

# *N*<sup>2</sup>-methylguanosine modifications on human tRNAs and snRNA U6 are important for cell proliferation, protein translation and pre-mRNA splicing

Can Wang<sup>1</sup>, Nathalie Ulryck<sup>1</sup>, Lydia Herzel<sup>2</sup>, Nicolas Pythoud<sup>3</sup>, Nicole Kleiber<sup>2</sup>, Vincent Guérineau<sup>4</sup>, Vincent Jactel<sup>5</sup>, Chloé Moritz<sup>3</sup>, Markus T. Bohnsack<sup>2,6</sup>, Christine Carapito<sup>3</sup>, David Touboul<sup>4,7</sup>, Katherine E. Bohnsack<sup>2,\*</sup> and Marc Graille<sup>1,\*</sup>

<sup>1</sup>Laboratoire de Biologie Structurale de la Cellule (BIOC), CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France

<sup>2</sup>Department of Molecular Biology, University Medical Center Göttingen, 37073 Göttingen, Germany

<sup>3</sup>Laboratoire de Spectrométrie de Masse BioOrganique, CNRS, Université de Strasbourg, IPHC UMR 7178, Infrastructure Nationale de Protéomique ProFI, FR2048 Strasbourg, France

<sup>4</sup>Université Paris-Saclay, CNRS, Institut de Chimie des Substances Naturelles, UPR 2301, 91198 Gif-sur-Yvette, France

<sup>5</sup>Laboratoire de Synthèse Organique (LSO), CNRS, École polytechnique, ENSTA, Institut Polytechnique de Paris, 91120 Palaiseau, France

<sup>6</sup>Cluster of Excellence “Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells” (MBExC), Göttingen, Germany

<sup>7</sup>Laboratoire de Chimie Moléculaire (LCM), CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France

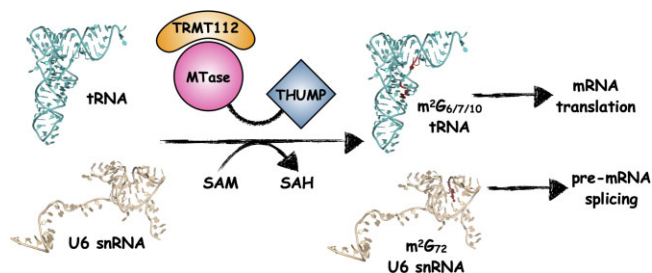
\*To whom correspondence should be addressed. Tel: +33 169 334890; Email: marc.graille@polytechnique.edu

Correspondence may also be addressed to Katherine E. Bohnsack. Tel: +49 551 3969305; Email: katherine.bohnsack@med.uni-goettingen.de

## ABSTRACT

Modified nucleotides in non-coding RNAs, such as tRNAs and snRNAs, represent an important layer of gene expression regulation through their ability to fine-tune mRNA maturation and translation. Dysregulation of such modifications and the enzymes installing them have been linked to various human pathologies including neurodevelopmental disorders and cancers. Several methyltransferases (MTases) are regulated allosterically by human TRMT112 (Trm112 in *Saccharomyces cerevisiae*), but the interactome of this regulator and targets of its interacting MTases remain incompletely characterized. Here, we have investigated the interaction network of human TRMT112 in intact cells and identify three poorly characterized putative MTases (TRMT11, THUMP3 and THUMP2) as direct partners. We demonstrate that these three proteins are active *N*<sup>2</sup>-methylguanosine (*m*<sup>2</sup>G) MTases and that TRMT11 and THUMP3 methylate positions 10 and 6 of tRNAs, respectively. For THUMP2, we discovered that it directly associates with the U6 snRNA, a core component of the catalytic spliceosome, and is required for the formation of *m*<sup>2</sup>G, the last ‘orphan’ modification in U6 snRNA. Furthermore, our data reveal the combined importance of TRMT11 and THUMP3 for optimal protein synthesis and cell proliferation as well as a role for THUMP2 in fine-tuning pre-mRNA splicing.

## Graphical abstract



## INTRODUCTION

In recent years, the epitranscriptome, a term collectively describing modified nucleotides throughout the transcriptome, has emerged as a key aspect of eukaryotic gene expression regulation (1). Indeed, almost all RNA types involved in splicing (small nuclear (sn)RNAs) or translation (transfer (t)RNAs and ribosomal (r)RNAs) of messenger (m)RNAs are decorated with numerous chemically diverse modifications often referred to as epitranscriptomic marks

(2–6). These marks are often conserved during evolution but many of their functions have remained elusive due to the lack of obvious phenotypes or because they are ‘orphan’ modifications for which the cognate writer enzymes remain unknown. Importantly, as writer enzymes are progressively identified, the critical roles of these RNA modifications in cellular development and in human pathologies, such

Received: December 21, 2022. Revised: April 21, 2023. Editorial Decision: May 16, 2023. Accepted: May 22, 2023

© The Author(s) 2023. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

as neurodevelopmental disorders and cancers, are being elucidated (4,7–15).

