nature chemical biology

ARTICLES

https://doi.org/10.1038/s41589-021-00874-8

Activity-based RNA-modifying enzyme probing reveals DUS3L-mediated dihydrouridylation

Wei Dai[1,5], Ang Li[1,5], Nathan J. Yu[1], Thao Nguyen[2,3,4], Robert W. Leach[3], Martin Wühr[2,3] and Ralph E. Kleiner[1]✉

Epitranscriptomic RNA modifications can regulate RNA activity; however, there remains a major gap in our understanding of the RNA chemistry present in biological systems. Here we develop RNA-mediated activity-based protein profiling (RNABPP), a chemoproteomic strategy that relies on metabolic RNA labeling, mRNA interactome capture and quantitative proteomics, to investigate RNA-modifying enzymes in human cells. RNABPP with 5-fluoropyrimidines allowed us to profile 5-methylcytidine (m^5C) and 5-methyluridine (m^5U) methyltransferases. Further, we uncover a new mechanism-based crosslink between 5-fluorouridine (5-FUrd)-modified RNA and the dihydrouridine synthase (DUS) homolog DUS3L. We investigate the mechanism of crosslinking and use quantitative nucleoside liquid chromatography–tandem mass spectrometry (LC–MS/MS) analysis and 5-FUrd-based crosslinking and immunoprecipitation (CLIP) sequencing to map DUS3L-dependent dihydrouridine (DHU) modifications across the transcriptome. Finally, we show that DUS3L-knockout (KO) cells have compromised protein translation rates and impaired cellular proliferation. Taken together, our work provides a general approach for profiling RNA-modifying enzyme activity in living cells and reveals new pathways for epitranscriptomic RNA regulation.

Postsynthetic modifications on biological macromolecules play important roles in biological processes. Among the central macromolecules of life, RNA exhibits perhaps the greatest diversity in modification chemistry[1]. RNA modifications are conserved throughout biology and are critical for the proper function of tRNA and rRNA[2]. In addition, studies from the past decade have revealed that mRNA is subject to diverse post-transcriptional modifications (known as the 'epitranscriptome'). These modifications, best exemplified by N^6-methyladenosine (m^6A), can impact mRNA behavior and regulate gene expression[3]. Moreover, studies of epitranscriptomic modifications have indicated that their dysregulation may be involved in disease etiology[4].

Our understanding of m^6A on RNA has been rapidly maturing, largely due to powerful technologies for mapping individual modification sites[5] and characterization of associated writer, eraser and reader proteins[6]. By contrast, insights into other RNA modifications have accumulated more slowly. For example, modifications on pyrimidines, including m^5C, m^5U, DHU, pseudouridine (ψ) and others, are abundant in non-coding RNA and conserved throughout evolution, but we lack insight into their biological role. For many modifications, progress has been hampered by a dearth of reliable strategies for mapping modifications across the transcriptome and identifying writer enzymes, as is needed for functional studies. Further, while nature has evolved rich biosynthetic machinery for modifying tRNA, the extent to which these modifications exist on mRNA is largely unknown.

Identifying RNA writer enzymes and their substrates is a major challenge, and these proteins are typically deduced based on homology or through screening. An alternative approach to enzyme discovery and profiling, known as activity-based protein profiling (ABPP)[7], relies on reactive small-molecule probes that form covalent adducts with enzyme families based on distinctive chemical features. ABPP provides a general platform for proteomic studies based on chemical reactivity but is typically limited to examples where a small molecule can mimic the native substrate, making its application to RNA-modifying enzymes challenging.

Here we develop RNABPP (Fig. 1a), a reactivity-based approach to profile RNA-modifying enzymes in living cells. RNABPP relies on metabolic labeling with 5-fluorocytidine (5-FCyd) (Fig. 1b), a cytidine analog that is efficiently incorporated into RNA and can form stable, mechanism-based adducts with RNA-modifying enzymes. We combine this warhead with mRNA interactome capture[8,9] and quantitative MS[10] to profile RNA-modifying enzymes in their native context. In addition to identifying mammalian m^5C RNA methyltransferases, which are predicted to form crosslinks with 5-FCyd-containing RNA based on their catalytic mechanism, our approach captures m^5U methyltransferases as well as DUS3L, the mammalian homolog of the yeast DUS, DUS3. We investigate the crosslinking mechanism to establish that 5-halopyrimidines can function as activity-based probes for DUS3L. Further, we characterize its cellular RNA substrates using quantitative nucleoside LC–MS/MS and individual nucleotide-resolution CLIP (iCLIP)-based sequencing[11]. Finally, we show that DUS3L regulates cell proliferation and protein translation efficiency. Our work provides an unbiased, reactivity-based platform for RNA-modifying enzyme discovery and characterization and expands our understanding of epitranscriptomic modifications in human RNA.

Results

Metabolic labeling with 5-FCyd. To develop our strategy (Fig. 1a), we needed a versatile chemical warhead that is efficiently incorporated into cellular RNA and forms stable, covalent adducts with RNA-modifying enzymes. For this purpose, we chose 5-FCyd, a mechanism-based inhibitor of RNA m^5C methyltransferases[12] (Fig. 1b). Because 5-FCyd is isosteric with cytidine, we reasoned that it should be efficiently incorporated into cellular RNA and

[1]Department of Chemistry, Princeton University, Princeton, NJ, USA. [2]Department of Molecular Biology, Princeton University, Princeton, NJ, USA. [3]Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA. [4]Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ, USA. [5]These authors contributed equally: Wei Dai, Ang Li. ✉e-mail: rkleiner@princeton.edu

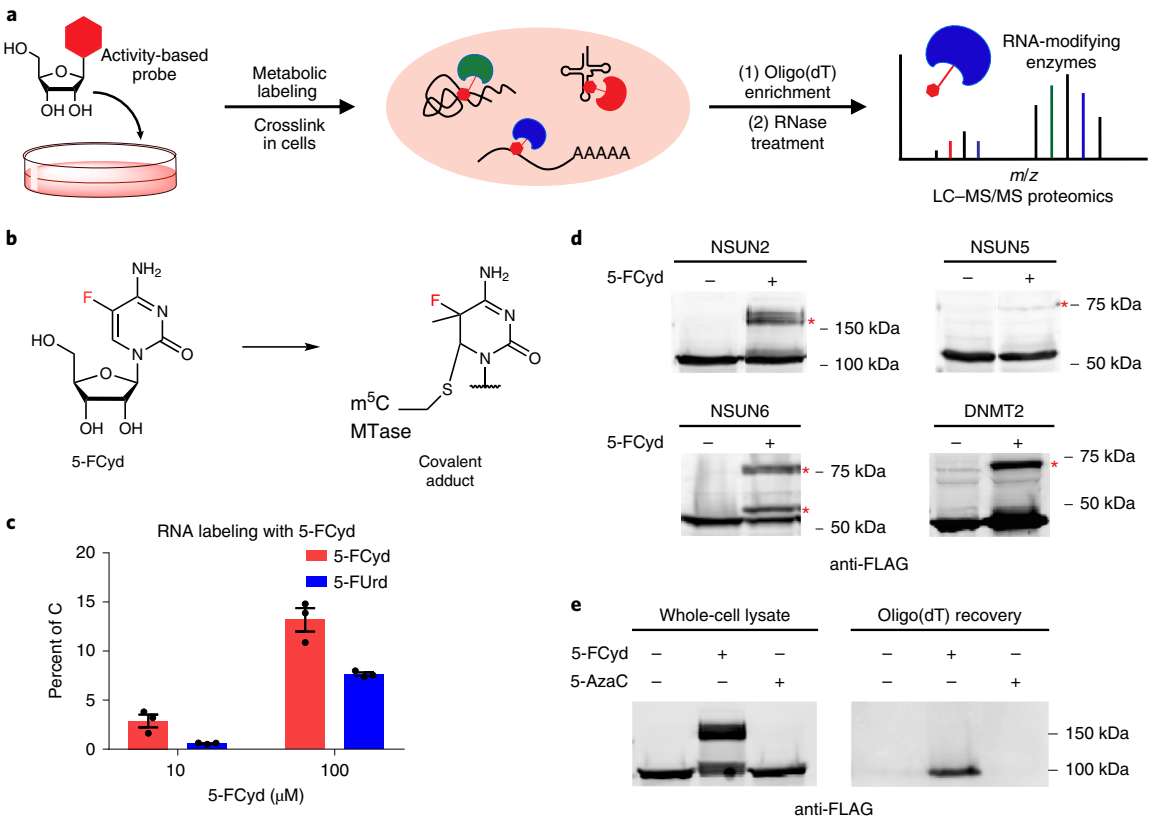

Fig. 1 | RNABPP enables the chemoproteomic analysis of RNA-modifying enzymes in living cells. a, Schematic representation of the RNABPP approach. Metabolic labeling with a nucleoside-based activity probe facilitates the incorporation of a chemical 'warhead' into cellular RNA. Crosslinked RNA-modifying enzymes on mRNA can then be identified through oligo(dT)-based enrichment and quantitative proteomics analysis. **b**, Chemical structure of 5-FCyd and proposed RNA–protein crosslinking between 5-FCyd-labeled RNA and m^{5}C RNA methyltransferases; MTase, methyltransferase. **c**, LC–MS/MS analysis of 5-FCyd and 5-FUrd levels in total RNA after metabolic labeling of HEK293T cells for 12 h with 5-FCyd. Three independent biological replicates were analyzed. Data represent mean values ± s.e.m. **d**, Western blot analysis of protein–RNA crosslinking in cells after 5-FCyd labeling. NSUN2, NSUN5, NSUN6 and DNMT2 were expressed as 3 × FLAG fusion proteins, and cells were treated with 10 μM 5-FCyd (for NSUN2 and NSUN6) and 100 μM 5-FCyd (for DNMT2 and NSUN5) for 12 h. The red asterisk indicates a slower migrating band. **e**, Oligo(dT)-based isolation of NSUN2 using RNABPP with 5-FCyd. Cells were fed with 10 μM 5-FCyd or 10 μM 5-AzaC or were untreated and then subjected to the RNABPP workflow. For **d** and **e**, the experiments were repeated three times independently with similar results.

minimally perturb recognition by RNA-modifying enzymes. Further, we expected that adducts formed between 5-FCyd-RNA and m^{5}C methyltransferases (Fig. 1b) would be stable, in contrast to the widely used inhibitor 5-azacytidine (5-AzaC)[13]. Finally, we suspected that 5-FCyd would be deaminated in cells to generate 5-FUrd[14], enabling reactivity-based profiling of enzymes modifying either cytidine or uridine.

To investigate metabolic labeling with 5-FCyd in cells, we used quantitative LC–MS/MS. Gratifyingly, treatment of HEK293T cells with 5-FCyd for 12 h resulted in efficient total RNA incorporation with 3% 5-FCyd/C at 10 μM treatment and 13% 5-FCyd/C at 100 μM treatment (Fig. 1c and Supplementary Table 1) without overt cytotoxicity (Supplementary Fig. 1). Further, we measured metabolic conversion of 5-FCyd to 5-FUrd with 0.6% and 8% RNA labeling with 5-FUrd using 10 μM and 100 μM 5-FCyd treatment conditions, respectively (Fig. 1c and Supplementary Table 1). Incorporation of 5-FCyd into mRNA proceeded with similar efficiency (Supplementary Fig. 2).

Moving forward, we evaluated crosslinking between 5-FCyd-labeled RNA and m^{5}C methyltransferases in cells. While 5-FCyd-modified RNA oligonucleotides crosslink with bacterial RNA m^{5}C methyltransferases in vitro[12], cellular crosslinking has been typically accomplished using 5-AzaC[13]. To test 5-FCyd-mediated crosslinking, we expressed several m^{5}C methyltransferases in HEK293T or Flp-In 293 cells, including NOP2/Sun RNA methyltransferase 2 (NSUN2), NSUN5, NSUN6 and DNA methyltransferase 2 (DNMT2), treated cells with 5-FCyd and evaluated RNA–protein crosslinking by western blotting. In all cases, we observed a slower migrating band following 5-FCyd treatment (Fig. 1d), consistent with crosslinking likely to abundant RNAs of uniform size, such as tRNA substrates of NSUN2, NSUN6 and DNMT2. Further, with NSUN2, we observed no crosslinking when 5-AzaC was used in place of 5-FCyd (Fig. 1e), likely due to the instability of the 5-AzaC adduct.

