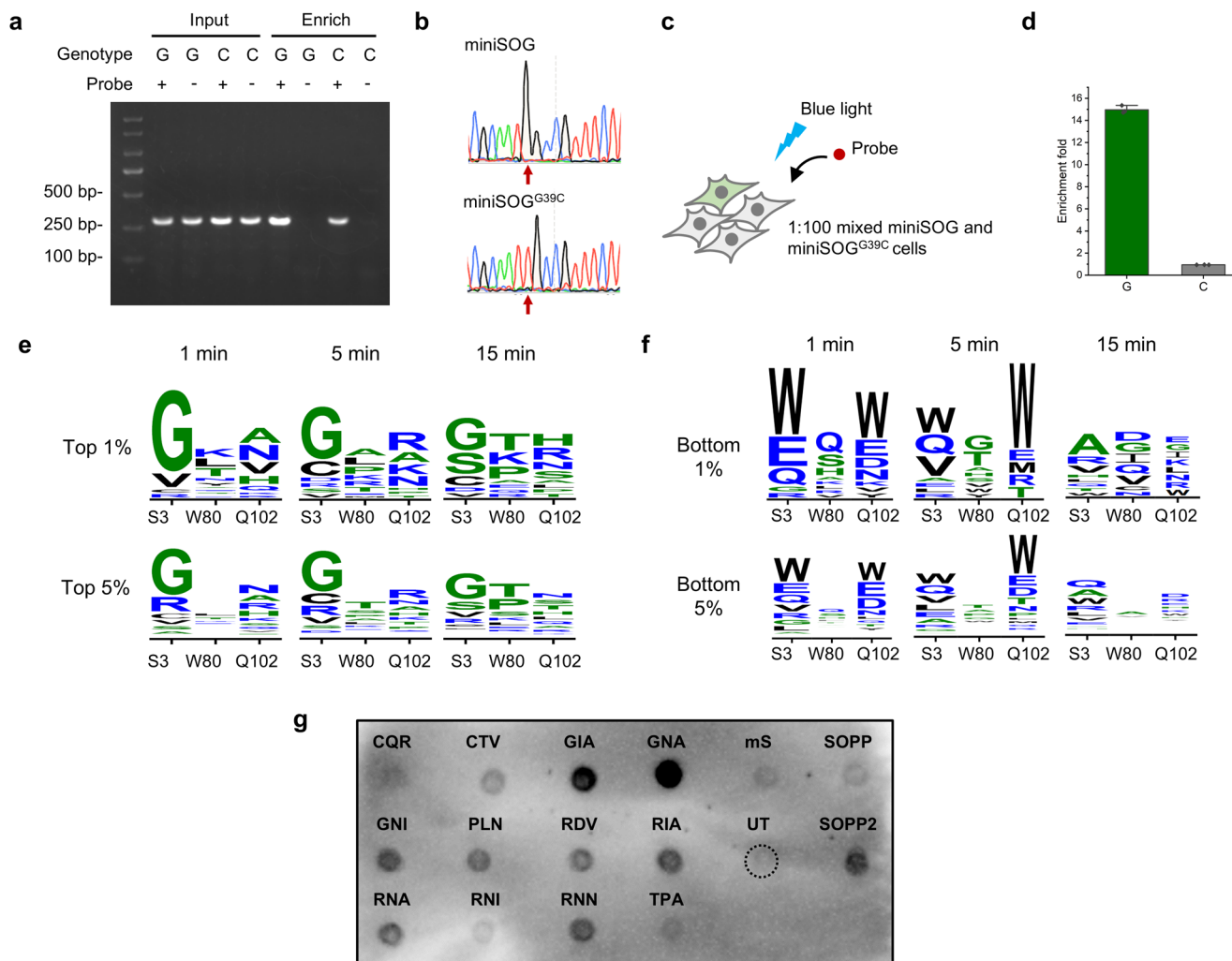
Extended data is available for this paper at <https://doi.org/10.1038/s41589-026-02313-y>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41589-026-02313-y>.

Correspondence and requests for materials should be addressed to Peng Zou.

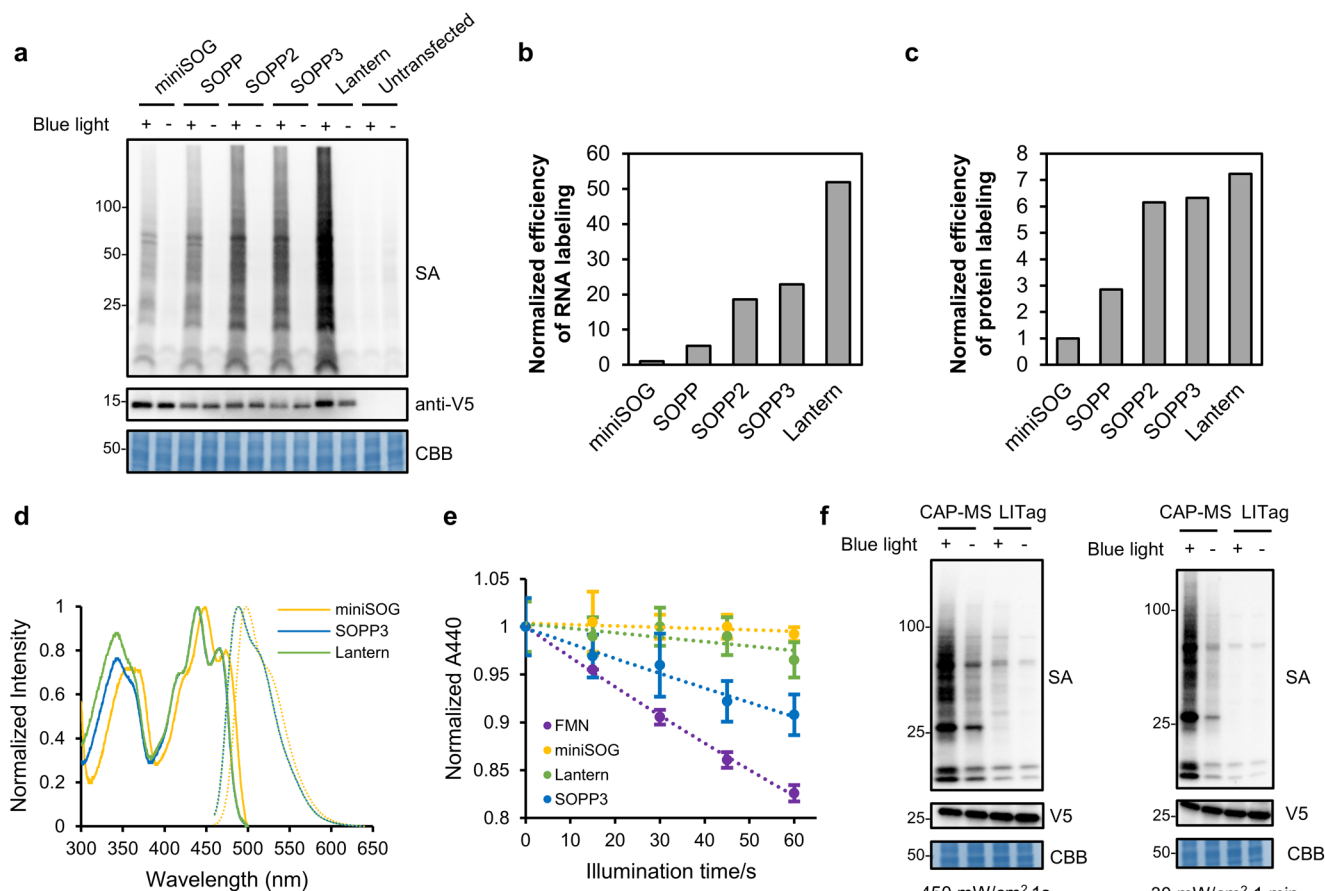
Peer review information *Nature Chemical Biology* thanks Michiel Vermeulen and the other, anonymous reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.