The most frequent RNA modification is methylation. It is primarily catalyzed by *S*-adenosyl-L-methionine (SAM)-dependent MTases, which structurally belong to two main classes: class I enzymes adopting a Rossmann fold and class IV with a SPOUT trefoil knotted fold (16). In eukaryotes, many class I enzymes are active as holoenzymes composed of a catalytic subunit associated with at least one activating partner (17,18). A prominent example is METTL3-METTL14, the major m<sup>6</sup>A mRNA MTase complex (19) but it is also the case for many tRNA MTases, such as the METTL1-WDR4 (Trm8–Trm82 in *S. cerevisiae* (yeast)), FTSJ1-WDR6 (Trm7–Trm732 in yeast) and TRMT61A-TRMT6 (Trm61–Trm6 in yeast) complexes (17,18).

Interestingly, TRMT112 proteins, which are conserved in all three domains of life, acts as an allosteric regulator of several archaeal and eukaryotic MTases targeting diverse substrates (tRNAs, rRNAs and proteins) involved in protein synthesis (17,20,21). The TRMT112 protein and its associated MTases are best characterized in *S. cerevisiae* while less is known about the human homologues. In yeast, Trm112 interacts with four MTases: Bud23 (BUD23 or WBSCR22 in human), Mtq2 (HEMK2 or MTQ2 in human), Trm9 (ALKBH8 in human) and Trm11 (probably TRMT11 in human). The Trm112–Bud23 complex catalyzes the formation of 7-methylguanosine (m<sup>7</sup>G) on the 18S rRNA and participates in the 40S ribosomal subunit biogenesis pathway (22,23). The Trm112–Mtq2 complex modifies the side chain of the glutamine residue of the universally conserved GGQ motif from the translation termination factor eRF1, which recognizes stop codons present in the ribosomal A-site (24,25,26,27). This eRF1 motif enters into the peptidyl transferase center of ribosomes to trigger the release of newly synthesized proteins (28). This complex has also recently been shown to be important for the biogenesis of the large ribosomal subunit (29). In humans, this complex also modifies the lysine 12 side chain of histone H4 and is important for prostate cancer cell proliferation (30). The yeast Trm112–Trm9 and Trm112–Trm11 complexes contribute to translation elongation by modifying specific tRNAs (31–35). Trm9 is responsible for adding a methyl group to 5-carboxymethyl-(2-thio)-uridine (cm<sup>5</sup>(s<sup>2</sup>)U) to form 5-methoxycarbonylmethyl-(2-thio)-uridine (mcm<sup>5</sup>(s<sup>2</sup>)U) at position 34 of some tRNAs (32–35). In humans, two proteins share significant similarities with yeast Trm9 and both interact with TRMT112: ALKBH8 modifies tRNAs in a similar manner to Trm9 (36,37) while the biochemical function of TRMT9B (also known as KIAA1456 or hTrm9L), a protein acting as a tumor suppressor, remains to be clarified (38,39). Trm11 catalyzes the formation of m<sup>2</sup>G at position 10 of some tRNAs in both yeast and several archaea (31,40–44). Identification and characterization of these TRMT112-containing holoenzymes have highlighted common features. For example, TRMT112 (i) stabilizes these MTases *in cellulo*, (ii) activates them by enhancing SAM binding, and (iii) contributes to substrate binding. Furthermore, the MTases share a common mode of interaction with TRMT112 such that their interactions are mutually exclusive (17,20).

In metazoans, the METTL5 MTase, a protein with no orthologue in *S. cerevisiae* yeast, was identified as a direct partner of TRMT112 and subsequently, the TRMT112–METTL5 complex was revealed to mediate the formation

of N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) on the 18S rRNA (45–50). This demonstrated that the interaction network of human TRMT112 extends beyond the homologues of the MTases known from yeast. Indeed, a recent study uncovered seven MTases (BUD23, METTL5, MTQ2, ALKBH8, TRMT11, THUMP2 and THUMP3) associated with TRMT112 (51) and TRMT112-THUMP3 was shown to target specific tRNAs, where it introduces m<sup>2</sup>G modifications at position 6 (or position 7 in the case of tRNA<sup>Trp</sup>) (52).