Proteomic profiling of 5-FCyd-reactive proteins. After establishing 5-FCyd-m^{5}C methyltransferase crosslinking in cells, we tested whether we could isolate crosslinked RNA–protein complexes for unbiased proteomic analysis. For this purpose, we chose mRNA interactome capture[8,9], which relies on oligo(dT) hybridization to isolate protein–mRNA complexes (typically generated by photocrosslinking) directly from lysate. Because this approach should primarily enrich for mRNA m^{5}C methyltransferases, we tested this strategy with NSUN2, which methylates both tRNA[13,15] and mRNA[16,17]. In brief, cells expressing epitope-tagged NSUN2 were treated with 5-FCyd for 12 h, and covalent RNA–protein complexes

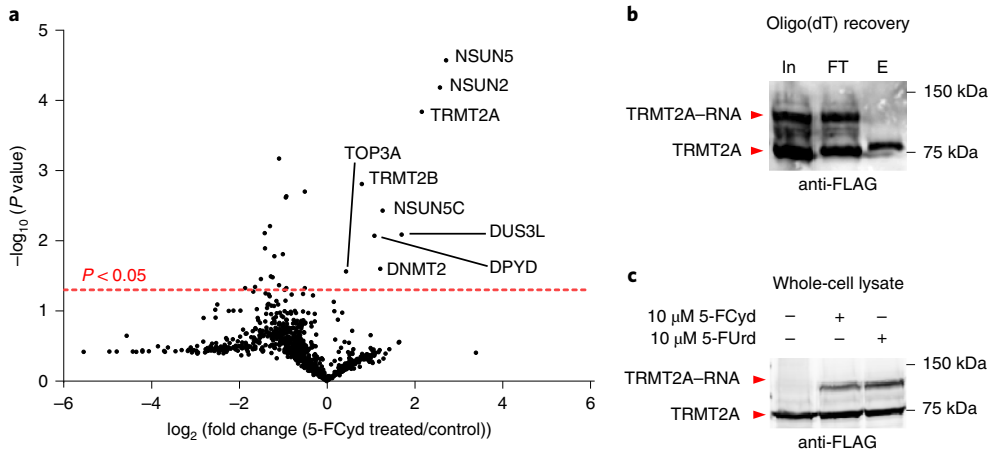

Fig. 2 | Proteomic analysis of 5-FCyd-reactive proteins on mRNA using RNABPP. a, Volcano plot showing enrichment of 5-FCyd-reactive proteins. Three independent biological replicates were analyzed using TMT-based isobaric tags. Multiple Student's *t*-tests (unpaired) were performed to evaluate the statistical significance; $P = 0.000065$ for NSUN2; $P = 0.000027$ for NSUN5; $P = 0.00015$ for TRMT2A; $P = 0.0015$ for TRMT2B; $P = 0.0037$ for NSUN5C; $P = 0.0081$ for DUS3L; $P = 0.0084$ for DPYD; $P = 0.025$ for DNMT2; $P = 0.027$ for TOP3A. **b**, Western blot validation of proteomics result for m^{5}U methyltransferase TRMT2A; In, input; FT, flow through; E, eluate. **c**, Western blot analysis of TRMT2A–RNA crosslinking after metabolic labeling with 10 μM 5-FCyd or 5-FUrd for 12 h. For **b** and **c**, the experiments were repeated three times independently with similar results.

were isolated under denaturing conditions using oligo(dT)-based pulldown. Following bead elution, we characterized poly(A) RNA enrichment by Bioanalyzer analysis (Supplementary Fig. 3) and measured protein enrichment by western blotting after RNA digestion. Gratifyingly, we were able to observe strong enrichment of NSUN2 in the 5-FCyd-treated cells compared to control cells or those treated with 5-AzaC (Fig. 1e), establishing that metabolic labeling with 5-FCyd combined with oligo(dT)-based enrichment can be applied to isolate mRNA-modifying enzymes.

Having validated our approach, we next profiled 5-FCyd-reactive proteins using quantitative MS-based proteomics. We set up a comparative analysis pipeline to measure protein enrichment in 5-FCyd-treated cells compared to untreated cells. Independent samples were labeled with tandem mass tags (TMT)-based isobaric tags[18] and quantified in one LC–MS/MS run (Fig. 2a and Supplementary Dataset 1). In total, we found nine proteins enriched with a *P* value of <0.05; all but one are known or putative pyrimidine-modifying enzymes. Of these, four are m^{5}C methyltransferases, including NSUN2 (refs. [16,17]), which we used to validate the method. We also identified NSUN5, NSUN5C and DNMT2. NSUN5 and NSUN5C are the mammalian homologs of Rcm1, a yeast rRNA m^{5}C methyltransferase[19]. Recently, it was found that NSUN5 methylates analogous residues in human rRNA[20]; its activity on other RNA species is unexplored, although it can bind to G-quadruplex structures in *NRAS* mRNA[21]. DNMT2 homologs have been primarily implicated in m^{5}C formation at C38 on tRNA[22], although recent evidence has implicated human DNMT2 in mRNA modification[23].

The 5-FCyd-treated samples also showed enrichment of several enzymes likely to modify uridine. Because 5-FCyd feeding generates 5-FUrd in RNA, these proteins are likely captured by 5-FUrd. The enzymes tRNA methyltransferase 1 homolog A (TRMT2A) and TRMT2B are the mammalian homologs of yeast TRM2 (ref. [24]) and were recently shown to install m^{5}U at U54 on cytosolic[25] and mitochondrial tRNAs[26], respectively. These proteins use a similar catalytic mechanism as m^{5}C methyltransferases[25] and should crosslink with 5-FUrd-modified RNA. Our data suggest that these proteins may generate m^{5}U on mRNA, consistent with a recent finding by Feng and co-workers[27]. Finally, we were surprised to find DYPD and DUS3L, enzymes from the dihydropyrimidine dehydrogenase family and tRNA-DUS family, respectively, enriched following 5-FCyd treatment. Both of these enzymes should catalyze a similar reaction, reduction of C5–C6 double bond of uracil; but, DYPD[28] is known to modify the free nucleobase, and DUS3L is predicted to modify tRNA by homology to yeast DUS3 (ref. [29]). Notably, DHU has not previously been reported to occur on human mRNA.

Characterization of m^{5}C and m^{5}U methyltransferases. To further explore our findings, we validated interactions of our top hits, NSUN2, NSUN5 and TRMT2A, with mRNA by analyzing crosslinking and protein recovery by western blotting. Gratifyingly, we were able to recover all three proteins using the RNABPP workflow, validating our proteomics results (Fig. 2b and Supplementary Figs. 4 and 5). For TRMT2A, we also observed protein–RNA crosslinking following 5-FUrd feeding (Fig. 2c), and TRMT2A could be recovered using 5-FUrd RNABPP (Supplementary Fig. 6a), suggesting that this enzyme was most likely enriched by metabolic conversion of 5-FCyd to 5-FUrd.

Next, we characterized the contributions of NSUN2, NSUN5 and TRMT2A to cellular m^{5}C and m^{5}U formation. We generated KO cells for each gene using CRISPR/Cas9 technology (Supplementary Table 2 and Supplementary Figs. 7 and 8), extracted RNA and measured m^{5}C and m^{5}U modification levels by nucleoside LC–MS/MS. A rigorous purification protocol using double poly(A) pulldown and small RNA and rRNA depletion was used to prepare the mRNA fraction (Fig. 3a and Supplementary Fig. 9a). We monitored N^6-isopentenyladenosine (i^{6}A) levels, which is characteristic to tRNA[30], to ensure efficient depletion of small RNA (Supplementary Fig. 9b,c). Consistent with previous RNA bisulfite sequencing and m^{5}C MS data that have implicated NSUN2 as the major mRNA m^{5}C-forming enzyme[16,17], we found an 82% reduction in mRNA m^{5}C levels following NSUN2 KO (Fig. 3c, Supplementary Figs. 10 and 11 and Supplementary Tables 3 and 4)[17]. We also found that NSUN2 installs the majority of m^{5}C sites on total RNA (76% reduction following NSUN2 KO) (Fig. 3b), likely reflective of abundant tRNA m^{5}C sites[13,15]. While our LC–MS/MS data together with published sequencing maps[16,17] support the existence of NSUN2-dependent m^{5}C sites on mRNA, interpretation of nucleoside LC–MS data from poly(A)-enriched fractions must be treated carefully. We have used a rigorous poly(A) mRNA purification protocol; however, because m^{5}C is ~40-fold higher in total RNA than in poly(A)-RNA and comparably reduced in each sample following NSUN2 KO (Fig. 3b,c),

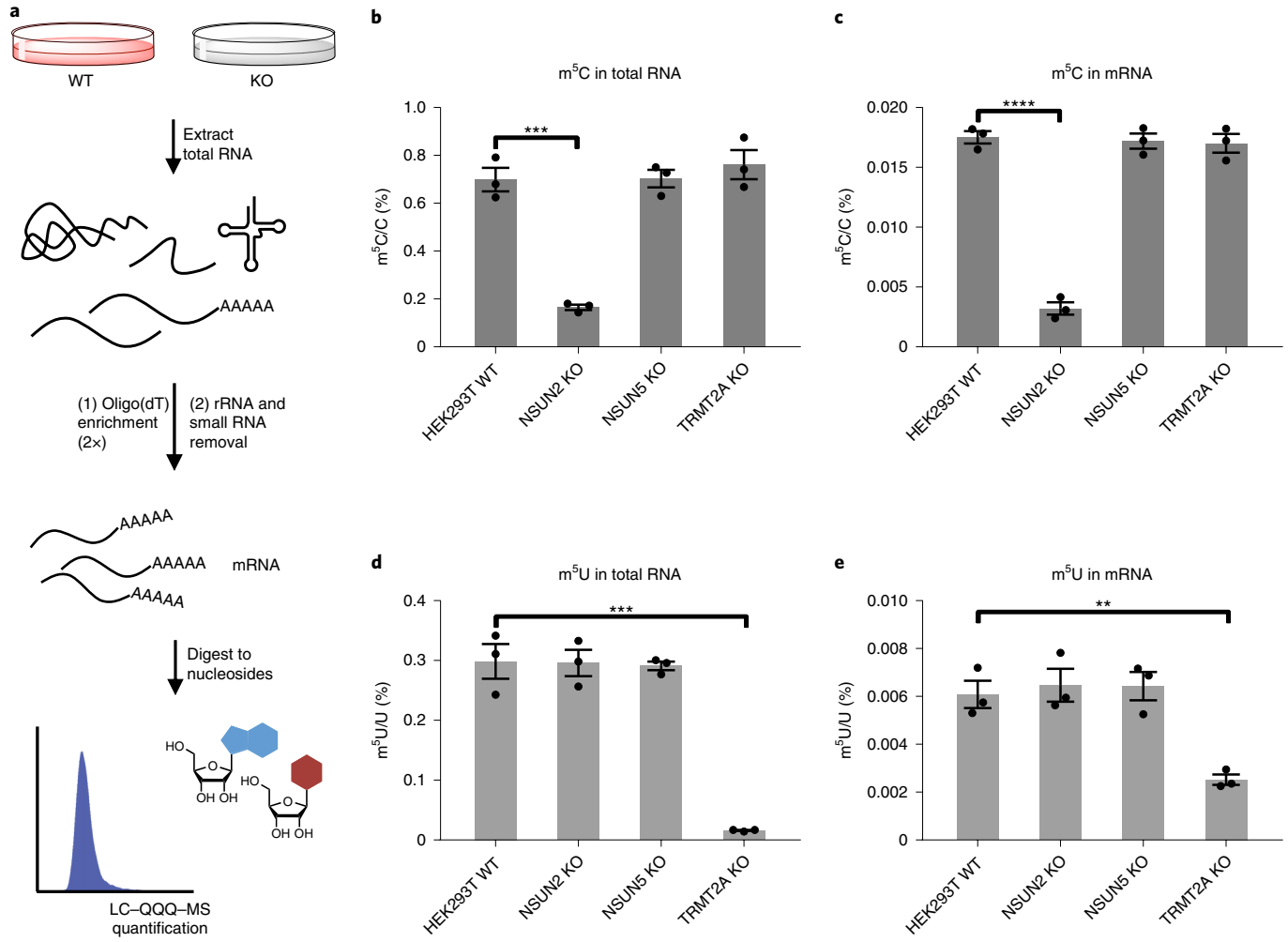

Fig. 3 | NSUN2 and TRMT2A are the major RNA m^{5}C and m^{5}U methyltransferases, respectively. a, Schematic representation of the mRNA isolation protocol for liquid chromatography–triple quadrupole–tandem mass spectrometry (LC–QQQ–MS) quantitation of modified nucleotides in HEK293T wild-type (WT) and KO cells. **b**–**e**, Quantitation of m^{5}C and m^{5}U levels in total RNA (**b** and **d**) and mRNA (**c** and **e**) extracted from WT, NSUN2 KO, NSUN5 KO and TRMT2A KO cell lines. Three independent biological replicates were analyzed. Data represent mean values ± s.e.m. An unpaired *t*-test (two-tailed) was performed to evaluate the statistical significance; $P = 0.00045$ in **b**; $P = 0.000040$ in **c**; $P = 0.00063$ in **c**; $P = 0.0043$ in **d**.

we cannot exclude the possibility that contaminating tRNA/rRNA nucleosides may contribute to the measured m^{5}C content.