Extended Data Fig. 1 | Directed evolution of Lantern using a mammalian cell-based selection platform. (a) DNA gel electrophoresis analysis of RT-PCR products prepared from pre- and post-enrichment RNAs. Cells stably expressing tdMCP-miniSOG or tdMCP-miniSOG^{G39C} were labeled separately with 10 mM PA under 15 min blue-light irradiation at 30 mW/cm². After labeling, total RNA was extracted, conjugated with biotin, and enriched with streptavidin-coated beads. Eluted RNA was reverse transcribed and amplified with primers targeting the miniSOG sequence. The experiment was repeated independently twice with similar results. **(b)** Sanger sequencing analysis of recovered DNA around the mutation site. **(c)** Schematic of the model selection using mixed miniSOG- and

miniSOG^{G39C}-expressing cells. Mixed cells were labeled with 10 mM PA under 15 min blue-light irradiation at 30 mW/cm². **(d)** Enrichment folds of miniSOG and miniSOG^{G39C}, calculated from changes in sequence abundance after selection. Error bars: mean $\pm$ SD, $n = 3$ biological replicates. **(e-f)** Top-enriched **(e)** and bottom-enriched **(f)** amino acid residues in the first-round selection under 1, 5, and 15 min irradiation. **(g)** Streptavidin dot blot analysis of RNA labeling by the 12 top-enriched mutants. The three-letter abbreviations denote the amino acid residues at S3, W80, and Q102, respectively. UT, untransfected cells. The top mutant, 'GNA', was named Lantern.

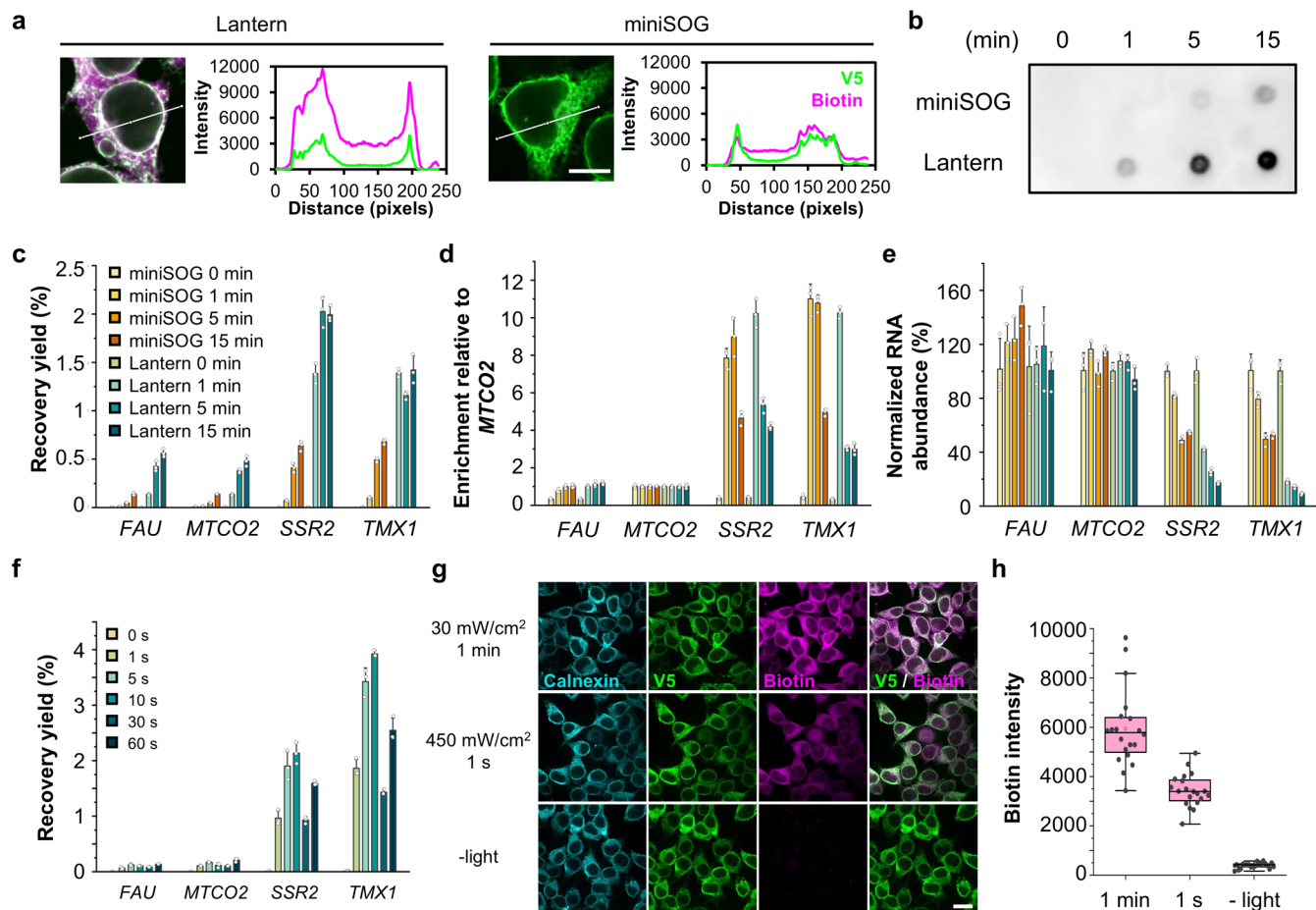


Extended Data Fig. 2 | Comparison of Lantern with other photocatalysts.