Here, we identified human TRMT112-interacting proteins using an unbiased proximity labelling approach in intact cells, revealing not only associated MTases, but also numerous other factors. We focus on three TRMT112-MTase complexes, TRMT112-TRMT11, -THUMP2 and -THUMP3 complexes as the roles of the tRNA m<sup>2</sup>G<sub>6/7</sub> modifications installed by TRMT112-THUMP3 in translation regulation are not yet understood, TRMT11 has not yet been characterized in human cells and the function of THUMP2 remains elusive. We characterize these three holoenzymes as m<sup>2</sup>G MTases modifying small RNAs and show that their absence affects the proliferation of HCT116 colon cancer cells. TRMT11 and THUMP3 are responsible for m<sup>2</sup>G formation at positions 10 and 6, respectively, of specific tRNAs. We show that lack of TRMT11, THUMP3 or both m<sup>2</sup>G tRNA MTases does not significantly affect substrate tRNA stability, folding or aminoacylation. However, we demonstrate that the absence of m<sup>2</sup>G in tRNAs impairs protein synthesis. Notably, our data indicate that m<sup>2</sup>G modifications at positions 6/7 and 10 function synergistically. Excitingly, we also discover that THUMP2 binds the U6 snRNA, the central snRNA involved in the pre-mRNA splicing reaction, and is essential for m<sup>2</sup>G formation on this snRNA, an ‘orphan’ modification identified more than 40 years ago (5,53,54). In rodents (5,53,54), this m<sup>2</sup>G is found at position 72, which is located within the spliceosome catalytic core, and we reveal that this enzyme is important for pre-mRNA splicing in human cells.

## MATERIALS AND METHODS

### Statistical analyses

Unless otherwise stated, all data are expressed as mean ± standard deviation (SD) calculated from three (growth curves, enzymatic activities, AHA incorporation, RNA levels) or five (polysome profiles by sucrose gradient, m<sup>2</sup>G and m<sup>2,2</sup>G quantifications) biological replicates. Statistical differences were determined using an independent two-tailed Student’s *t*-test to compare values for a given mutant to those of the controls (WT samples). \*\*\*\* *P* < 0.0001; \*\*\* *P* < 0.001; \*\* *P* < 0.01; \* 0.01 < *P* < 0.05; ns: not statistically significant. Details of the statistical analyses performed during bioinformatics analyses are given in the corresponding sections below.

### Molecular cloning

Plasmids used in this study are listed in Supplementary Table S1. Plasmids generated in this study were created using standard cloning methods involving the oligonucleotides listed in Supplementary Table S2, restriction enzyme digestions and T4 ligation. The human *TRMT112* coding sequence was amplified from pFF6 (kind gift from Dr. V.

Heurgué-Hamard). *In vitro* synthesized, codon optimized DNA fragments encoding the human *THUMP2* or *THUMP3* coding sequences (obtained from Integrated DNA Technologies, Belgium) were used as templates in the preparation of plasmids encoding for *THUMP2* or *THUMP3*, respectively. For co-expression of the *Bos taurus* (*Bt*) *THUMP2*-TRMT112 complex in *E. coli*, an *in vitro* synthesized DNA fragment (Integrated DNA Technologies, Belgium) containing the coding sequences of both proteins (His-tagged *BtTHUMP2* and untagged *BtTRMT112*) was used as a template to generate pMG1032. The coding sequence of human *TRMT11* was amplified by PCR from the HsCD00515724 plasmid (DNASU repository, (55)). Plasmids for expression of guide RNAs for CRISPR-Cas genome editing were based on pSpCas9(BB)-2A-puro vector and were generated by cloning of annealed oligonucleotides with appropriate restriction enzyme cleavage site overhangs. To generate stably transfected HEK293 cell lines for the inducible expression of N-terminally 2 × Flag-His<sub>6</sub> or C-terminally His<sub>6</sub>-2 × Flag tagged *THUMP2*, the *THUMP2* CDS was cloned into appropriate pcDNA5-based plasmids (Supplementary Table S1). Site-directed mutagenesis was performed according to Zheng *et al.* (56) where necessary using oligonucleotides listed in Supplementary Table S2.

### BioID (proximity-dependent biotin identification)-mass spectrometry (MS)

The BioID experiments were performed mostly as described in Chapat *et al.* (57). Briefly,  $8 \times 10^6$  HEK293T cells were cultured in DMEM complete medium (DMEM medium (Gibco) supplemented with 10% fetal bovine serum (Gibco), 200 U/ml of penicillin and 200 µg/ml streptomycin) in 15 cm plates at 37°C and 5% of CO<sub>2</sub>. 24 h after seeding or when the cells reached 80% confluence, pcDNA3.1-based plasmids for the expression of BirA(R118G)\*-HA tagged eGFP (pMG1127), TRMT112 (pMG891) or TRMT112-T5R (pMG906) or the BirA\*-HA cassette alone (Addgene #36047) were transiently transfected (Supplementary Table S1). For transient transfection, 5 µg of plasmids (2 µg of pcDNA3.1 MCS-BirA(R118G)-HA or of pMG1127 supplemented with 3 µg of empty pcDNA3.1 plasmid or 5 µg of pMG891 or of pMG906) in a total volume of 300 µl of Opti-MEM medium (Gibco) were mixed with 300 µl of Opti-MEM medium containing 15 µl of Lipofectamine 2000 (Invitrogen) and added gently to the plates. After 24 h, 50 µM biotin was added to the culture medium and cells were grown for a further 24 h. Cells were washed with 1xPBS (Gibco) and harvested in lysis buffer (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1% Nonidet P-40, 0.4% SDS, 1.5 mM MgCl<sub>2</sub>, 1 mM EGTA, benzamide, and cOmplete EDTA-free Protease inhibitor (1 × PIC)). Lysates were rotated at 4°C for 30 min and then centrifuged at 16 000 × g for 20 min at 4°C. For each condition, equal amounts of proteins were incubated with anti-Streptavidin high performance affinity resin (Cytiva; #17511301) pre-equilibrated in lysis buffer and incubated overnight at 4°C. Next, the supernatant was discarded and the beads were washed once with 500 µl of SDS wash buffer (50 mM Tris-HCl, pH 7.5, and 2% SDS), twice with lysis buffer, three times with 50 mM ammonium bicarbonate (pH 8.0). Proteins aggregated on beads were resuspended in 100 µl of NH<sub>4</sub>HCO<sub>3</sub> at 25 mM containing 2 µg of Trypsin/Lys-C mix (Mass Spec Grade, Promega,