We next investigated the contribution of NSUN5 and TRMT2A in m^{5}C formation. We found no statistically significant change in m^{5}C levels following KO of either protein (Fig. 3b,c). This is not unexpected for TRMT2A, which is a known m^{5}U methyltransferase and likely enriched by 5-FUrd. NSUN5 has been shown to generate m^{5}C at a single rRNA site through bisulfite mapping approaches[20], and our data suggest it is unlikely to have many additional substrates. It is also plausible that 5-FCyd-NSUN5 crosslinks occur through kinetic trapping of the enzyme–RNA complex and do not reflect bonafide m^{5}C modifications.

The identification of m^{5}U methyltransferases TRMT2A and TRMT2B in our dataset suggests the presence of m^{5}U on human mRNA. To investigate further, we performed LC–MS/MS analysis of m^{5}U levels in RNA extracted from WT HEK293T and TRMT2A KO cells as described above (Supplementary Tables 3 and 4). In total RNA, we found a dramatic decrease in m^{5}U following TRMT2A knockout with a 95% reduction of the modification (Fig. 3d), consistent with published data[25]. Further, we detected m^{5}U in the enriched mRNA fraction, albeit at lower levels than in total RNA, and observed a 59% reduction of m^{5}U following TRMT2A KO (Fig. 3e). Our results suggest that TRMT2A may install m^{5}U on cellular mRNA, consistent with a recent study[27]. In addition, residual m^{5}U in the mRNA-enriched fraction following TRMT2A ablation suggests the existence of other m^{5}U methyltransferases. We propose that TRMT2B is a likely candidate, given its enrichment in our proteomic dataset and evidence that it modifies tRNA and rRNA[26]. At the current time, however, given the lack of sequencing data supporting the existence of m^{5}U on mRNA and the low levels of this modification measured in poly(A)-enriched RNA fractions compared to high levels in non-coding RNAs, nucleoside LC–MS data alone should be treated cautiously[31].

5-Halopyrimidines crosslink DUS3L. We next turned to DUS3L, the mammalian homolog of the yeast DUS, DUS3 (ref. [29]). Phizicky and co-workers previously demonstrated that yeast DUS3 installs DHU at U47 in tRNA^{Tyr}[29]; however, its mammalian homolog, DUS3L, has been uncharacterized, and we lack transcriptome-wide maps of DHU sites in any organism. Further, to our knowledge, DHU has been found exclusively in tRNA and rRNA, and there is no precedent for its occurrence in mRNA. To probe our finding further, we analyzed crosslinking and RNABPP protein enrichment with an anti-DUS3L antibody. In lysate generated from 5-FCyd-treated

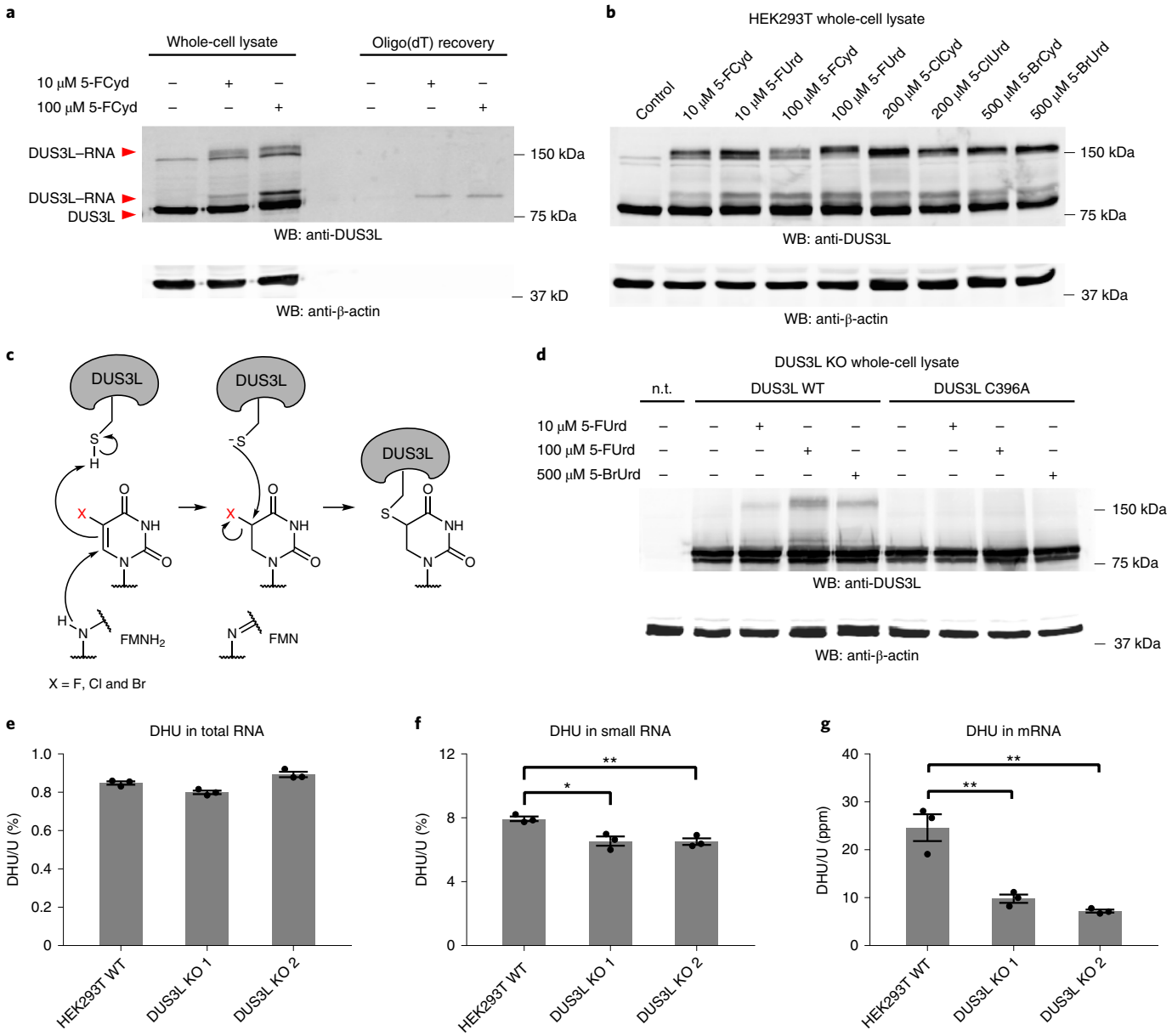

Fig. 4 | DUS3L installs DHU on human RNA. a, Western blot analysis of 5-FCyd-mediated enrichment for DUS3L. After metabolic labeling with 5-FCyd, oligo(dT)-enriched samples were analyzed by western blotting using an anti-DUS3L antibody. **b**, Western blot analysis of DUS3L–RNA crosslinking after metabolic labeling with a panel of C5-halogenated pyrimidines; 5-ClCyd, 5-chlorocytidine; 5-ClUrd, 5-chlorouridine; 5-BrCyd, 5-bromocytidine; 5-BrUrd, 5-bromouridine. **c**, Proposed mechanism of DUS3L–RNA crosslinking mediated by C5-halogenated uridine derivatives. **d**, Western blot analysis of protein–RNA crosslinking between WT DUS3L or C396A DUS3L and 5-FUrd- or 5-BrUrd-labeled RNA. WT or mutant DUS3L transgenes were introduced by transfection of plasmid constructs into DUS3L-KO cells. For **a**, **b** and **d**, the experiments were repeated three times independently with similar results; n.t., no transfection control. **e**–**g**, Quantitation of DHU levels in total RNA (**e**), small RNA (**f**) and mRNA (**g**) extracted from WT and DUS3L-KO cells. Three independent biological replicates were analyzed. Data represent mean values ± s.e.m. An unpaired *t*-test (two-tailed) was performed to evaluate the statistical significance; $^{*}P = 0.012$ and $^{**}P = 0.0046$ in **f**; $^{**}P = 0.0072$ (1) and $P = 0.0035$ (2) in **g**.

cells, we observed slower migrating species (Fig. 4a), consistent with covalent RNA–protein adducts. Further, in 5-FCyd-labeled samples, but not in the control samples, we were able to detect DUS3L after oligo(dT) pulldown (Fig. 4a and Supplementary Fig. 12). Taken together, these results demonstrate 5-FCyd-dependent RNA crosslinking of DUS3L and support its interaction with polyadenylated RNA.

Next, we investigated the RNA–protein crosslinking mechanism. Because DUS enzymes modify uridine, we tested whether metabolic labeling with 5-FUrd would result in crosslinking and protein enrichment. Our results demonstrate similar amounts of crosslinking (Fig. 4b) and protein recovery (Supplementary Fig. 6b) using either 5-FCyd or 5-FUrd, which are suggestive of 5-FUrd as the major crosslinking entity. We also evaluated crosslinking with modified pyrimidine analogs containing chloro (5-ClCyd and 5-ClUrd) or bromo (5-BrCyd and 5-BrUrd) groups at the C5 position. In all cases, we observed efficient DUS3L crosslinking (Fig. 4b). How do 5-halopyrimidines facilitate crosslinking of DUS3L? Based on structural and biochemical data for DUS enzymes[32,33], we propose that crosslinking begins with enzymatic