(a) Western blot analysis comparing labeling by miniSOG variants in HEK293T cells. Cells were transiently transfected with mitochondrial matrix-targeted miniSOG variants and labeled with 10 mM PA under 1 min irradiation. Molecular weight markers (kDa) are shown on the left. Similar results were observed in two biological replicates. **(b-c)** Quantification of RNA **(b)** and protein **(c)** labeling efficiencies of miniSOG variants. Labeling efficiency was calculated as the biotin intensity relative to V5 intensity and normalized to the labeling efficiency of miniSOG. RNA biotinylation data are from Fig. 1d. Protein biotinylation data and V5 intensities are from **(a)**. **(d)** Absorption and fluorescence spectra of miniSOG,

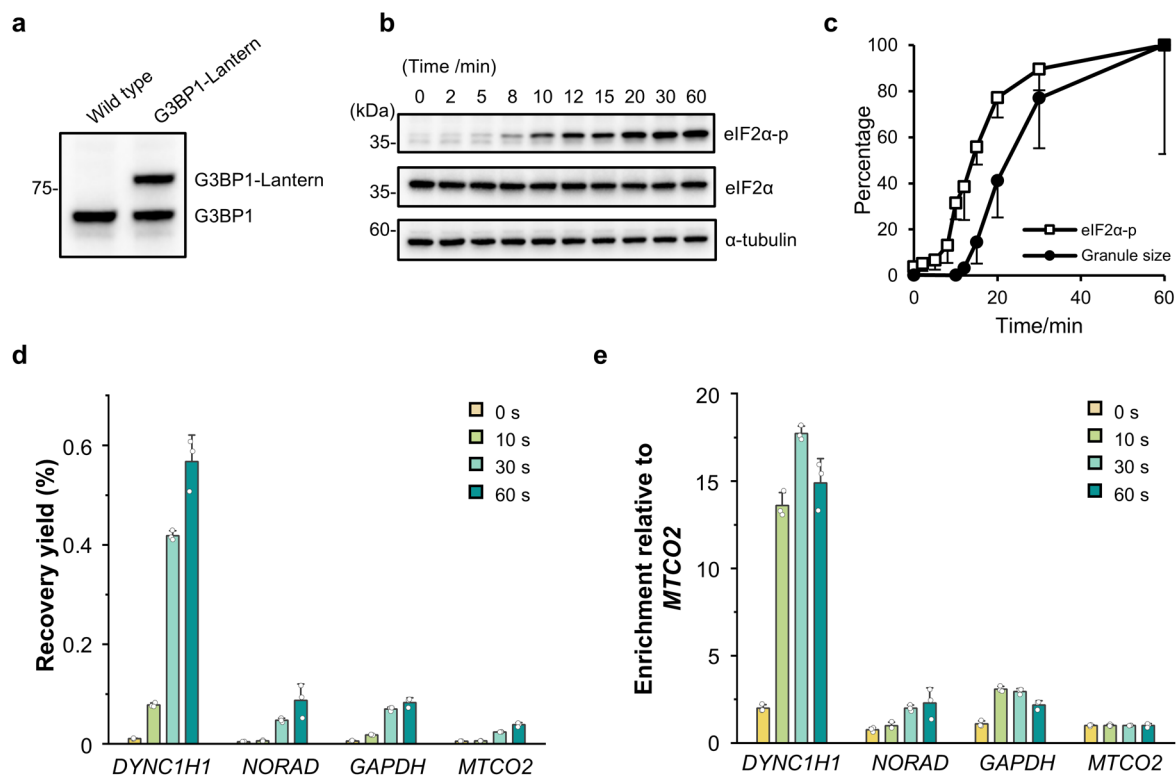
SOPP3, and Lantern. **(e)** Photobleaching curves of different photosensitizers upon blue-light irradiation. Error bars: mean $\pm$ SD, $n = 3$ technical replicates.

(f) Western blot analysis comparing CAP-MS and LITag labeling. Cells were transiently transfected with H2B-fused Lantern for CAP-MS or H2B-fused LOV* for LITag, and labeled with 1 mM 3-EA for CAP-MS or 0.5 mM BP for LITag under the same irradiation conditions. Left, 450 mW/cm² irradiation for 1 s; right, 30 mW/cm² irradiation for 1 min. Molecular weight markers (kDa) are shown on the left. Original Western Blot data of **(a)** and **(f)** have been provided in Source Data Fig. 1.



Extended Data Fig. 3 | Lantern-mediated RNA labeling at the ERM. **(a)** Line-scan analysis of labeling with ERM-Lantern and ERM-miniSOG under 1 min irradiation at 30 mW/cm². Scale bar: 10 μ m. **(b)** Streptavidin dot blot analysis of RNA labeling with ERM-miniSOG or ERM-Lantern. Cells were labeled with 10 mM PA under 1 to 15 min irradiation at 30 mW/cm². **(c-f)** RT-qPCR analysis of RNA labeling with ERM-miniSOG or ERM-Lantern. Cells stably expressing miniSOG-Sec61 β or Lantern-Sec61 β were labeled with 10 mM PA under 1 to 15 min blue-light irradiation at 30 mW/cm². Recovery yield **(c)** and normalized enrichment **(d)** were calculated from C_t values of samples before and after enrichment. RNA abundance before affinity purification is shown in **(e)**, with values scaled to the