Madison, WI, USA). Digestion was performed under agitation at 37°C during 4 h. Then, the same enzyme amount was added to the solution and the second digestion step was conducted overnight at 37°C under agitation. Supernatant was collected after a centrifugation at 500 × g for 2 min, and beads were washed with water. Both solutions were pooled. The enzymatic digestion was stopped by adding 6 µl of pure formic acid (FA, 2% final concentration). Samples were then centrifuged at 16 000 × g for 5 min and supernatants were collected (90% of the solution). Finally, the samples were vacuum dried, and then resolubilized in 400 µl of H<sub>2</sub>O/ACN (acetonitrile)/FA (98/2/0.1 v/v/v) prior to nanoLC-MS/MS analysis. Each condition was performed in five replicates.

NanoLC-MS/MS analyses were performed on a nanoAcquity UPLC device (Waters Corporation, Milford, USA) coupled to a Q-Exactive HF-X mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). Peptides were loaded on a symmetry C18 pre-column (20 mm × 180 µm with 5 µm diameter particles, Waters) before being separated on an ACQUITY UPLC BEH130 C18 column (250 mm × 75 µm with 1.7 µm diameter particles). The solvent system consisted of 0.1% FA in water (solvent A) and 0.1% FA in ACN (solvent B). Samples were loaded into the enrichment column over 3 min at 5 µl/min with 99% of solvent A. Peptides were then eluted at 450 nl/min with the following gradient of solvent B: from 1 to 8% over 2 min, 8 to 35% over 77 min, and 35 to 90% over 1 min. The system was operated in a data-dependent acquisition (DDA) mode with automatic switching between MS (mass range 375–1500 *m/z* with *R* = 120 000 at 200 *m/z*, automatic gain control fixed at  $3 \times 10^6$  ions, and a maximum injection time set at 60 ms) and MS/MS (mass range 200–2000 *m/z* with *R* = 15 000 at 200 *m/z*, automatic gain control fixed at 1.105, and the maximal injection time set to 60 ms) modes. The twenty most abundant peptides were selected on each MS spectrum for further isolation and higher energy collision dissociation (normalized collision energy set to 27), excluding unassigned, singly charged and over seven times charged ions. The dynamic exclusion time was set to 40 s.

Raw nanoLC-MS/MS data were processed using MaxQuant software (version 1.6.6.0). Peaks were assigned with the Andromeda search engine with trypsin specificity. The database used for the searches was extracted from UniProtKB-SwissProt and included all *Homo sapiens* entries (22 July 2019; Taxonomy ID = 9606; 20 409 entries). The minimum peptide length required was seven amino acids and a maximum of one missed cleavage was allowed. The precursor mass tolerance was set to 20 ppm for the first search and 4.5 ppm for the main search. The fragment ion mass tolerance was set to 20 ppm. Methionine oxidation and acetyl (Protein N-term) were set as variable modifications. The maximum false discovery rate was 1% for peptides and proteins with the use of a decoy strategy. The ‘match between runs’ option was deactivated. Unique peptides were used but modified peptides as well as their unmodified counterparts, were excluded from protein quantification. We used the ‘proteingroups.txt’ file with intensities (non-normalized intensities). The dataset was deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD038997 (58). Statistical analyses were performed using ProStaR software (59). Only proteins for which 5 intensity values were available in a single condition were kept. After log<sub>2</sub> transformation, intensities were normalized within

condition using vsn method and imputation of missing values were performed. For each sample, the slsa algorithm was used for the POV (Partially Observed Values) imputation, and missing values were replaced by the 2.5 percentile value for the MEC (Missing on the Entire Condition); statistical testing was performed using Limma. Benjamini–Hochberg method was used to adjust p-values for multiple testing and differentially expressed proteins were sorted out using a p-value threshold that guarantees an FDR below 1%.