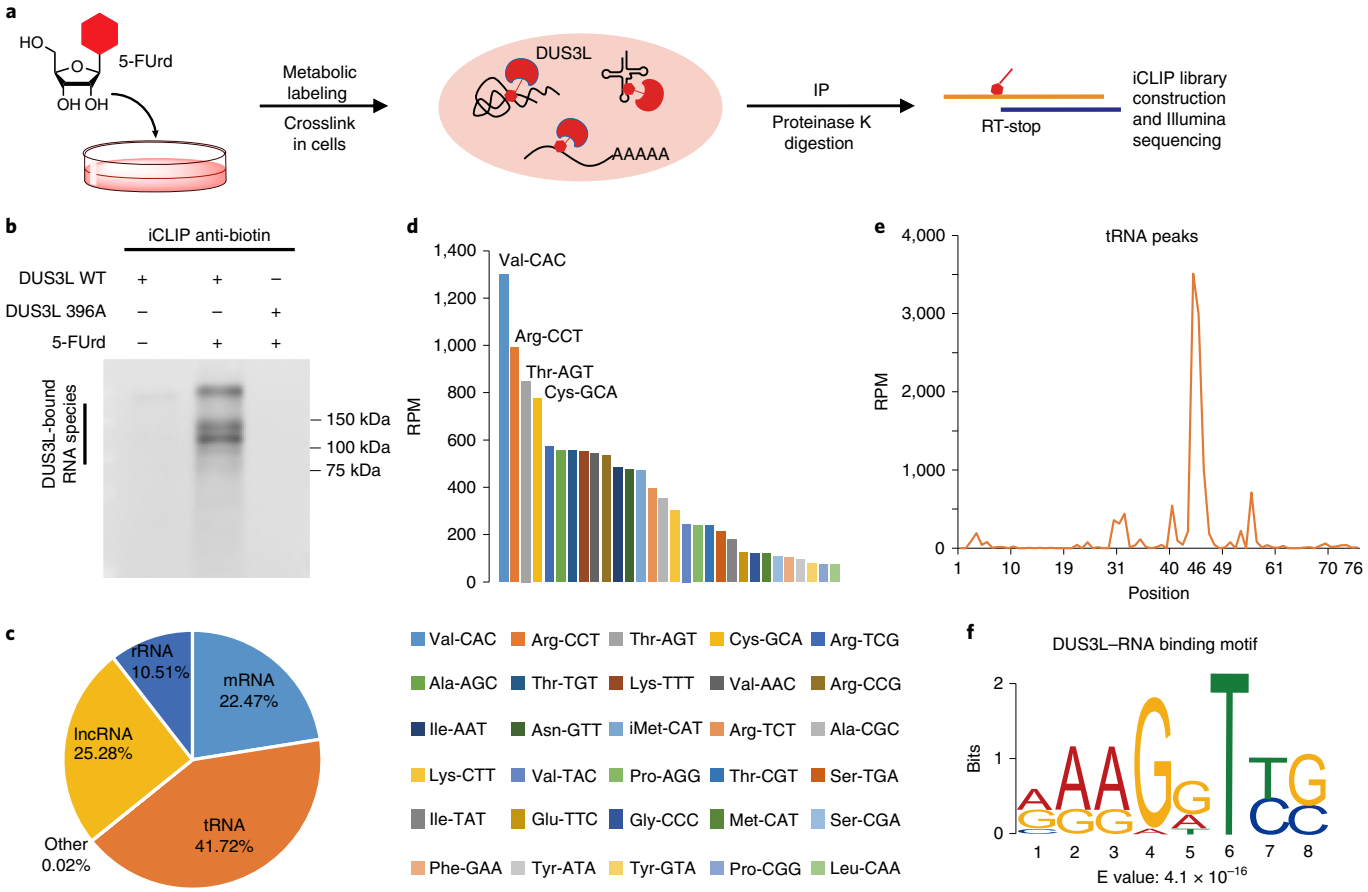

Fig. 5 | 5-FUrd-iCLIP sequencing of DUS3L substrates. a, Schematic of 5-FUrd-iCLIP workflow; IP, immunoprecipitation; RT, reverse transcription. **b**, Analysis of RNA–protein crosslinking in DUS3L iCLIP samples. Cells expressing DUS3L or DUS3L C396A were treated with 5-FUrd or left untreated, and covalently linked RNA was detected by anti-biotin western blotting after immunoprecipitation and RNase fragmentation. For the data presented in **b**, the experiment was repeated three times independently with similar results. **c**, Composition of RNA identified by 5-FUrd-iCLIP according to uniquely mapped reads; lncRNA, long non-coding RNA. **d**, Abundance of tRNA species as enriched by 5-FUrd-iCLIP; RPM, reads per million mapped reads; iMet, initiator methionine. **e**, Coverage of all tRNA peaks according to their relative position within the mature tRNA. **f**, Consensus motif detected by MEME using all RNA peaks identified by 5-FUrd-iCLIP.

reduction of 5-halouridine to 5-halodihydrouridine followed by nucleophilic attack of a conserved catalytic cysteine (C396 in DUS3L) on the C5 position with halide serving as a leaving group (Fig. 4c). To directly test this hypothesis, we generated a stable DUS3L-KO cell line (Supplementary Figs. 7 and 8) using CRISPR/Cas technology and transfected plasmids encoding either WT DUS3L or DUS3L containing the C396A mutation. Following feeding with 5-FUrd or 5-BrUrd, we were able to observe crosslinking to the WT transgene, but not to the C396A mutant, suggesting that the catalytic cysteine residue is important for crosslinking. (Fig. 4d)

LC–MS quantification of DHU levels in human RNA. We next studied the role of DUS3L in cellular DHU formation. We isolated RNA from WT and two independent DUS3L-KO cell lines generated using CRISPR/Cas technology and measured DHU levels in total RNA, mRNA and small RNA using LC–MS/MS (Supplementary Figs. 13, 14 and 15 and Supplementary Tables 5, 6 and 7). In total RNA, we found an appreciable level of DHU, corresponding to 0.8% of all uridine residues. Interestingly, DHU levels in total RNA of both DUS3L-KO strains were largely unchanged relative to WT, indicating that DUS3L catalyzes a minor fraction of DHU modification on total RNA (Fig. 4e); it is likely that the majority of DHU is installed by the other mammalian DUS homologs (that is, DUS1L, DUS2L and DUS4L). In the small RNA fraction (17 to 300 nucleotides (nt)), which contains tRNA, we observed a dramatically higher level of DHU, corresponding to ~8% of total uridine residues (Fig. 4f), and an 18% decrease in DHU levels in the DUS3L-KO strains (Fig. 4f), establishing a role for DUS3L in DHU formation most likely on mammalian tRNA. This would be analogous to yeast DUS3, which modifies U47 on yeast $tRNA^{Tyr}$ [29]. Further, we isolated mRNA from WT and DUS3L-KO strains using the previously described approach and analyzed DHU levels. We detected DHU in the mRNA-enriched fraction, although levels were far lower than in total RNA and small RNA. Interestingly, despite a modest decrease of DHU in total RNA or small RNA following DUS3L KO, we observed a 60–70% decrease in DHU levels in poly(A)-enriched RNA for both DUS3L-KO strains. While these data support the existence of DHU on human mRNA (Fig. 4g), given the extremely low levels that we measure and the possibility of contaminating DHU from abundant non-coding RNA species, further evidence is needed to unambiguously establish the presence of DHU on human mRNA.

Mapping DUS3L-dependent DHU sites. Because the biological substrates of DUS3L are unknown, we adapted the iCLIP method[11] to map DUS3L-dependent DHU sites across the transcriptome using metabolic labeling with 5-FUrd to crosslink substrate RNAs (Fig. 5a). Analogous approaches have been applied to map m^5C and

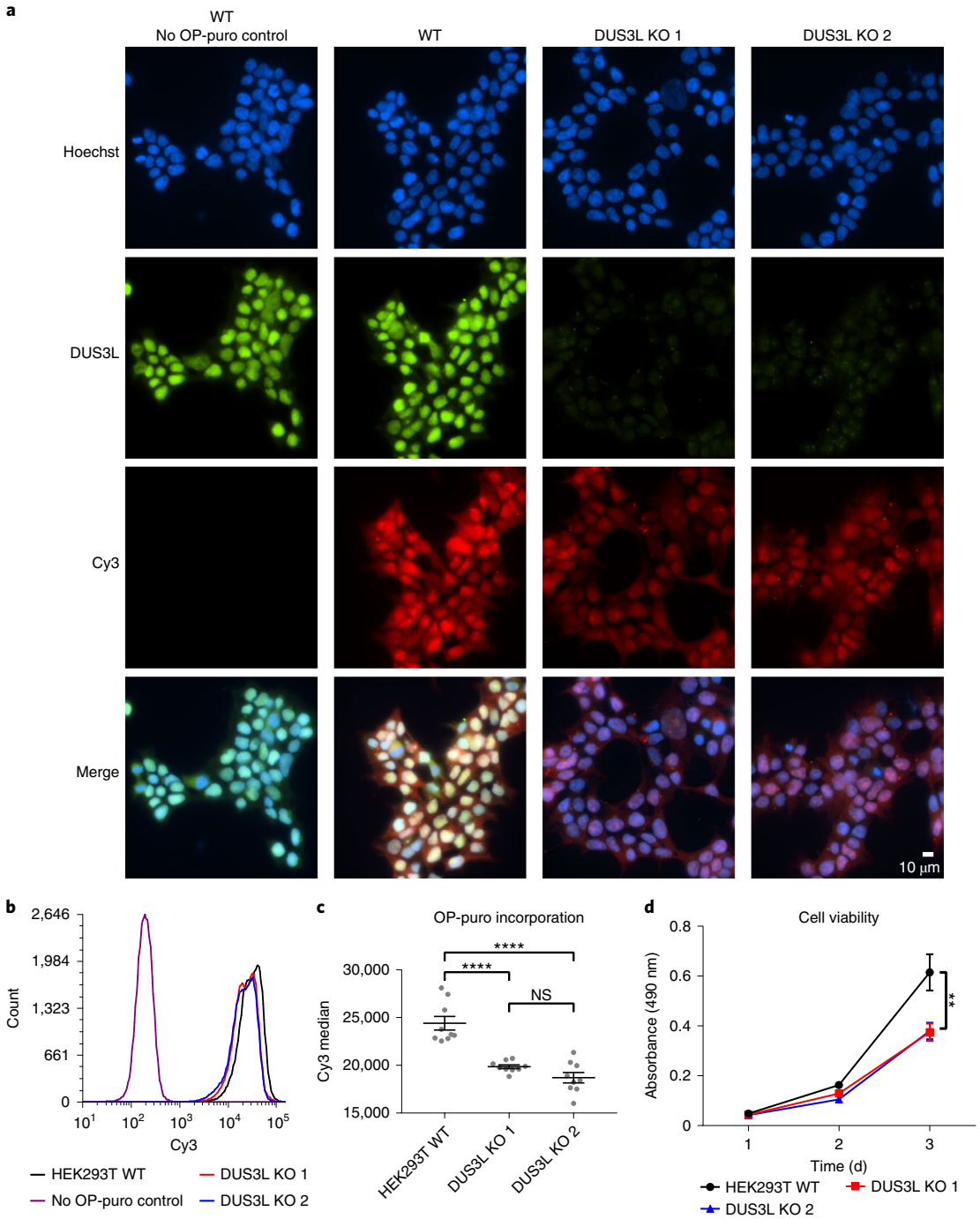

Fig. 6 | DUS3L regulates cell proliferation and protein translation efficiency. a, Fluorescence microscopy analysis of DUS3L localization and OP-puro incorporation by WT and DUS3L-KO cells. OP-puro labeling was imaged following CuAAC reaction with Cy3-azide. Two independent biological replicates were analyzed for each cell line with similar results. **b**, Flow cytometry analysis of global protein translation using OP-puro in WT and DUS3L-KO cells. **c**, Plot of OP-puro incorporation as measured by Cy3 fluorescence from **b**. The median fluorescence intensity and s.e.m. values from three independent biological replicates (50,000 cells analyzed per sample, three technical replicates per sample) are shown. An unpaired *t*-test (two-tailed) was performed to evaluate the statistical significance; ****$P = 0.00013$ (1) and $P = 0.0000094$ (2); NS, $P = 0.064$. **d**, Cell viability of WT and DUS3L-KO cells measured by MTS assay. Twelve independent biological replicates were analyzed. Data represent mean values ± s.e.m. An unpaired *t*-test (two-tailed) was performed to evaluate the statistical significance; **$P = 0.0075$ (1) and $P = 0.0068$ (2).

m^5U methyltransferase substrates using metabolic labeling with modified pyrimidines[13,25] or methyltransferase variants that cannot release RNA substrate[15]. In brief, we fed cells with 5-FUrd and immunoprecipitated DUS3L together with covalently linked RNA. Analysis of DUS3L-associated RNA demonstrated clear enrichment over cells not treated with 5-FUrd or cells expressing C396A DUS3L (Fig. 5b and Supplementary Fig. 16a), indicating the specificity of RNA–protein crosslinking. Next, we generated iCLIP libraries from 5-FUrd-treated cells following literature precedent (Supplementary Fig. 16b,c)[34]. Control RNA libraries were prepared from untreated

cells expressing WT DUS3L. Based on the number of PCR cycles required to generate detectable PCR products, we estimate ~tenfold more RNA in the 5-FUrd-treated samples than in the control samples. RNA libraries were subjected to Illumina sequencing and deduplicated, and uniquely aligned reads were used to identify crosslink sites based on reverse transcription-stop signatures. Crosslink sites were further aggregated into crosslink peaks using Paraclu[35], and only peaks present in multiple independent biological replicates that showed at least a fourfold enrichment over the control peaks were used for further analysis (Supplementary Dataset 2).