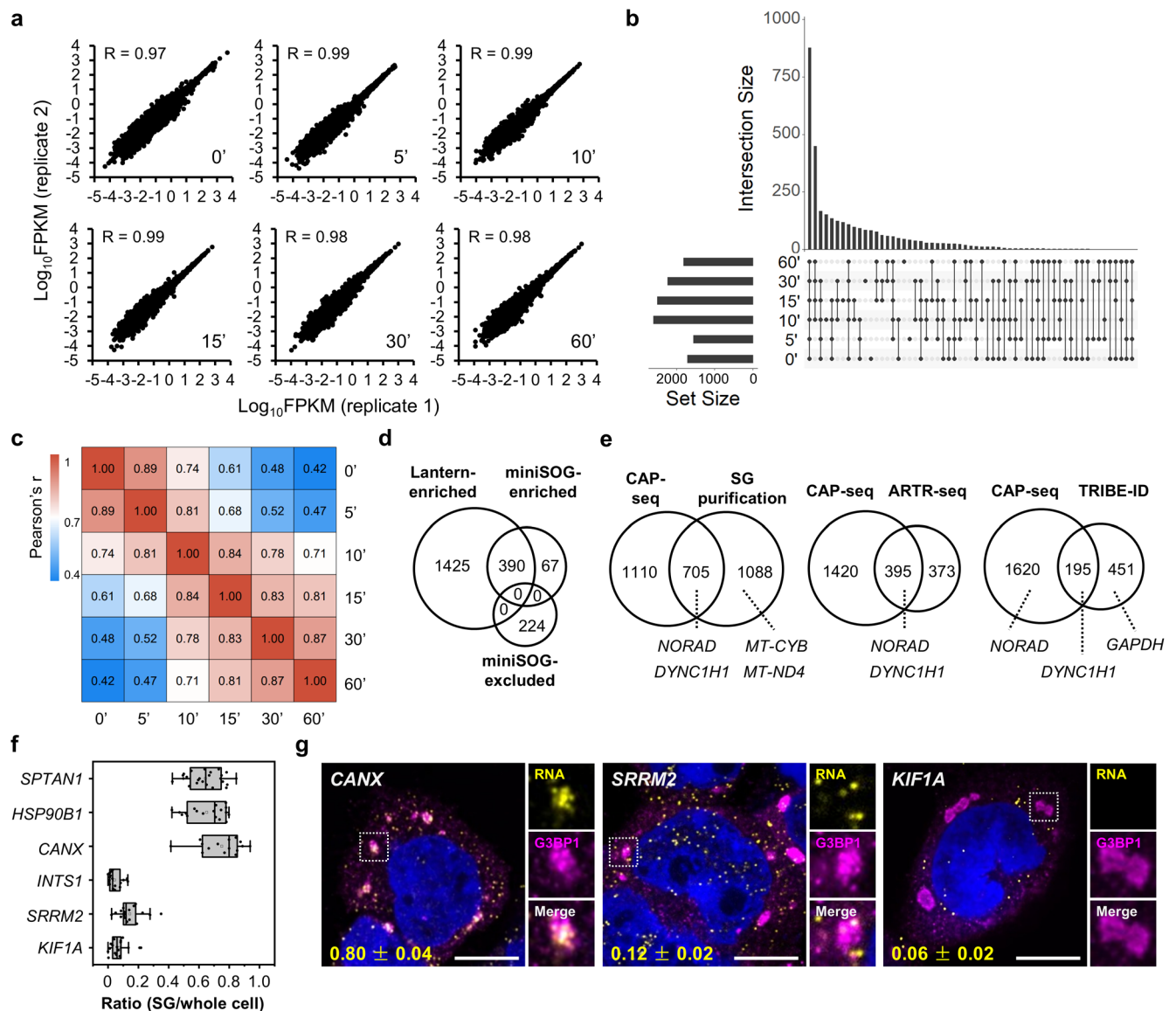
omit-light negative control ($t = 0$ min). **(f)** RT-qPCR analysis showing recovery yields of detected genes by Lantern-mediated CAP-seq. ERM-Lantern cells were labeled with 10 mM PA under 1 to 60 s irradiation at 450 mW/cm². *SSR2* and *TMX1* are secretory genes, whereas *FAU* and *MTCO2* are non-secretory genes. Error bars: mean $\pm$ SD; $n = 3$ technical replicates. **(g)** Representative immunofluorescence images of ERM-Lantern cells labeled with 10 mM PA and irradiated with blue light at 30 mW/cm² for 1 min or 450 mW/cm² for 1 s. Scale bar: 20 μ m. **(h)** Quantification of biotin signal intensities in **(g)**. Boxes indicate the first and third quartiles; whiskers indicate the outermost data points within 1.5 $\times$ the interquartile range; middle lines indicate the median. $n = 20$ cells.

**Extended Data Fig. 4 | Lantern-mediated RNA labeling at stress granules.**

(a) Western blot analysis of G3BP1 expression level in the G3BP1-Lantern stable cell line. Molecular weight markers (kDa) are shown on the left. The experiment was repeated independently three times with similar results. **(b)** Western blot analysis characterizing eIF2α phosphorylation levels during SG assembly. Original Western Blot data have been provided in Source Data Fig. 2. **(c)** Quantification of eIF2α phosphorylation and SG growth during SG assembly. Error bars represent one standard deviation below the mean. The eIF2α

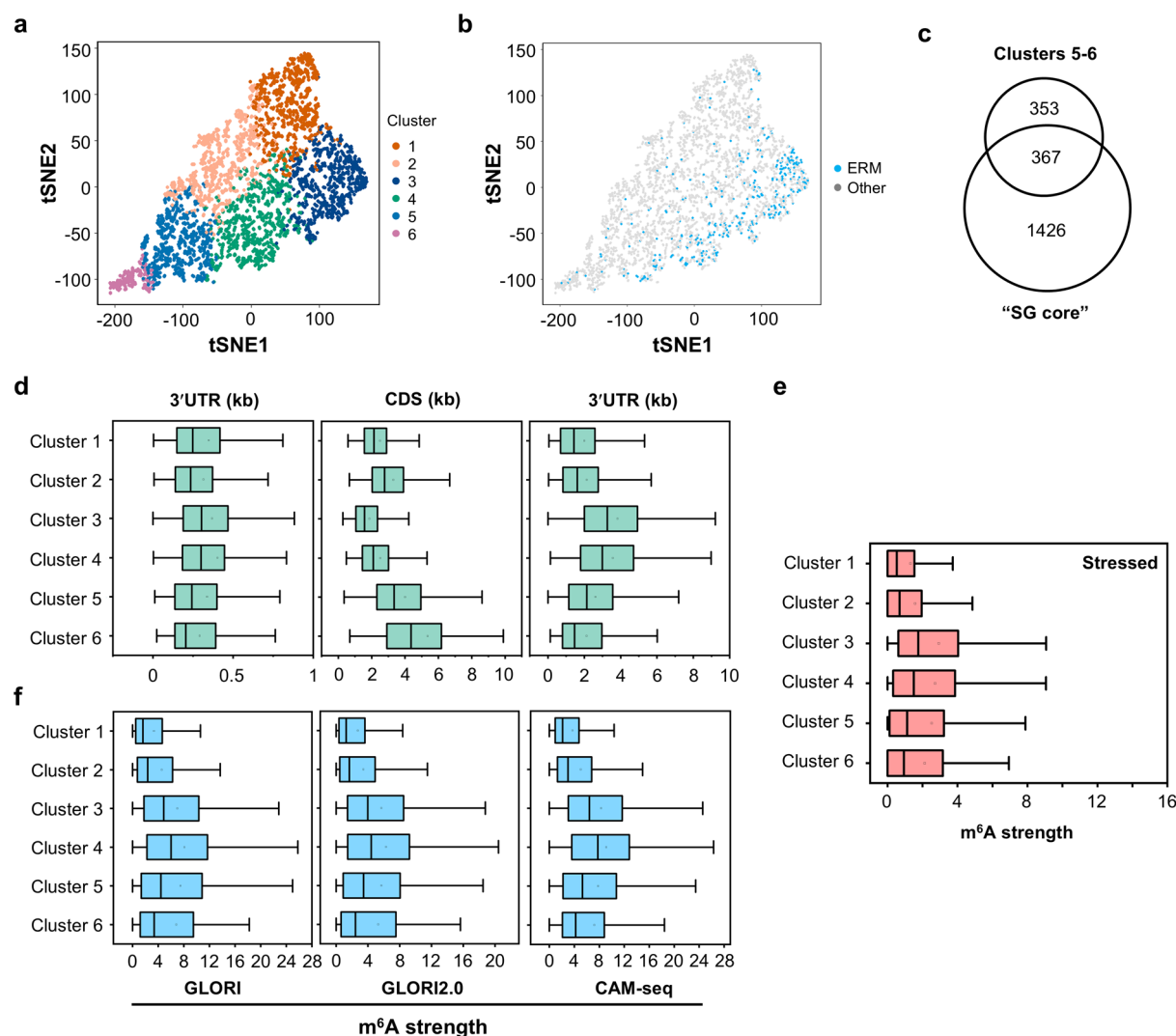
phosphorylation blot was performed with three biological replicates. The average granule size at each time point was calculated from 20 cells.