### Co-immunoprecipitation (Co-IP) and western blotting

$3.5 \times 10^6$  HEK293T cells were seeded in 10 cm plates and the next day, were transfected with 2  $\mu$ g of plasmids using Lipofectamine 2000 as described above. After 24 h, cells were harvested and cleared cell extracts were prepared as described above. Cell extracts were incubated overnight at 4°C with either anti-Flag M2 affinity agarose gel (Sigma-Aldrich; #A2220) or anti-HA agarose beads (Pierce™, #26181) pre-equilibrated with lysis buffer (50 mM HEPES–KOH pH 7.5, 100 mM KCl, 2 mM EDTA, 0.1% NP40, 10% glycerol, 1 mM PMSF, 1 mM DTT supplemented with 1  $\times$  PIC). After several washing steps with the wash buffer (50 mM HEPES–KOH pH 7.5, 100 mM KCl, 2 mM EDTA), the beads were mixed with 2  $\times$  loading buffer (80 mM Tris–HCl pH 6.8, 20% glycerol, 2% SDS, 0.01% bromophenol blue, 2% 2-mercaptoethanol) and eluates were separated by sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) prior to transfer on nitrocellulose membrane for western blotting using antibodies described in Supplementary Table S3.

### IP-MS to detect proteins associated with THUMPD2

$8 \times 10^6$  HEK293T cells were seeded in 15 cm dishes and the next day, were transiently transfected with 5  $\mu$ g of plasmids (pMG976 or pMG1128; Supplementary Table S1) as described above. Cells were collected in lysis buffer (50 mM HEPES–KOH pH 7.5, 100 mM KCl, 2 mM EDTA, 0.1% NP40, 10% glycerol, 1 mM PMSF, 1 mM DTT supplemented with 1  $\times$  PIC) and a freeze-thaw strategy involving freezing cells in liquid nitrogen for 5 min and thawing at 37°C was used for lysis. The lysate was then incubated at 4°C for 20 min followed by centrifugation (20 000  $\times$  g, 4°C, 15–30 min) to remove cellular debris. Cleared cell extracts were incubated with pre-washed anti-Flag beads overnight at 4°C following which, the beads were washed five times with lysis buffer complemented with 1  $\times$  PIC. The washed beads were directly frozen in liquid nitrogen and stored at -80°C. Samples were sonicated prior to transfer to 1.5 ml Eppendorf tubes compatible with a magnetic rack. After incubating the beads on the magnetic rack, the supernatant was discarded. The beads were washed three times with 200  $\mu$ l of 50 mM ammonium bicarbonate and finally resuspended in 100  $\mu$ l of ammonium bicarbonate (50 mM) before the addition of 2  $\mu$ g of Trypsin/Lys-C. Proteins were digested for 4 h at 37°C, 300 rpm. Then 2  $\mu$ g of Trypsin/Lys-C was added again and the samples were incubated overnight at 37°C, 300 rpm. The next day, supernatants were collected after incubation on the magnetic holder. The beads were washed with 100  $\mu$ l of LC–MS grade water and pooled with the supernatants. Digestion was stopped by the addition of formic acid at a final concentration of 2%. Samples were transferred

back to a new tube after incubation on the magnetic stand to remove any residual beads, before being evaporated to dryness and resuspended in 10  $\mu$ l of H<sub>2</sub>O/ACN/FA (98/2/0.1).

Peptides were analyzed using a nanoAcquity (Waters) coupled to a Q Exactive HF-X (ThermoFisher Scientific) in the same condition as upper described BioID samples.

Protein identifications were performed using Mascot (v.2.6.2). The database used contains all UniProtKB–SwissProt *Homo sapiens* entries, common MS contaminants and decoys (November 2019). A maximum of one missed cleavage was allowed, the precursor tolerance was set at 5 ppm and the fragment tolerance at 0.05 Da. Carbamidomethylation of cysteine residues was set as a fixed modification. N-term acetylation and oxidation of methionine residues were defined as variable modifications. Protein identifications were validated and relative label-free quantification was performed with Proline (v2.0; (60)) using only specific peptides without modifications and applying a 1% FDR at protein and PSM levels. Differential analysis was performed in Prostar (v1.22.6). Filtering was set to at least four values in one condition. A VSN normalisation between conditions and a quantile imputation (Quantile 2.5, Factor 1) were applied. The hypothesis test was performed using a Limma test to compare both conditions. The *P*-value was calibrated using the Benjamini–Hochberg calibration and *P*-value filtering was applied to achieve an FDR of 1.36% for a *P*-value = 1E-04. The complete dataset was deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD038967 (58,61).