Analysis of 5-FUrd-treated DUS3L samples showed that the largest fraction of unique reads mapped to cytoplasmic tRNA genes (Fig. 5c). By contrast, the majority of reads in control samples originated from rRNA, likely reflecting non-specific binding (Supplementary Fig. 16d). In total, we identified 180 tRNA peaks corresponding to 45 different tRNA isoacceptors (Fig. 5d and Supplementary Dataset 2). Further, alignment of all tRNA peaks according to their relative position within the mature tRNA transcript demonstrated strong enrichment at positions 45–47 (Fig. 5e and Supplementary Dataset 2), with 28 of the 45 tRNA isoacceptors containing peaks between positions 44 and 49 (Supplementary Dataset 2). Because the nearest U residue occurs at position 46–48 in most tRNAs, we conclude that DUS3L reacts with 5-FUrd at residues U46–48, and reverse transcription proceeds on average one additional nucleotide into the DUS3L–RNA crosslinked adduct. In addition, we observed minor peaks near position 58, likely due to abundant m^1A modification at this position, as well as 5′ to the U46–48 site, which may result from partial read-through of the DUS3L-5-FUrd adduct (Fig. 5e and Supplementary Dataset 2). Multiple expectation maximizations for motif elicitation (MEME)[36] analysis of the crosslink peaks identified a consensus motif that matches the variable loop sequence of many tRNA isoacceptors (Fig. 5f). Taken together, our iCLIP data show that U46–48 in the tRNA variable loop is a major dihydrouridylation site for human DUS3L. Phizicky and co-workers[29] reached a similar conclusion in their studies of yeast DUS3 with individual tRNAs, indicating that the tRNA-modifying function of this enzyme is conserved in eukaryotic evolution.

Because we identified DUS3L through RNABPP oligo(dT) pulldown and measured DHU in the poly(A) RNA fraction, we also investigated DUS3L peaks mapping to non-tRNA genes. Among different classes of RNA identified in the 5-FUrd-treated samples, we found an approximately twofold increase in the fraction of uniquely mapped reads from mRNA and lncRNA compared to the control samples (Fig. 5c and Supplementary Fig. 16d) and identified 830 total non-tRNA crosslink peaks (Supplementary Dataset 2). Most of these peaks mapped to introns and non-coding RNA and, on average, showed ~tenfold lower normalized read counts than DUS3L tRNA peaks (Supplementary Dataset 2). Among non-tRNA peaks, we found 80 peaks in mature mRNA transcripts distributed evenly across 5′- and 3′-untranslated regions (UTRs) and coding sequences (Supplementary Dataset 2 and Supplementary Fig. 16e). Another 44 peaks were located in retained introns, which can be incorporated into mature transcripts. MEME analysis of mRNA peaks yielded the consensus motif GGGTCC (Supplementary Fig. 16f), similar to the tRNA DHU site at residues 46–48. While these peaks do suggest DHU modification in mRNA, given the propensity of reverse transcription-stop-based analyses for generating false positives[31] and the relatively low number and abundance of these peaks, orthogonal validation will be required to firmly establish DUS3L-mediated DHU modifications outside of the tRNA variable loop.

DUS3L regulates protein translation and cell viability. Having mapped the cellular substrates of DUS3L, we next studied the effect of DUS3L KO on protein translation and cell viability. To measure the protein translation rate, we fed WT and DUS3L-KO cells *O*-propargyl-puromycin (OP-puro)[37] and measured its accumulation into cellular proteins using fluorescence microscopy and flow cytometry. In DUS3L-KO strains, we observed a 18–23% reduction in OP-puro labeling (Fig. 6a–c and Supplementary Table 8), indicating that protein translation was impaired in these cell lines and implicating DUS3L in this process. We also found that DUS3L exhibited primarily nuclear localization (Fig. 6a). To quantify cell growth and metabolism, we utilized an MTS-based assay and observed a 40% reduction in viability in both DUS3L-KO strains after 3 d (Fig. 6d and Supplementary Table 9). Given that we have identified widespread DUS3L-mediated dihydrouridylation in the tRNA variable loop, it is likely that the observed effects on protein translation and cell proliferation in DUS3L-KO cells are the result of global hypomodification of tRNAs. Kato et al.[38] previously demonstrated that hDUS2 knockdown impaired the growth of lung cancer cells, reinforcing the importance of DHU modification and DUS enzymes in fundamental cellular processes.

Discussion

In this manuscript, we develop an activity-based profiling strategy to study RNA-modifying enzymes in living cells. While ABPP has been applied to diverse classes of cellular enzymes[7], its application to RNA-modifying enzymes, an emerging class of disease-relevant pharmacologic targets, has been limited. Here, by applying metabolic labeling with 5-FCyd, we profile three distinct classes of pyrimidine-modifying enzymes: m^5C RNA methyltransferases, m^5U RNA methyltransferases and DUSs. The versatility of 5-FCyd is made possible by its partial conversion to 5-FUrd in cells as well as diverse reactivity modes, including latent electrophilicity and inhibition of β-elimination during enzyme–substrate release.

The presence of mRNA m^5C and the identity of its writers has been controversial. While tRNA m^5C sites have been well established[13,15,22,39], RNA bisulfite sequencing of mRNA has led to inconsistent m^5C maps[40], and CLIP-based studies with NSUN2 have primarily identified tRNA modification sites[13,15]. Here we provide two complementary data points supporting the existence of m^5C on mRNA and implicating NSUN2 as its writer. First, we recover NSUN2 using oligo(dT)-based pulldown with 5-FCyd. Second, LC–MS/MS analysis of poly(A)-enriched RNA demonstrates the NSUN2-dependent presence of m^5C at levels of ~1 in 5,000 cytidine bases. LC–MS analysis of bulk nucleoside digests has known limitations in quantifying modification levels[31], particularly when modifications are more abundant in tRNA/rRNA than in mRNA (as is the case for m^5C). In addition, oligo(dT)-based pulldown approaches can recover non-mRNA-associated proteins. Therefore, while neither of these measurements alone is conclusive, they are consistent with two recent bisulfite sequencing studies that show that NSUN2 is the major mRNA m^5C writer[16,17].

RNABPP with both 5-FCyd and 5-FUrd identified the m^5U methyltransferase TRMT2A. In addition, we were able to detect m^5U in bulk nucleoside digests from poly(A)-enriched RNA, although levels were ~tenfold lower than m^5C. Recent work using CLIP-based sequencing has shown that TRMT2A primarily installs m^5U U54 on tRNA but did not identify TRMT2A-dependent m^5U sites on mRNA[25], although a small number of mRNAs were found to associate with the protein. Similar to our study, Feng and co-workers measured m^5U in bulk nucleoside digests from poly(A)-enriched RNA by MS. However, they did not observe a substantial reduction in m^5U levels following short interfering RNA (siRNA)-mediated knockdown of TRMT2A[27] (perhaps due to inefficient knockdown). Given these contradictory findings, and the lack of sequencing strategies for profiling m^5U sites across the transcriptome, at the present time, we feel that the existence of m^5U on mRNA awaits further confirmation.

DHU is an abundant RNA modification found in all kingdoms of life[41–44]; yet, our understanding of this modified base in mammals is lacking. Our work provides a transcriptome-wide

characterization of the substrates of a mammalian DUS. Taken together, LC–MS analysis of fractionated RNA in WT and DUS3L-KO cells and sequencing of 5-FUrd-induced DUS3L–RNA crosslinks indicate that U46–48 in the tRNA variable loop is the major dihydrouridylation substrate of DUS3L. Our results align with studies by Phizicky and co-workers on yeast DUS3 (ref. [29]), which used microarray analysis to implicate the enzyme in DHU formation at U47 on a small number of tRNAs, including $tRNA^{Tyr}$, $tRNA^{Val}$, $tRNA^{Cys}$, $tRNA^{Ile}$, $tRNA^{Trp}$, $tRNA^{Thr}$ and $tRNA^{Asn}$; modification of $tRNA^{Tyr}$ was further confirmed by primer extension analysis. Our analysis of human DUS3L identified U46–48 sites on 28 tRNA isoacceptors, including all of the tRNAs identified by the Phizicky study with the exception of $tRNA^{Trp}$. We did not identify any major DUS3L peaks in the DHU-rich 'D-arm', which indicates that other mammalian DUS enzymes likely install DHU at these sites. Interestingly, while we observed reduced cell proliferation and protein translation in human DUS3L-KO cells, DUS3 deletion alone in budding yeast does not cause overt growth defects[29], except when it is combined with the deletion of other tRNA-modifying enzymes[45]. This may be due to an increased reliance on dihydrouridylation in mammals for regulating tRNA structure, context-specific roles in the evaluated cell lines and potential differences in the ability of yeast and mammalian enzyme to modify non-tRNA substrates. Notably, our 5-FUrd-iCLIP data show DUS3L crosslinking peaks on non-tRNA substrates, including a small number found in mature mRNA. While these peaks suggest potential DUS3L-catalyzed DHU modification sites on mammalian mRNA, given the relatively low abundance of reads at DUS3L mRNA peaks compared to DUS3L tRNA peaks and the indirect nature of reverse transcription-stop-based analyses[31], orthogonal validation will be required to establish the existence of DHU on mammalian mRNA. Further investigation of DHU sites across the transcriptome at single-nucleotide resolution will likely require the development of DHU-specific sequencing technology, analogous to approaches that have been used for other epitranscriptomic bases[5], and will be critical to further understand the role that this modification and associated proteins play in biological processes.

Our study demonstrates an unexpected reactivity of DUS enzymes toward electrophilic nucleotides. Because most DUS enzymes possess a catalytic cysteine residue, metabolically incorporated 5-halopyrimidines may serve as a general class of activity-based probes for the entire family, providing a chemical strategy for profiling DUS activity and substrate scope in vivo. In addition, cysteine reactivity found in DUS enzymes[46] may be a starting point for the development of small-molecule tools for this enzyme class[47,48]. We also envision that the modified nucleosides used in this work can be incorporated into synthetic oligonucleotides to develop probes for studying and modulating RNA-modifying enzyme activity in vitro and in vivo. Finally, with appropriate nucleoside selection, we propose that RNABPP can be applied broadly in a targeted (mechanism-based) or non-targeted (reactivity-based) manner to profile diverse RNA-associated proteins in different biological contexts and at amino acid resolution[49]. Electrophilic small-molecule ABPP probes have seen widespread use[47], and placement of these functional groups at the C5 position of pyrimidines should enable their integration into the RNABPP method. Such studies are underway in our group and will provide new insights into transcriptomic and epitranscriptomic processes in biology.

Online content

Any methods, additional references, Nature Research 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/s41589-021-00874-8.