(d-e) RT-qPCR analysis showing recovery yields **(d)** and enrichment folds **(e)** of representative genes. G3BP1-Lantern cells were treated with 0.5 mM NaAsO₂ for 60 min and then labeled with 10 mM MPA under 10 to 60 s irradiation at 30 mW/cm². *DYNC1H1* and *NORAD* are previously reported SG-enriched transcripts. *GAPDH* is a house-keeping gene normally localized in the cytosol. *MTCO2* belongs to the mitochondrial transcriptome. Error bars: mean ± SD, n = 3 technical replicates.



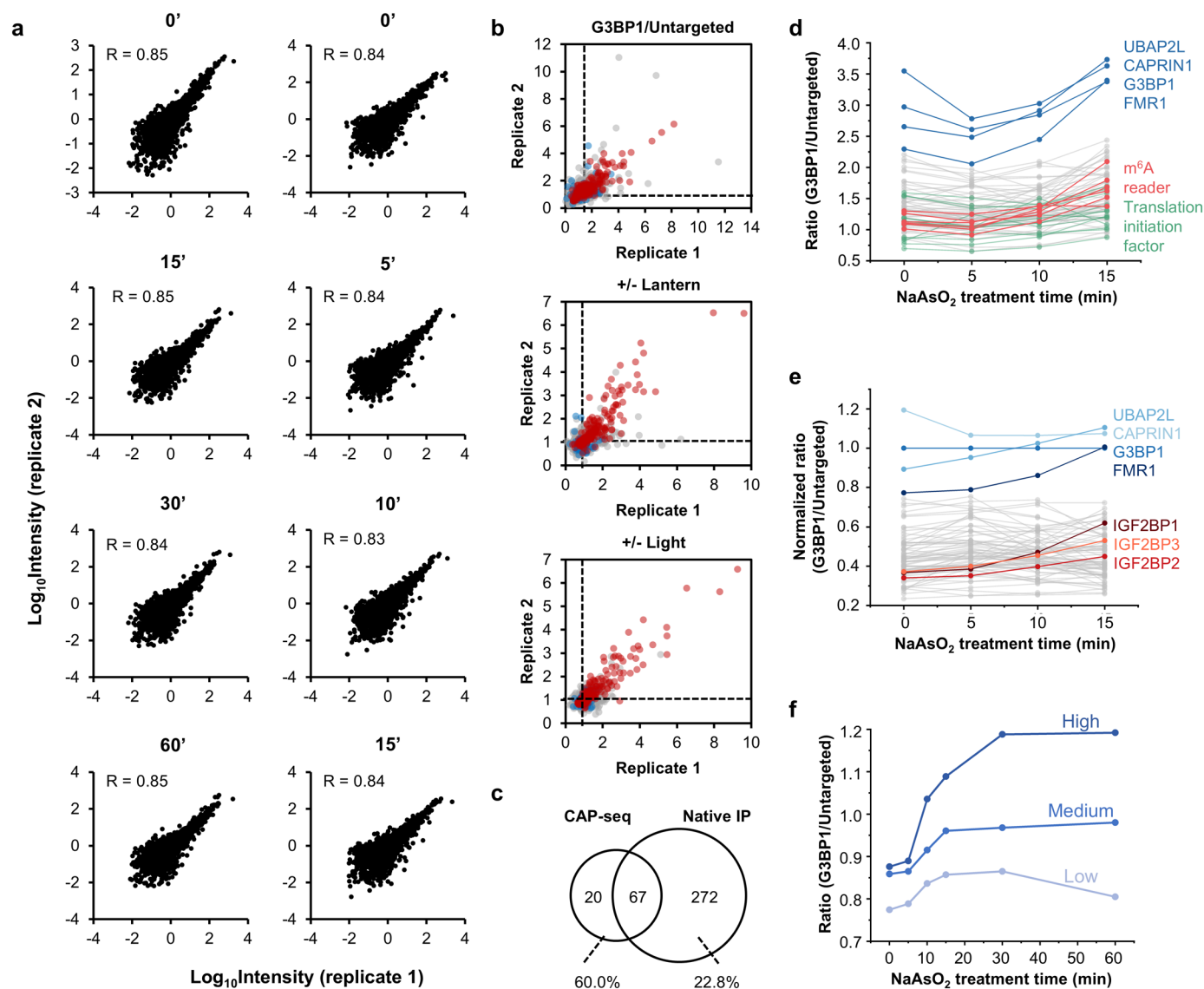
Extended Data Fig. 5 | Lantern CAP-seq during SG assembly in HEK293T cells. (a) FPKM values of transcripts commonly detected in two replicates. Pearson's correlation coefficients between replicates are indicated in each plot. (b) Temporal overlap of G3BP1-proximal RNAs during SG assembly. Horizontal bars on the left indicate the total number of enriched transcripts at each time point. Vertical bars show the number of transcripts shared by the time-point combinations indicated by the connected dots below. (c) Heatmap showing Pearson's correlation coefficients between Log₂FC (G3BP1 vs. Untargeted) values determined at different time points during SG assembly. (d) Venn diagram comparing the SG transcriptome identified by Lantern-mediated CAP-seq with SG-proximal and SG-excluded datasets determined by miniSOG-mediated CAP-seq¹⁸. (e) Venn diagram comparing the SG transcriptome identified by CAP-seq

with datasets from SG purification³⁶, ARTR-seq³⁵, and TRIIBE-ID³⁷. The ARTR-seq dataset was analyzed using DESeq2 cutoffs of log₂FC > 1 and $P < 0.05$ (two-sided Wald test). The TRIIBE-ID dataset was obtained after 1 h treatment with rapamycin and sodium arsenite. *DYNC1H1* and *NORAD* are true positives; *GAPDH*, *MT-CYB* and *MT-ND4* are true negatives. (f) Quantification of the ratio of smFISH signals within SGs versus the whole cell. Boxes indicate the first and third quartiles; whiskers indicate the outermost data points within 1.5× the interquartile range; open squares denote the mean; middle lines indicate the median. $n = 15$ cells. (g) Representative smFISH images of the enriched SG transcript *CANX* and non-enriched transcripts *SRRM2* and *KIF1A*. Scale bar: 10 μ m. The colocalization ratio of each transcript is indicated at the bottom left as median $\pm$ S.E.M.