### Heterologous expression and purification of mammalian proteins

To obtain the human THUMPD3–TRMT112 complex, *E. coli* BL21 (DE3) Gold cells (Agilent technologies) were co-transformed with plasmids for the expression of His<sub>6</sub>-ZZ tagged THUMPD3 (pMG889) and untagged TRMT112 (pFF6) (Supplementary Table S1). Protein expression was performed at 37°C in 1 l of terrific broth auto-inducible (TBAI, ForMedium; #AIMTB0260) media supplemented with kanamycin (100  $\mu$ g/ml) and chloramphenicol (25  $\mu$ g/ml). Bacteria were collected by centrifugation (3300  $\times$  g, 4°C, 20 min) and resuspended in lysis buffer (100 mM HEPES–NaOH pH 7.5, 200 mM NaCl; 5 mM 2-mercaptoethanol). Cell lysis was performed by sonication and cell debris were removed by centrifugation (20 000  $\times$  g, 4°C, 30 min). Soluble proteins were incubated with NiNTA agarose beads (pre-equilibrated with lysis buffer; Macherey–Nagel) for 1 h at 7°C. Washing steps, first with lysis buffer supplemented with 20 mM imidazole pH 7 and then with wash buffer (100 mM HEPES–NaOH pH 7.5, 1 M NaCl; 5 mM 2-mercaptoethanol) were used to remove non-specific proteins and nucleic acids bound to the proteins of interest. The retained proteins were eluted with lysis buffer supplemented with 400 mM imidazole pH 7 and incubated overnight at 4°C with human rhinovirus 3C protease to remove the His<sub>6</sub>-ZZ tag under dialysis conditions against lysis buffer. After a second incubation with NiNTA beads, the unbound proteins were collected, concentrated and diluted into buffer A (100 mM HEPES–NaOH pH 7.5, 50 mM NaCl; 5 mM 2-mercaptoethanol) before injection onto a Hitrap–Q column for ion-exchange chromatography

purification. The proteins were eluted in the lysis buffer using a gradient of NaCl from 50 mM to 1 M NaCl. Finally, purification was performed on a S200-16/60 size-exclusion column (GE Healthcare) equilibrated in lysis buffer. Eluted proteins were concentrated and used for *in vitro* functional assays.

To obtain the *Bos taurus* (*Bt*) THUMPD2–TRMT112 complex, a plasmid for the co-expression of His<sub>6</sub>-ZZ tagged *Bt*THUMPD2 and untagged *Bt*TRMT112 was utilized (pMG1032; Supplementary Table S1). The *Bt*THUMPD2–TRMT112 complex was co-expressed in *E. coli* BL21 (DE3) Gold cells growing at first 37°C for 3 h and then transferred to 18°C for overnight culture in TBAI media containing kanamycin (100 µg/ml). For purification of the complex, the same procedure as for the human THUMPD3–TRMT112 complex was applied with the difference that the composition of the lysis buffer was 20 mM Tris–HCl pH 8, 200 mM NaCl, 5 mM 2-mercaptoethanol, 10 µM ZnCl<sub>2</sub>.

A pNEA-vH-based plasmid for the expression of C-terminally His<sub>6</sub> tagged TRMT11 (pMG715) and the pPF6 plasmid for expression of untagged TRMT112 were co-transformed in *E. coli* BL21 (DE3) Gold cells for large-scale cell culture (1 l). Bacteria were first incubated at 37°C for 3 h and then shifted to 18°C for overnight incubation in TBAI media containing ampicillin (100 µg/ml) and chloramphenicol (25 µg/ml). The human TRMT11–TRMT112 complex was purified through affinity chromatography using NiNTA resin, ion-exchange chromatography using a heparin column and then size exclusion chromatography (S200-16/60 size-exclusion column) as described above. In this case, the lysis buffer composition was 20 mM Tris–HCl pH 7.5, 200 mM NaCl, 5 mM 2-mercaptoethanol.

### Size exclusion chromatography-multi-angle laser light scattering (SEC-MALLS)

The molecular weight of the different complexes in solution were determined by SEC-MALLS as previously described (44). Briefly, for each complex, a 100 µl sample of complex (1 mg/ml) was injected on a Superdex™ 200 Increase 10/300 GL column (GE-Healthcare) at a flow rate of 0.75 ml/min. As running buffers, we used the buffer used for the size-exclusion chromatography step performed on each complex (see above). Elution was followed by a UV-visible spectrophotometer, a RID-20A refractive index detector (Shimadzu), a MiniDawn TREOS detector (Wyatt Technology). The data were collected and processed with the program ASTRA 6.1 (Wyatt Technology). *M<sub>w</sub>* was directly calculated from the absolute light scattering measurements using a *dn/dc* value of 0.183.