Received: 1 September 2020; Accepted: 2 August 2021;

References

1. Boccaletto, P. et al. MODOMICS: a database of RNA modification pathways. 2017 update. *Nucleic Acids Res.* **46**, D303–D307 (2018).
2. Nachtergaele, S. & He, C. The emerging biology of RNA post-transcriptional modifications. *RNA Biol.* **14**, 156–163 (2017).
3. Roundtree, I. A., Evans, M. E., Pan, T. & He, C. Dynamic RNA modifications in gene expression regulation. *Cell* **169**, 1187–1200 (2017).
4. Delaunay, S. & Frye, M. RNA modifications regulating cell fate in cancer. *Nat. Cell Biol.* **21**, 552–559 (2019).
5. Li, X., Xiong, X. & Yi, C. Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat. Methods* **14**, 23–31 (2016).
6. Zaccara, S., Ries, R. J. & Jaffrey, S. R. Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 608–624 (2019).
7. Cravatt, B. F., Wright, A. T. & Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **77**, 383–414 (2008).
8. Castello, A. et al. Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**, 1393–1406 (2012).
9. Baltz, A. G. et al. The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol. Cell* **46**, 674–690 (2012).
10. Pappireddi, N., Martin, L. & Wuhr, M. A review on quantitative multiplexed proteomics. *ChemBioChem* **20**, 1210–1224 (2019).
11. Konig, J. et al. iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat. Struct. Mol. Biol.* **17**, 909–915 (2010).
12. Liu, Y. & Santi, D. V. m^5C RNA and m^5C DNA methyl transferases use different cysteine residues as catalysts. *Proc. Natl Acad. Sci. USA* **97**, 8263–8265 (2000).
13. Khoddami, V. & Cairns, B. R. Identification of direct targets and modified bases of RNA cytosine methyltransferases. *Nat. Biotechnol.* **31**, 458–464 (2013).
14. Lu, L. J., Tseng, W. C. & Randerath, K. Effects of 5-fluorocytidine on mammalian transfer RNA and transfer RNA methyltransferases. *Biochem. Pharmacol.* **28**, 489–495 (1979).
15. Hussain, S. et al. NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. *Cell Rep.* **4**, 255–261 (2013).
16. Huang, T., Chen, W., Liu, J., Gu, N. & Zhang, R. Genome-wide identification of mRNA 5-methylcytosine in mammals. *Nat. Struct. Mol. Biol.* **26**, 380–388 (2019).
17. Yang, X. et al. 5-Methylcytosine promotes mRNA export—NSUN2 as the methyltransferase and ALYREF as an m^5C reader. *Cell Res.* **27**, 606–625 (2017).
18. Thompson, A. et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **75**, 1895–1904 (2003).
19. Schosserer, M. et al. Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. *Nat. Commun.* **6**, 6158 (2015).
20. Heissenberger, C. et al. Loss of the ribosomal RNA methyltransferase NSUN5 impairs global protein synthesis and normal growth. *Nucleic Acids Res.* **47**, 11807–11825 (2019).
21. Herdy, B. et al. Analysis of NRAS RNA G-quadruplex binding proteins reveals DDX3X as a novel interactor of cellular G-quadruplex containing transcripts. *Nucleic Acids Res.* **46**, 11592–11604 (2018).
22. Tuorto, F. et al. RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. *Nat. Struct. Mol. Biol.* **19**, 900–905 (2012).
23. Xue, S. et al. Depletion of TRDMT1 affects 5-methylcytosine modification of mRNA and inhibits HEK293 cell proliferation and migration. *Biochem. Biophys. Res. Commun.* **520**, 60–66 (2019).
24. Nordlund, M. E., Johansson, J. O., von Pawel-Rammingen, U. & Bystrom, A. S. Identification of the *TRM2* gene encoding the $tRNA(m^5U_{54})$ methyltransferase of *Saccharomyces cerevisiae*. *RNA* **6**, 844–860 (2000).
25. Carter, J. M. et al. FICC-Seq: a method for enzyme-specified profiling of methyl-5-uridine in cellular RNA. *Nucleic Acids Res.* **47**, e113 (2019).
26. Powell, C. A. & Minczuk, M. TRMT2B is responsible for both tRNA and rRNA m5U-methylation in human mitochondria. *RNA Biol.* **17**, 451–462 (2020).
27. Cheng, Q. Y. et al. Chemical tagging for sensitive determination of uridine modifications in RNA. *Chem. Sci.* **11**, 1878–1891 (2020).
28. Lu, Z. H., Zhang, R. & Diasio, R. B. Purification and characterization of dihydropyrimidine dehydrogenase from human liver. *J. Biol. Chem.* **267**, 17102–17109 (1992).
29. Xing, F., Hiley, S. L., Hughes, T. R. & Phizicky, E. M. The specificities of four yeast dihydrouridine synthases for cytoplasmic tRNAs. *J. Biol. Chem.* **279**, 17850–17860 (2004).

30. Schweizer, U., Bohleber, S. & Fradejas-Villar, N. The modified base isopentenyladenosine and its derivatives in tRNA. *RNA Biol.* **14**, 1197–1208 (2017).
31. Wiener, D. & Schwartz, S. The epitranscriptome beyond m^{6}A. *Nat. Rev. Genet.* **22**, 119–131 (2021).
32. Yu, F. et al. Molecular basis of dihydrouridine formation on tRNA. *Proc. Natl Acad. Sci. USA* **108**, 19593–19598 (2011).
33. Rider, L. W., Ottosen, M. B., Gattis, S. G. & Palfey, B. A. Mechanism of dihydrouridine synthase 2 from yeast and the importance of modifications for efficient tRNA reduction. *J. Biol. Chem.* **284**, 10324–10333 (2009).
34. Huppertz, I. et al. iCLIP: protein–RNA interactions at nucleotide resolution. *Methods* **65**, 274–287 (2014).
35. Frith, M. C. et al. A code for transcription initiation in mammalian genomes. *Genome Res.* **18**, 1–12 (2008).
36. Bailey, T. L. et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).
37. Liu, J., Xu, Y., Stoleru, D. & Salic, A. Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. *Proc. Natl Acad. Sci. USA* **109**, 413–418 (2012).
38. Kato, T. et al. A novel human tRNA-dihydrouridine synthase involved in pulmonary carcinogenesis. *Cancer Res.* **65**, 5638–5646 (2005).
39. Schaefer, M. et al. RNA methylation by Dnmt2 protects transfer RNAs against stress-induced cleavage. *Genes Dev.* **24**, 1590–1595 (2010).
40. Bohnsack, K. E., Hobartner, C. & Bohnsack, M. T. Eukaryotic 5-methylcytosine (m^{5}C) RNA methyltransferases: mechanisms, cellular functions, and links to disease. *Genes* **10**, 102 (2019).
41. Kuchino, Y. & Borek, E. Tumour-specific phenylalanine tRNA contains two supernumerary methylated bases. *Nature* **271**, 126–129 (1978).
42. Madison, J. T. & Holley, R. W. The presence of 5,6-dihydrouridylic acid in yeast 'soluble' ribonucleic acid. *Biochem. Biophys. Res. Commun.* **18**, 153–157 (1965).
43. Xing, F., Martzen, M. R. & Phizicky, E. M. A conserved family of *Saccharomyces cerevisiae* synthases effects dihydrouridine modification of tRNA. *RNA* **8**, 370–381 (2002).
44. Bishop, A. C., Xu, J., Johnson, R. C., Schimmel, P. & de Crecy-Lagard, V. Identification of the tRNA-dihydrouridine synthase family. *J. Biol. Chem.* **277**, 25090–25095 (2002).
45. Alexandrov, A. et al. Rapid tRNA decay can result from lack of nonessential modifications. *Mol. Cell* **21**, 87–96 (2006).
46. Lanning, B. R. et al. A road map to evaluate the proteome-wide selectivity of covalent kinase inhibitors. *Nat. Chem. Biol.* **10**, 760–767 (2014).
47. Backus, K. M. et al. Proteome-wide covalent ligand discovery in native biological systems. *Nature* **534**, 570–574 (2016).
48. Erlanson, D. A. et al. Site-directed ligand discovery. *Proc. Natl Acad. Sci. USA* **97**, 9367–9372 (2000).
49. Trendel, J. et al. The human RNA-binding proteome and its dynamics during translational arrest. *Cell* **176**, 391–403 (2019).

Methods

Chemicals. 5-AzaC, 5-FCyd, 5-FUrd, 5-ClCyd, 5-ClUrd, 5-BrCyd and 5-BrUrd were all purchased from Carbosynth. All other chemicals were purchased from Sigma-Aldrich or Fisher Scientific unless otherwise indicated.

Plasmids. NSUN2, NSUN5, NSUN6, TRMT2A, DNMT2 and DUS3L cDNA were obtained from Dharmacon (MHS6278–213243840, MHS6278–202826182, MHS6278–202802090, MHS6278–202760193, MHS6278–202756140 and MHS6278–202830580). The C396A mutation was introduced into DUS3L using overlap extension PCR with mutagenic primers. For transient transfection and construction of Flp-In cell lines, NSUN2, NSUN5, NSUN6, TRMT2A, DNMT2 and DUS3L cDNA were cloned into a modified pcDNA5/FRT/TO (Life Technologies, V6520–20) vector containing an N-terminal 3×FLAG tag. For generation of the KO cell lines, DNA oligos containing guide RNA (gRNA) sequences were cloned into pX330-U6-Chimeric_BB-CBh-hSpCas9 (Addgene, 42230). The DNA oligos were phosphorylated with T4 PNK (NEB, M0201), annealed and ligated into a BbsI-digested (NEB, R3539) px330 backbone with T4 DNA ligase (NEB, M0202).

General cell culture. HEK293T WT and KO cells were cultured at 37 °C in a humidified atmosphere with 5% CO_2 in DMEM (Thermo Fisher, 11995073) supplemented with 10% fetal bovine serum (Bio-Techne, S12450H), 1× penicillin–streptomycin (Thermo Fisher, 15070–063) and 2 mM L-glutamine (Thermo Fisher, 25030–081).

Generation of stable cell lines. To generate stable cell lines expressing 3×FLAG-tagged NSUN2, NSUN5 and TRMT2A, Flp-In T-Rex 293 cells were seeded at 0.6 million cells per well in six-well plates and cotransfected with pOG44 (2 µg; Thermo Fisher, V600520) and pcDNA5/FRT/TO plasmid containing 3×FLAG-NSUN2, 3×FLAG-NSUN5 or 3×FLAG-TRMT2A (0.2 µg). Following selection in 100 µg ml^{-1} hygromycin B and 15 µg ml^{-1} blasticidin, colonies were expanded. To test the expression efficiency of 3×FLAG-tagged constructs, cells were induced with tetracycline (0 to 1 µg ml^{-1}) for 24 h. Cells were collected and lysed in cell extraction buffer (Invitrogen, FNN0011; added fresh, 1 mM phenylmethyl sulfonyl fluoride (PMSF) and protease inhibitor tablet (Sigma, 11836170001)). The proteins were separated on SDS–PAGE gels and analyzed by western blotting (anti-FLAG M2, 1:1,000 dilution; Sigma, F1804).

Generation of KO cell lines. HEK293T WT cells (0.8 million) were seeded in a six-well cell culture dish the day before transfection. Two micrograms of px330 plasmid containing the gRNA for the target protein and 200 ng of pcDNA3-FKBP-EGFP-HOTag3 (Addgene, 106924) were cotransfected using Lipofectamine 2000 (Thermo Scientific, 11668027). Cells were sorted by FACS 2 d after transfection. The top 95% of cells displaying green fluorescent protein (GFP) signals were sorted as single cells into 96-well dishes. Genomic PCR and western blot (anti-NSUN2, Proteintech, 20854–1-AP, 1:2,000; anti-NSUN5, Proteintech, 15449–1-AP, 1:2,000; anti-TRMT2A, Proteintech, 16199–1-AP, 1:1,000; anti-DUS3L, Proteintech, 15643–1-AP, 1:2,000; anti-β-actin, Cell Signaling, 8H10D10, 1:10,000 as a loading control) were performed to confirm KO.

Poly(A) pulldown of 5-FCyd-treated cells for mass spectrometry. HEK293T cells were treated with 10 µM 5-FCyd at 80% confluency for 12 h. Cells from 10×10 cm dishes were used per replicate. Poly(A) pulldown was performed following literature precedent[8,9] with minor modifications. Ten milliliters of fresh lysis buffer (20 mM Tris, pH 7.5, 500 mM LiCl, 0.5% lithium dodecyl sulfate, 1 mM EDTA and a fresh protease inhibitor tablet (Sigma, 11836170001) and 5 mM DTT) and 750 µl of pre-equilibrated oligo(dT)$_{25}$ magnetic beads (NEB, S1419S) were used per sample. Lysis, pulldown and washes were performed at room temperature. Buffers used for washing were lysis buffer, NP-40 wash buffer (20 mM Tris, pH 7.5, 140 mM LiCl, 1 mM EDTA, 0.5% NP-40 (vol/vol); added freshly: 0.5 mM DTT), and NP-40-free wash buffer (20 mM HEPES, pH 7.5, 140 mM LiCl, 1 mM EDTA; added freshly, 0.5 mM DTT). The bound poly(A) RNA was eluted in 350 µl of elution buffer (20 mM HEPES, pH 7.5, 1 mM EDTA) by heating at 55 °C for 3 min. Benzonase nuclease (250 U; Millipore Sigma, 70746) was added to concentrated eluate and incubated at 37 °C for 1 h to release the crosslinked proteins from RNA. The protein concentration in the samples was determined by bicinchoninic acid (BCA) assay (Thermo Scientific, 23225).