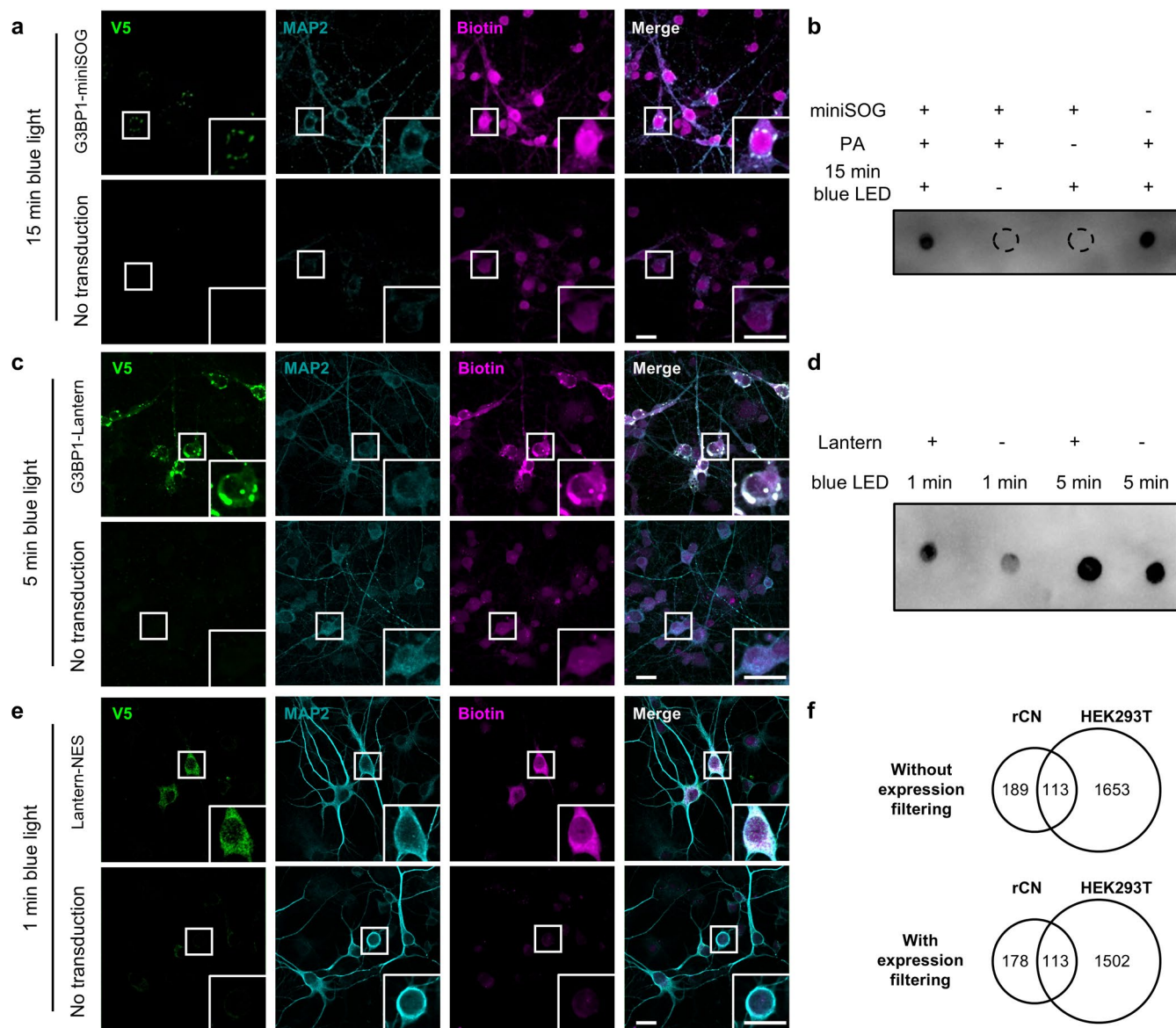
Extended Data Fig. 6 | Cluster analysis of the SG transcriptome during assembly. (a–b) t-SNE analysis of 3,289 G3BP1-interacting RNAs based on their $\log_2$ FC (G3BP1/NES) values across six time points during SG assembly. RNAs are annotated by the six assigned clusters (a) or ERM localization identified in unstressed cells (b). (c) Venn diagram comparing transcripts from clusters 5 and 6 with the SG core transcriptome obtained by SG core purification³⁶. (d–f) Box plots comparing gene length (d) and m⁶A strength (e–f) for RNAs in clusters 1–6.

Gene length features were obtained from Ensembl. m⁶A strength was calculated either from m⁶A maps generated in this study by applying the GLORI2.0 protocol to HEK293T cells treated with 0.5 mM NaAsO₂ for 1 h (e), or from published GLORI³⁸, GLORI2.0³⁹, and CAM-seq⁴⁰ datasets (f). Boxes indicate the first and third quartiles; whiskers indicate the outermost data points within 1.5× the interquartile range; open squares denote the mean; middle lines indicate the median. The n value for each box is the same as in Fig. 3e.