### Radioactivity-based enzymatic assay

The *in vitro* tRNA methylation assays were performed as previously described (44). Briefly, The enzymatic assays were performed by mixing 75 pmol of *E. coli* tRNA<sub>i</sub><sup>Met</sup> (purified as previously described; (62)) with 3 pmol of enzyme and 2 µM of SAM ((Sigma-Aldrich, #7007) including 0.2 µM of [<sup>3</sup>H]-SAM, Perkin Elmer) in MTase buffer (25 mM K-phosphate pH 7.5, 50 µM EDTA, 5 mM MgCl<sub>2</sub>, 5 mM NH<sub>4</sub>Cl) in a total volume of 50 µl. The reaction was incubated overnight at 37°C and stopped by adding 5 ml of cold trichloroacetic

acid (TCA 5%) containing 0.5% of methionine, then 20 µl of RNA carrier (4 mg/ml, Colnbrook-Bucks-England), and followed by filtration on glass microfiber filters (Whatman GF/C). Beckman Coulter LS6500 scintillation counter was used to determine [<sup>3</sup>H] incorporation.

### Characterization of RNA modifications by MS

m<sup>2</sup>G nucleosides formed in tRNAs upon incubation with different enzymes were detected by High-performance liquid chromatography (HPLC)–MS using a modified version of the protocol published in (44). Briefly, 20 µg of tRNA substrate (for example, *E. coli* tRNA<sub>i</sub><sup>Met</sup>; corresponding to 834 pmol of tRNA) were incubated overnight with 0.8 µg of TRMT112–MTase complexes (34 pmol of enzyme) supplemented with 126 pmol of SAM at 37°C. The negative control consisted of only tRNA in the same conditions. RNA digestion into nucleosides and HPLC–MS procedures were as previously reported (44).

For nucleoside quantification, the samples were analyzed by LC–high resolution MS (HRMS) instrument based on the specific elution time and *m/z* value of the respective nucleosides. The experiments were performed with a LC 1260 Prime (Agilent Technologies, Waldbronn, Germany) coupled to a high-resolution tandem mass spectrometer QTOF 6546 (Agilent Technologies, Waldbronn, Germany). The column (HSS T3, 150 mm – 2.1 mm × 2.7 µm, Waters, USA) was chosen for an efficient separation of the three nucleosides without the use of any buffer such as ammonium acetate in the mobile phase. The elution gradient started from 100% water + 0.1% formic acid to 50% acetonitrile in 8 min then 100% acetonitrile during 2 min and finally back to the initial conditions. The injected volume was fixed at 2 µl. For electrospray (ESI) analysis, mass spectra were recorded in positive ion mode with the following parameters: gas temperature 325°C, drying gas flow rate 10 l min<sup>-1</sup>, nebulizer pressure 30 psi, sheath gas temperature 400°C, sheath gas flow rate 10 l min<sup>-1</sup>, capillary voltage 3500 V, nozzle voltage 500 V, fragmentor voltage 110 V, skimmer voltage 45 V, Octopole 1 RF Voltage 750 V. For ESI, internal calibration was achieved with two calibrants purine and hexakis (1 h,1 h,3 h-tetrafluoropropoxy) phosphazene (*m/z* 121.0509 and *m/z* 922.0098) providing a high mass accuracy better than 3 ppm. First of all, standards, including G (Sigma-Aldrich; #G6752), m<sup>2</sup>G (Sigma-Aldrich; #M4004) and m<sup>2,2</sup>G (Berry and associates; #PR-3702) used in this experiment, were analyzed by LC–HRMS to ensure the elution time and calibration curves from 22 to 66 000 fmol injected in triplicate were plotted for each nucleoside. For quantification, the nucleoside samples were diluted into different concentrations (as determined by the OD at 254 nm) to enable quantification on either G or m<sup>2</sup>G/m<sup>2,2</sup>G based on their relative concentrations.

The m<sup>2</sup>G modifications were mapped on *E. coli* tRNA<sub>i</sub><sup>Met</sup> by matrix-assisted laser desorption ionisation time-of-flight (MALDI-TOF) MS–MS as previously described (44).

### Knock-out HCT116 cell lines generation by CRISPR-cas9 technique

The protocol used was adapted from Ran *et al.* (63). HCT116 parental cell lines and the generated knock-out cell lines were cultured in McCoy 5A complete medium (McCoy 5A medium (Gibco) supplemented with 10% fetal bovine serum (FBS,

Gibco) and 100 U/ml Penicillin and 100 µg/ml Streptomycin). The sgRNAs were designed using the CHOPCHOP website (<https://chopchop.cbu.uib.no/>; (64)). For each target gene, three sgRNAs selected from the top of the proposed list were cloned into the pSpCas9(BB)-2A-puro vector (Addgene plasmid #62988; (63); Supplementary Table S1).

To generate knock-out cell lines,  $0.5 \times 10^6$  HCT116 parental cells were seeded on 6-well plates and after 24 h incubation, transfected with 500 ng of plasmids for expression of the different sgRNAs. After 24 h, transfected cells were selected with 1.5 µg/ml of puromycin for a further 72 h. Clonal cell lines were generated either by diluting into 0.8 cells per 100 µl and seeding into a 96-well plate, or by inoculating specific numbers of cells (for example, 100, 200 and 400 cells) into 15 cm dishes. Cells were cultured for 2–3 weeks until single colonies were visible. The single colonies were transferred into 12-well plates and cultured to 80% confluence. Cells were then collected and lysed to detect the specific protein expression by western blotting. Colonies showing no expression of the protein of interest were selected for genomic DNA extraction using the DNeasy® blood and tissue kit (Qiagen) according to the manufacturer's instruction. The targeted genome regions were amplified by PCR using primers designed according to the CHOPCHOP website (Supplementary Table S2) and products were subjected to Sanger sequencing (Eurofins-MWG).