Mass spectrometry-based proteomics. Samples were prepared mostly as previously described[50]. Eluted samples were dried using a vacuum evaporator at room temperature and taken up with 10 µl of 6 M guanidinium chloride in 200 mM EPPS, pH 8.0, to a final concentration of 0.2 µg µl^{-1}. The samples were subsequently digested with 20 ng µl^{-1} Lys-C (Wako) and 10 ng µl^{-1} trypsin (Promega) and labeled with TMT 6-plex (Thermo Fisher Scientific) and cleaned for LC–MS analysis with stage tips[51]. Approximately 2 µg of the sample was analyzed via LC–MS on an Orbitrap Fusion Lumos with a TMT-MS2 method as previously described[52].

Mass spectrometry-based proteomics data analysis. MS data analysis was performed essentially as previously described[53,54]. The MS data in the Thermo RAW format were analyzed using the Gygi Lab software platform (GFY Core version 3.8) licensed through Harvard University. The ratio in each channel was normalized by the channel's median change for the following proteins, which are experimentally expected to be the same across the conditions: SRSF3, SNRPE, SRSF10, SRSF1, SRSF4, CPSF1, SRSF7, SNRPB, SNRPF, SNRPD2, ZRANB2, POLRMT, TROVE2, FAM120A, SLTM, GEMIN5, MTPAP, MRPL43 and SNRPD1. The normalized values are then subjected to downstream statistical analysis.

RNABPP validation by western blotting. For NSUN2, NSUN5 and TRMT2A, Flp-In T-Rex 293 cell lines expressing the 3×FLAG-tagged proteins were used to validate the proteomics results. For DUS3L, HEK293T WT cells were used. 5-FCyd treatment was performed as described previously. For poly(A) pulldown, washes with NP-40-free buffer were omitted, and RNase cocktail (Invitrogen, AM2288) instead of benzonase was used for RNA digestion. The eluates were separated on an SDS–PAGE gel and analyzed by western blotting.

Gel-based crosslinking assay. For the gel-based crosslinking assay, 3×FLAG-tagged proteins were expressed by transient transfection in HEK293T cells or by tetracycline induction in the corresponding Flp-In T-Rex 293 cell lines for NSUN2, NSUN5, NSUN6, DNMT2 and TRMT2A. For DUS3L, HEK293T WT cells were used for crosslinking of DUS3L with the panel of 5-halogenated pyrimidine compounds. To test the crosslinking between DUS3L C396A-mutant and uridine analogs, WT and DUS3L C396A were expressed in DUS3L-KO cells by transient transfection. A 12-h treatment was performed for all cell lines and compounds at the indicated concentrations. Cells were lysed in cell extraction buffer (Invitrogen), and the proteins were separated on a 10% SDS–PAGE gel followed by western blot analysis.

RNA isolation and nucleoside liquid chromatography–tandem mass spectrometry. Total RNA was extracted using TRIzol reagent (Thermo Fisher, 15596018) following the manufacturer's protocol. Small RNA was isolated using Zymo RNA Clean & Concentrator-5 (Zymo Research, R1016) following the manufacturer's protocol with a minor modification (adjusted RNA-binding buffer was made by mixing three parts of RNA-binding buffer and two parts of ethanol). For mRNA isolation, total RNA was subjected to two rounds of poly(A) selection using oligo-(dT)$_{25}$ beads (NEB, S1419S). The poly(A) RNA was further subjected to ribodepletion to remove rRNA contamination using the NEBNext rRNA Depletion kit (NEB, E6310) or custom synthesized probes[55]. Small RNA was removed from ribodepleted poly(A) RNA using Zymo RNA Clean & Concentrator-5 (Zymo Research). The RNA was digested and dephosphorylated with nuclease P1 (Wako, 145–08221) and Antarctic phosphatase (NEB, M0289) before LC–MS analysis. Briefly, 5–10 µg of RNA was digested in a 30-µl reaction with 2 U of nuclease P1 at 37 °C for 2 h (buffer composition: 7 mM sodium acetate, pH 5.2, 0.4 mM $ZnCl_2$). Dephosphorylation was performed next by adding 1.5 µl of Antarctic phosphatase and 3.5 µl of 10× Antarctic phosphatase buffer and incubating the reaction at 37 °C for another 2 h. The dynamic multiple reaction monitoring method was used to perform quantitative LC–QQQ–MS analysis of modified nucleosides on an Agilent 1260 Infinity II HPLC coupled to an Agilent 6470 triple quadrupole mass spectrometer in positive ion mode. A Hypersil GOLD C18 Selectivity HPLC Column (Thermo Fisher, 25003–152130; 3-µm particle size, 175-Å pore size, 2.1×150 mm, 36 °C) was used for all analyses, with a gradient composed of 0.1% formic acid in water (A) and acetonitrile (B) at a flow rate of 0.4 ml min^{-1} following literature precedent[56]. The following operating parameters for the mass spectrometer were used: gas temperature of 325 °C, gas flow rate of 12 liters min^{-1}, nebulizer pressure at 20 psi and capillary voltage at 2,500 V, with fragmentor voltage and collision energy optimized for each nucleoside. MS1 (parent ion) to MS2 (deglycosylated base ion) transition for each nucleoside was set as follows: m/z 262 → 130 for 5-FCyd, m/z 263 → 131 for 5-FUrd, m/z 258 → 126 for m^{5}C, m/z 259 → 127 for m^{5}U, m/z 247 → 115 for DHU, m/z 336 → 204 for i^{6}A, m/z 268 → 136 for A, m/z 244 → 112 for C and m/z 245 → 113 for U. Commercially available ribonucleosides were used to generate standard curves. The levels of 5-FCyd, 5-FUrd, m^{5}C, m^{5}U, DHU and i^{6}A were determined by normalizing the concentration of modified nucleosides to the concentration of the corresponding canonical nucleosides in the sample.

5-FUrd-iCLIP. Library preparation for iCLIP was adapted from the literature[34]. Flp-In T-Rex 293 cells expressing 3×FLAG-DUS3L (20×10 cm with 5-FUrd treatment and 40×10 cm for untreated control) were treated with 1 µg ml^{-1} tetracycline once the cells reached 60% confluency. After 12 h, the medium was changed to fresh medium containing 1 µg ml^{-1} tetracycline (control) or 1 µg ml^{-1} tetracycline and 100 µM 5-FUrd and cultured for another 12 h. Cells were washed twice with cold PBS, and 500 µl of lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% NP-40, 0.1% SDS and 0.5% sodium deoxycholate) was added to each dish. Lysates were treated with Turbo DNase (AM2239) and a low (1:200) or high (1:50) concentration of RNaseI (AM2295) at 37 °C for 3 min with rotation, after which they were centrifuged at 13,000 rpm for 15 min at 4 °C. The supernatant was incubated with anti-FLAG beads (200 µl of Protein G beads and 15 µg of anti-FLAG M2 antibody) at 4 °C overnight with rotation, and the immunoprecipitated material was then washed three times with high-salt buffer and once with CutSmart buffer (NEB). The beads were treated with Quick-CIP (NEB) at 37 °C for 30 min with rotation, washed twice with high-salt buffer and incubated with preadenylated and biotin-labeled L3 linker by using T4 RNA ligase I (NEB) at 16 °C overnight.

After washing three times with high-salt buffer, the beads were boiled with 60 µl of sample buffer, and 3 µl was analyzed by western blot for biotinylation with the chemiluminescent nucleic acid detection module (Thermo). The rest of the sample was gel purified based on the position of the biotin signal. Immunoprecipitated RNA was recovered using a D-tube dialyzer midi (Merck-Millipore) and digested with proteinase K (Roche). Reverse transcription was performed using oligonucleotides containing randomized barcodes (unique molecular identifiers (UMIs)) and two inversely oriented adaptor regions separated by a BamHI restriction site. cDNAs were size purified (70–80 nt, 80–100 nt and 100–150 nt) on TBE-Urea gels and then circularized by CircLigase II (Epicentre). Circularized cDNAs were digested with BamHI, and linearized cDNAs were amplified using Solexa primers and submitted for Illumina sequencing.

Bioinformatic analysis. The iCLIP data were processed using the iCount Primary Analysis pipeline (consensus mapping) on the iMAPs web server (https://imaps.genialis.com/iclip). Briefly, random UMIs were used to distinguish and discard PCR duplicate reads, and adaptor/barcode sequences were then removed. Trimmed reads were first mapped to tRNA/rRNA with STAR (v.2.7.0f). The unmapped reads were further mapped to GRCh38 with STAR (v.2.7.0f), and only uniquely mapping reads were used for further analysis. iCount-generated raw crosslinking sites were used for peak calling analysis by Paraclu with the following parameters: a minimal sum of scores inside a cluster of 10, a maximal cluster size of 4 and a minimal density increase of 2. Read counts were normalized per million uniquely mapping reads (RPM). Peaks that were unique to the 5-FUrd-treated sample or showed a fold change of greater than four (5-FUrd treatment versus control) were kept after intersection. Only peaks with read counts >10 in at least two of three replicates were selected for further analyses. Sequences five bases upstream and downstream were extracted from the reference and used for motif analysis by MEME (V.5.3.3) using -mod zoops -nmotifs 3 -minw 6 -maxw 50 -objfun classic -revcomp -markov_order 0.

Global protein translation assay. The global protein translation efficiency in HEK293T WT and DUS3L-KO cells was assessed using OP-puro (Click chemistry tool, 1407–5) based on literature precedent[57]. In brief, the cells were labeled with OP-puro by replacing the medium with fresh medium containing 50 µM OP-puro (20 mM reconstituted stock solution in DMSO) and incubating the cells at 37 °C for 1 h. Cells incubated in fresh medium without OP-puro were used as a negative control. Cu(I)-catalyzed azide/alkyne cycloaddition was performed to conjugate OP-puro to Cy3-azide (Click chemistry tools, AZ119–1) on fixed cells. Cells were resuspended in PBS containing 1 mM $Cu(II)SO_4$, 2 mM THPTA, 10 µM Cy3-azide and 10 mM sodium ascorbate (freshly prepared) and incubated at room temperature for 2 h in the dark. The reaction solution was removed at the end of the reaction, and cells were washed with ice-cold 0.1% Triton X-100 in PBS (PBST) three times. The cells were then resuspended in PBS containing Hoechst 33342 (Thermo Scientific, 62249; 20 µg ml^{-1}) and kept on ice and protected from light until flow cytometry analysis (less than 24 h). Each single-cell suspension was run on the BD LSRII SORP Flow Cytometer (BD Biosciences). The fluorescent Cy3 was excited by a 561-nm, 40-mW laser, and its emitted light was detected after passing through a 582/15 bandpass filter. The fluorescent Hoechst 33342 was excited by a 355-nm, 30-mW laser, and its emitted light was detected after passing through a 450/50 bandpass filter. To confirm consistency among cytometer runs, each sample and control was run on the cytometer three times, producing three replicate raw data files for each. On a forward scatter area versus side scatter area scatter plot, only cells were included in the primary analysis gate. On a forward scatter area versus forward scatter height scatter plot, only single cells were included in the secondary analysis gate. The tertiary analysis gate further excluded aggregates on a scatter plot of the Hoechst 33342 width value versus the Hoechst 33342 area fluorescence intensity. Each raw data file contained the aforementioned fluorescence and scatter values of 50,000 single cells. The Hoechst 33342 area fluorescence intensities of all cells in the tertiary gate were plotted on a histogram to represent DNA content of each cell. The Cy3 area fluorescence intensities of the same cell population were plotted on another histogram to represent the quantity of Cy3 present in each cell. The raw data files were analyzed in BD FACSDiva software version 8.0.2 (BD Biosciences) and FCS Express version 7 (DeNovo Software). The median fluorescence intensities of Cy3 from each cytometer run were used to compare the global protein translation efficiency among the corresponding cell lines.