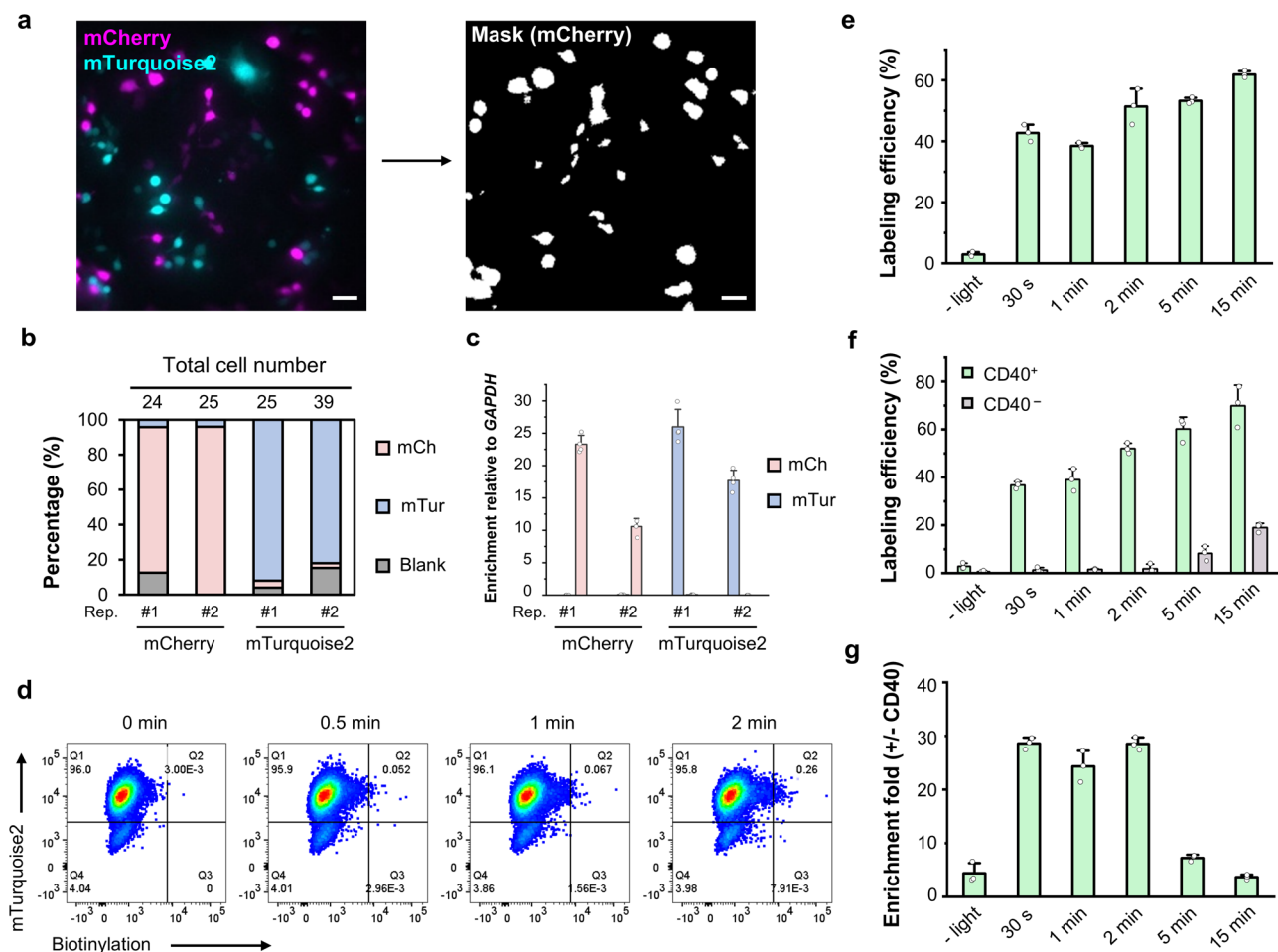
Extended Data Fig. 7 | Temporally resolved profiling of the SG proteome during assembly. (a) Correlation analysis of mass spectrometry intensities for proteins detected in replicate experiments. Pearson's correlation coefficients (R) are shown at the top. (b) Correlation analysis of SG protein enrichment in replicate labeled samples compared with controls. Enrichment folds were calculated from normalized protein intensities. Red: canonical SG proteins, defined as Tier 1 proteins in RNA Granule Database⁴⁴; blue: extracellular or mitochondrial matrix proteins, used as true negatives. Dashed lines denote the

cutoff ratio of 1.1. (c) Venn diagram comparing SG proteome datasets identified by Lantern-mediated CAP-MS and native immunoprecipitation (Native IP)⁴⁸, with the percentage of Tier 1 proteins indicated on the side. (d–e) G3BP1/Untargeted ratios of SG proteins before (d) and after normalization to G3BP1 (e) during the first 15 min of stress. (f) G3BP1/Untargeted ratios of human RNAs with low, medium, and high m⁶A strength during the 1 h stress time. The m⁶A map was generated from unstressed HEK293T cells using the GLORI2.0 procedure.



Extended Data Fig. 8 | Optimization of stress granule CAP-seq labeling in primary neuron culture. (a) Representative immunofluorescence images of rat cortical neurons labeled with G3BP1-miniSOG in the presence of 10 mM PA under 15 min irradiation. (b) Streptavidin dot blot analysis of RNAs labeled by miniSOG under 15 min irradiation. (c) Representative immunofluorescence images of rat cortical neurons labeled with G3BP1-Lantern in the presence of 10 mM PA under 5 min irradiation. (d) Streptavidin dot blot analysis of RNAs labeled by Lantern under 1 or 5 min irradiation. (e) Representative immunofluorescence images

of rat cortical neurons labeled with Lantern-NES in the presence of 10 mM PA under 1 min irradiation. (f) Venn diagram comparing SG transcriptome datasets generated by Lantern-mediated CAP-seq in rat cortical neurons (rCN) and HEK293T cells. Transcripts without homologous annotations were removed. The analysis was repeated after removing low-abundance transcripts with read counts <10 (bottom). All imaging experiments were repeated independently twice with similar results. Scale bar: 20 μ m.

**Extended Data Fig. 9 | Spatially resolved cellular tagging with Lantern.**

(a) DMD masks generated from fluorescence images of the mCherry or mTurquoise2 channels. **(b)** Percentage distribution of different cell types in the sorted cell population. **(c)** RT-qPCR analysis of normalized *mCherry* (pink) and *mTurquoise2* (blue) expression levels in sorted cells. Cells were selectively labeled based on mCherry or mTurquoise2 expression, with two replicates for each condition. Expression levels were normalized to *GAPDH*. Error bars:

mean $\pm$ SD; $n = 4$ technical replicates. **(d)** Flow cytometry analysis of Lantern-TMMC38 cells labeled with 2 mM Btn-NH₂ under 0 to 2 min irradiation.

(e-g) Quantitation of CD40/CD40L-mediated cell-cell labeling by flow cytometry. *Cis*-labeling indicates self-biotinylation of CD40L-Lantern bait cells **(e)**, whereas *trans*-labeling indicates biotinylation of CD40⁺-expressing prey cells **(f)**.

Normalized *trans*-labeling efficiency is shown in **(g)**. Error bars: mean $\pm$ SD; $n = 3$ technological replicates.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
☐	☒ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☐	☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒ Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection	Illumina NovaSeq 6000 system
Data analysis	HISAT2 (version 2.1.0), HTSeq (version 0.6.1), DESeq2 (version 1.42.1), pheatmap (version 1.0.12), samtools (version 1.6.1), ImageJ (version 2.3.0), MaxQuant (version 1.6.10)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The transcriptomic data generated in this study have been deposited in the GEO database under the accession code GSE292673. The proteomic data generated in this study have been deposited in the ProteomeXchange database under accession code IPX0014765000. Lantern-expressing plasmids are available through Addgene (mito-Lantern: 250922 and 250920; Lantern-HaloTag-NES: 250919; G3BP1-Lantern: 250918; Lantern-MAVS: 250917; Lantern-Sec61β: 250916). The gel-

based western blot data generated in this study are provided in the Source Data file. The raw data of all the bar charts in this study are provided in the Source Data file. The PDB entries used in this work could be found in the following hyperlinks: 6GPU (miniSOG). Source data are provided with this paper.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	This research doesn't involve human participants
Reporting on race, ethnicity, or other socially relevant groupings	This research doesn't involve human participants
Population characteristics	This research doesn't involve human participants
Recruitment	This research doesn't involve human participants
Ethics oversight	This research doesn't involve human participants