Fluorescence microscopy. For fluorescence microscopy experiments, a poly-L-lysine-coated (Sigma, P8920) 12-mm glass coverslip was placed in each well before cell seeding. The OP-puro labeling was performed as described previously. At the end of the labeling, cells were washed once with DPBS and fixed with 3% paraformaldehyde (PFA) in PBS for 20 min at room temperature and washed with PBS three times. The cells were then permeabilized with PBST for 20 min at room temperature and washed twice with PBS. Click chemistry with Cy3-azide was performed as described in the previous section. Coverslips were incubated upside down with drops (100 µl) of freshly prepared reaction mixture. The reaction was allowed to proceed for 2 h at room temperature in the dark. Cells were then washed five times with PBST for 10 min to remove non-specific binding. To stain for DUS3L, cells were blocked with 5% goat serum in PBST for 1 h. The coverslips were incubated with anti-DUS3L antibody (Proteintech, 15643–1-AP, 1:200) for 2 h and washed with PBST three times for 5 min each. Goat anti-rabbit Alexa 488 antibody (Jackson ImmunoResearch, 1:400) for 1 h was used for secondary antibody staining. The cells were washed twice with PBST for 5 min, stained with Hoechst 33342 (Thermo Scientific, 1 µg ml^{-1}) for 5 min and washed with PBS twice for 5 min. The coverslips were mounted in ProLong Gold AntiFade Reagent (Life Technologies) and sealed with nail polish. Images of fixed cells were acquired using NIS Elements AR software and a Nikon Eclipse Ti microscope equipped with a ×100 objective and CMOS camera. Images used for direct comparison were acquired using standardized illumination and exposure settings and displayed with identical lookup table settings.

Cell viability assay. HEK293T WT and DUS3L-KO cells were plated in 96-well culture plates (4,000 cells in 200 µl of medium per well) on day 0. Cell viability was measured daily using the MTS assay (CellTiter 96 Aqueous Non-Radioactive Cell Proliferation Assay; Promega, G5430) for a total of 3 d (day 1 to day 3). The absorbance was read at 490 nm using a Synergy H1 Microplate Reader (BioTek).

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

Data availability

The sequencing data reported in this paper have been deposited in the NCBI Gene Expression Omnibus (accession code GSE175825). The proteomics data reported in this paper are available via ProteomeXchange with identifier PXD022645. Source data are provided with this paper.

References

50. Gupta, M., Sonnett, M., Ryazanova, L., Presler, M. & Wuhr, M. Quantitative proteomics of *Xenopus* embryos I, sample preparation. *Methods Mol. Biol.* **1865**, 175–194 (2018).
51. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
52. Sonnett, M., Yeung, E. & Wuhr, M. Accurate, sensitive, and precise multiplexed proteomics using the complement reporter ion cluster. *Anal. Chem.* **90**, 5032–5039 (2018).
53. Sonnett, M., Gupta, M., Nguyen, T. & Wuhr, M. Quantitative proteomics for *Xenopus* embryos II, data analysis. *Methods Mol. Biol.* **1865**, 195–215 (1865).
54. Huttlin, E. L. et al. A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143**, 1174–1189 (2010).
55. Adiconis, X. et al. Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nat. Methods* **10**, 623–629 (2013).
56. Su, D. et al. Quantitative analysis of ribonucleoside modifications in tRNA by HPLC-coupled mass spectrometry. *Nat. Protoc.* **9**, 828–841 (2014).
57. Hidalgo San Jose, L. & Signer, R. A. J. Cell-type-specific quantification of protein synthesis in vivo. *Nat. Protoc.* **14**, 441–460 (2019).

Acknowledgements

We thank C. DeCoste and K. Rittenbach at the Princeton University Flow Cytometry Resource Facility for assistance with flow cytometry analysis. We thank L. Ryazanova (supported by the Princeton Catalysis Initiative and the Lewis–Sigler Collaboration Fund) for technical assistance. R.E.K. acknowledges support from a National Science Foundation CAREER award (MCB-1942565), the National Institute of Health (R01GM132189), the Sidney Kimmel Foundation and the Alfred P. Sloan Foundation. This work was supported by NIH grant R35 GM128813 (to M.W.). T.N. was supported by the American Heart Association. W.D. was generously supported by the Edward C. Taylor 3rd Year Graduate Fellowship in Chemistry. A.L. was supported by the Princeton Catalysis Initiative. All authors thank Princeton University for financial support.

Author contributions

R.E.K. conceived the study, analyzed data, wrote the manuscript and performed experiments. W.D. performed RNABPP experiments, nucleoside MS, crosslinking studies and protein translation assays. A.L. performed iCLIP experiments and bioinformatic analysis. R.W.L. performed the bioinformatic analysis. N.J.Y. performed cell viability assays and crosslinking studies. T.N. performed MS proteomics and associated data analysis. M.W. supervised T.N.

Competing interests

The authors declare no competing financial interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-021-00874-8.

Correspondence and requests for materials should be addressed to Ralph E. Kleiner.

Peer review information *Nature Chemical Biology* thanks Jing Yang 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.

nature research

Corresponding author(s): Ralph Kleiner

Last updated by author(s): Jun 9, 2021

Reporting Summary

Nature Research 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 Research 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 (*n*) 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 *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	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. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	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 *d*, Pearson's *r*), 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	Mass spectrometry proteomics data was collected using Orbitrap Fusion Lumos version 3.3.2782.34 (Thermofisher Scientific). Nucleoside LC-MS data was recorded using Agilent MassHunter Workstation Data Acquisition version 10.0. Flow cytometry data was collected using BD FACSDiva version 8.0.2 (BD Biosciences, San Jose, CA.). Fluorescence microscopy data acquired by NIS Elements AR software version 4.60.00. Cell viability data was acquired using Gen5 version 2.08 (BioTek). All codes and software used for iCLIP data collection were described in detain in methods part.
Data analysis	Mass spectrometry proteomics data analysis was performed using GFY Core version 3.8 - a suite of software tools developed in the Gygi Lab and licensed by Harvard University & Python version 3. SEQUEST algorithm version 28 (rev. 12) was also used for MS data analysis. Nucleoside LC-MS data was analyzed by Agilent MassHunter Qualitative Analysis version 10.0 and Agilent QQQ Quantitative Analysis. Flow cytometry data was analyzed by BD FACSDiva software version 8.0.2 (BD Biosciences, San Jose, CA.) and FCS Express Version 7 (DeNovo Software, Pasadena, CA.). Multiple t tests for the proteomics data and all the unpaired t tests for nucleoside LC-MS, flow cytometry experiment, and cell viability assay were performed using GraphPad Prism version 7.04. The iCLIP data was processed using the iCount Primary Analysis pipeline (consensus mapping) on iMAPs web server (https://imaps.genialis.com/iclip). Reads were mapped by using STAR (v.2.7.0f). Motif were generated by using MEME (V.5.3.3)

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 Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Mass spectrometry proteomics data is provided in a separate Excel file.
For nucleoside LC-MS experiment, nucleoside concentrations and calculated ratios are provided in SI. Examples of chromatograms and calibration curves are also provided in SI.
For flow cytometry analysis, median values of Hoechst and Cy3 signals of all the measurements were presented in a table in SI. Examples of Cy3 histograms are shown in Figure 5b. A gating example is shown in SI.
Cell viability data is provided in SI.
The sequencing data reported in this paper have been deposited in the NCBI Gene expression omnibus (accession code: GSE175825).
Other data can be obtained from the authors upon request.

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	For LC-MS proteomics experiment, LC-QQQ-MS experiments, and iCLIP experiment, n = 3 was selected to offer sufficient statistical power. For IF and imaging experiments, n = 2 was selected to guarantee the consistency of the images. For flow cytometry experiment, n = 3 was used for independent biological replicates, and 3 technical replicates were used for each biological replicates to offer sufficient statistical power. For cell viability assays, n = 12 was used to offer sufficient statistical power.
Data exclusions	For all the experiments reported for this manuscript, no data was excluded.
Replication	All experiments were independently replicated, with biological and technical replicates listed in the legends of the corresponding figures. All attempts at replication were successful.
Randomization	No randomization was performed because we assumed little selection bias in cell culture.
Blinding	No blinding was performed as there was no need to prevent the study participants (cell culture) from knowing the treatment conditions. Data collection was largely performed using automated analysis software.

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
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

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

Antibodies

Antibodies used	Monoclonal ANTI-FLAG® M2 antibody produced in mouse was purchased from Sigma (catalog #F1804).

Anti-β-actin (mouse) antibody was purchased from Cell Signaling (catalog #3700).
Anti-NSUN2 rabbit polyclonal antibody was purchased from proteintech (catalog #20854-1-AP, lot #00047872).
Anti-NSUN5 rabbit polyclonal antibody was purchased from proteintech (catalog #20854-1-AP, lot #00006626).
HTF9C rabbit antibody (TRMT2A: human, mouse, rat) was purchased from proteintech (catalog #16199-1-AP, lot #00016863).
Anti-DUS3L antibody (rabbit) was purchase from proteintech (catalog #15643-1-AP, lot #00006825).
Alexa Fluor 488 AffiniPure Goat Anti-Rabbit IgG (H+L) was purchased from Jackson Immunoresearch (111-545-003).

Validation

Validation of anti-FLAG M2 antibody can be obtained from Sigma website: https://www.sigmaaldrich.com/catalog/product/sigma/f1804?lang=en&region=US
Validation of anti-β-actin antibody can be obtained from Cell Signaling website: https://www.cellsignal.com/products/primary-antibodies/b-actin-8h10d10-mouse-mab/3700
Validation of anti-NSUN2 antibody can be obtained from proteintech website: https://www.ptglab.com/products/NSUN2-Antibody-20854-1-AP.htm
Validation of anti-NSUN5 antibody can be obtained from proteintech website: https://www.ptglab.com/products/NSUN5-Antibody-15449-1-AP.htm
Validation of HTF9C antibody can be obtained from proteintech website: https://www.ptglab.com/products/TRMT2A-Antibody-16199-1-AP.htm
Validation of anti-DUS3L antibody can be obtained from proteintech website: https://www.ptglab.com/products/DUS3L-Antibody-15643-1-AP.htm
Validation of Alexa Fluor 488 AffiniPure Goat Anti-Rabbit IgG (H+L) can be obtained from Jackson Immunoresearch website: https://www.jacksonimmuno.com/catalog/products/111-545-003

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

HEK293T WT cells (ATCC #CRL-3216) were a gift from Tom Muir at Princeton University.
Flp-In T-Rex 293 cells (ThermoFisher cat# R78007) were a gift from John LaCava at The Rockefeller University.
Knockout cell lines and stable cell lines expressing 3x-FLAG tagged constructs were generated in our lab.

Authentication

Stable cell lines expressing 3x-FLAG tagged constructs were confirmed by western blot.
Knockout cell lines were confirmed by genomic PCR and western blot.
HEK293T WT and parent Flp-In cell lines were not authenticated.

Mycoplasma contamination

Parent cell lines were tested for mycoplasma contamination.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

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

HEK293T WT and DUS3L KO cells were treated with 50 µM OP-puro for 1 h at 37 °C. The cells were then fixed with 1% PFA and permeabilized with PBST. CuAAC reaction was then performed on cells to conjugate OP-puro to Cy3 azide. After the reaction, cells were washed 3 times with PBST to remove free Cy3-azide. The cells were then resuspended in PBS containing 20 µg/mL Hoechst 33342 and kept on ice before the flow cytometry experiment (less than 24 h).

Instrument

BD LSRII SORP Flow Cytometer (BD Biosciences, San Jose, CA) was used for the flow cytometry experiment.

Software

BD FACSDiva software version 8.0.2 (BD Biosciences, San Jose, CA.) and FCS Express Version 7 (DeNovo Software, Pasadena, CA.) were used for data analysis.

Cell population abundance

50,000 single cells were analyzed for cytometer run, which account for 85 - 92 % of the total cell population analyzed by the flow cytometer.

Gating strategy	On an FSC-A (Forward Scatter-Area) vs. SSC-A (Side Scatter-Area) scatter plot, only cells were included in the primary analysis gate. On an FSC-A (Forward Scatter-Area) vs. FSC-H (Forward Scatter-Height) scatter plot, only single cells were included in the secondary analysis gate. The tertiary analysis gate further excluded aggregates on a scatter plot of the Hoechst 33342-Width value vs. the Hoechst 33342-Area fluorescence intensity.

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