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes of omic studies are indicated in methods. For RNA sequencing experiments, libraries constructed from total RNA were sequenced at a depth of 50 million paired reads per library to ensure a sufficient coverage of protein-coding RNAs. Libraries constructed from rRNA-depletion samples were sequenced at a depth of 10 million paired reads per library. The sample size of proteomic experiments was determined by Western Blot and Silver staining after biotin affinity precipitation.
Data exclusions	No data exclusion was performed.
Replication	The numbers of replicated experiments are indicated in methods. At least two biological replicates were performed for high-throughput sequencing and proteomic experiments. All the confocal images, blots and gels were repeated twice independently with similar results.
Randomization	Immortal cells and primary neurons were randomly seeded in culture dishes.
Blinding	Blinding is not relevant to our study. Given the high-throughput, objective nature of next-generation sequencing and proteomic analysis, these data are inherently quantitative and not susceptible to observer bias.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Antibody (vendor, catalog number, dilution): Rabbit anti-G3BP (Abcam, ab181150, 1:2000 for WB); Mouse anti-G3BP (Abcam, ab56574, 1:500); anti-G3BP2 (Abcam, ab86135, 1:500); anti-MAP2 (Abcam, ab5392, 1:5000); anti-TOMM20 (Abcam, ab56783, 1:1000); anti-Calnexin (Abcam, ab22595, 1:500); anti-DDX6 (Abclonal, A16270, 1:1000); anti-eIF2 α (CST, 5324, 1:3000); anti-EIF2S1 (Abcam, ab32157, 1:3000); Mouse anti-V5 (Biodragon, B1005, 1:1000 for IF and 1:5000 for WB); anti-Flag (Sigma-Aldrich, F1804, 1:1000); anti- α -Tubulin (Biodragon, B1052, 1:5000); Goat anti-mouse-Alexa Fluor 488 (ThermoFisher, A-11029, 1:1000); Goat anti-rabbit-Alexa Fluor 488 (ThermoFisher, A-11034, 1:1000); Streptavidin-Alexa Fluor 647 (ThermoFisher, S21374, 1:1000); Goat anti-mouse IgG-HRP (Biodragon, BF03001, 1:5000); Goat anti-rabbit IgG-HRP (Biodragon, BF03008, 1:5000); Streptavidin-HRP (ThermoFisher, 21124, 1:5000).
Validation	Used according to manufacturer's instructions. Rabbit anti-G3BP: https://www.abcam.cn/products/primary-antibodies/g3bp-antibody-epr13986b-ab181150 Mouse anti-G3BP: https://www.abcam.cn/products/primary-antibodies/g3bp-antibody-2f3-ab56574 anti-G3BP2: https://www.abcam.cn/products/primary-antibodies/g3bp2-antibody-ab86135 anti-MAP2: https://www.abcam.cn/products/primary-antibodies/map2-antibody-neuronal-marker-ab5392 anti-TOMM20: https://www.abcam.cn/products/primary-antibodies/tomm20-antibody-4f3-bsa-and-azide-free-ab56783 anti-Calnexin: https://www.abcam.cn/products/primary-antibodies/calnexin-antibody-er-marker-ab22595 anti-DDX6: https://abclonal.com.cn/catalog/A16270 anti-eIF2 α : https://www.cellsignal.com/products/primary-antibodies/eif2-alpha-d7d3-rabbit-monoclonal-antibody/5324 anti-EIF2S1: https://www.abcam.cn/products/primary-antibodies/eif2s1-phospho-s51-antibody-e90-ab32157 Mouse anti-V5: https://www.biodragon.cn/cn/goods/goodsView?GoodsId=1785 anti-Flag: https://www.sigmaaldrich.com/HK/zh/product/sigma/f1804 anti- α -Tubulin: https://www.biodragon.cn/cn/goods/goodsView?GoodsId=12552&Catalog=

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)	HEK293T cells (3101HUMSCSP5293) and U-2 OS cells (1101HUM-PUMC000028) were purchased from Cell Resource Center, Institute of Basic Medical Sciences, CAMS. MC38 cells were obtained from Prof. Peng Chen's laboratory at Peking University.
Authentication	Cells have not been authenticated.
Mycoplasma contamination	Cells used in this paper were test negative of mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	None.

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals	Neonatal Sprague Dawley rats (within 24 h post-birth) were used for cortical neuron dissection.
Wild animals	None.
Reporting on sex	Sex-based analyses were not performed in this study.
Field-collected samples	This study didn't involve samples collected from the field.
Ethics oversight	Approved by Peking University Institutional Animal Care and Use Committee (PKU IACUC).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Plants

Seed stocks	None.
Novel plant genotypes	None.
Authentication	None.

Flow Cytometry

Plots

Confirm that:

- ☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☒ All plots are contour plots with outliers or pseudocolor plots.
- ☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

For labeling of mTurquoise and mCherry co-culture system, HEK293T cells were transfected with Lantern-TM co-expressing mTurquoise or mCherry separately. Cells were mixed after transfection, labeled with the CAP-CELL procedure and digested with trypsin-EDTA (0.05%). Then the cells were stained with streptavidin-Alexa Fluor 647 and Calcein-AM, and analyzed by flow cytometry.

For labeling of migrating cancer cells, the MC38 cells stably expressing Lantern-TM were generated by lentiviral transduction. After the CAP-CELL labeling procedure, cells were digested and analyzed as described above.

For labeling of the CD40/CD40L-guided cell-cell interaction, HEK293T cells were transfected with CD40L-Lantern and CD40-EGFP separately. Cells were digested after the transfection was completed. 4e5 CD40L-Lantern cells were mixed with 2e5 CD40-EGFP cells, then incubate with 100 μ M Btn-NH2 probe in 1 mL PBS solution containing 10 mg/mL BSA at 37°C for 15 min. After the CAP-CELL labeling procedure, cells were resuspended, stained and analyzed as described above.

Instrument

BD FACSAriaIII Cell Sorter; BD LSRFortessa Cell Analyzer.

Software

FlowJo (version 10.6.2)

Cell population abundance

The abundance of mCherry+ and mTurquoise+ cells in the post sorted samples was determined by the fluorescence of mCherry and mTurquoise, respectively. The abundance of cell populations in each replicate was shown in Figure S6.

Gating strategy

The FSC/SSC gate was determined according to intensities of wild type HEK293T or MC38 cells. The biotinylation gate was determined according to the intensity of control samples. An illustration of gating strategy was shown in Figure S6.

- ☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.