### Cell growth curve

5000 cells were plated in 12-well plates for each experiment. Cell numbers were manually counted at different time points using a hemocytometer (KOVA Glasstic Slide 10 with quantitative grid) according to the manufacturer's instructions. Briefly, cells were collected following trypsin treatment and diluted to a suitable concentration (less than 50 cells in each grid after staining). Cells were stained by adding equal volume of 0.4% Trypan blue (Gibco) and the mixture was incubated at room temperature for 5 min. The sample was added to each grid and living cells were counted manually under a microscope. The total number of cells was determined using the following formula provided by the manufacturer: Total cell numbers = the average counting cells  $\times$  90 (factor)  $\times$  dilution times  $\times$  1000 (volume, µl)/81 (numbers of small grids).

### Colony formation assay

1000 cells were seeded in 6-well plates complemented with 3 ml medium. The cells were grown for around 10 days until the colonies became visible and reached a suitable size. The colonies were washed twice with PBS, fixed with 4% paraformaldehyde (PFA) for 15 min at room temperature, and stained with crystal violet solution for 10–30 min. The staining solution was then removed, and the colonies were washed gently with running water. The plate was dried at room temperature overnight and pictures were taken using a ChemiDoc (Bio-Rad).

### RNA extraction from human cell lines and enrichment of specific RNA populations

Total RNAs were extracted from the different human cell lines (grown in one dish of 10 cm or 15 cm plate) using TRIzol™ LS reagent (Invitrogen) according to manufacturer's

instructions. The total RNA pellet was resuspended with 20–50 µl of sterile or RNase-free water.

For small RNAs extraction, 0.1 volume of 5 M NaCl and 2 volumes of 100% cold ethanol were added to the total RNAs to precipitate the small RNAs. The RNA pellet was washed with 90% cold ethanol and then resuspended in 1 ml of 1 M NaCl solution followed by centrifugation at  $20\,000 \times g$  and 4°C for 30 min to remove large RNAs. The supernatant containing small RNAs was then precipitated by 2 volumes of 100% cold ethanol followed by dissolution into 1 ml of 1.8 M Tris-HCl pH 8. The tRNAs present in these small RNAs were deacylated by incubation at 37°C for 1.5 h and precipitated by adding 0.1 ml of 5 M NaCl and 2.2 ml of 100% cold ethanol prior to storage at –20°C. The small RNAs were harvested by centrifugation at  $20\,000 \times g$  and 4°C for 30 min and washed with 70% of cold ethanol. Finally, the RNA pellet was dried and dissolved in sterilized water. The composition of different RNA fractions was analyzed by a denaturing (urea) polyacrylamide gel electrophoresis (PAGE).

For total tRNA extraction, cells were lysed by incubation with buffer A (20 mM Tris-HCl pH 7.5, 10 mM KCl, 1.5 mM MgCl<sub>2</sub>, 1 mM DTT, 0.1% NP40) complemented with 1  $\times$  PIC at 4°C for at least 10 min. The soluble cell extract was centrifuged at  $20\,000 \times g$  and 4°C for 20 min and the supernatant transferred to a new tube. 1/9 volume of 125 mM EDTA was added and the mixture was incubated on ice for 15 min followed by addition of 1/9 volume of 5 M NaCl and incubation on ice for another 10 min. Next, the sample was transferred into a polycarbonate tube and subject to ultracentrifugation at  $200\,000 \times g$  for 2 h at 4°C to precipitate the large RNAs. The small RNAs present in the supernatant were extracted using TRIzol™. The extracted small RNAs (maximum amount of 100 µg of RNAs) were purified by size-exclusion chromatography column (Acclaim SEC-300; #079723) using a 100 mM NH<sub>4</sub>Ac buffer using a micro-Aktä machine and a 0.5 ml/min flow rate. *E. coli* tRNA<sub>i</sub><sup>Met</sup> was used as standard to determine the elution volume of tRNAs. The collected fractions were analyzed by denaturing PAGE to estimate the purity of the RNAs.

### tRNA quantification

Total cellular RNAs extracted using TRIzol™ LS reagent (Invitrogen) were treated with DNase I for 30 min at 37°C before a further purification step using the RNeasy® MinElute™ cleanup kit according to the manufacturer's instructions. tRNA demethylation and reverse transcription were performed with rtStar tRNA-optimized First-Strand Synthesis Kit (Arraystar, #AS-FS-004). qPCR for tRNAs was performed using the nrStar Human tRNA PCR Array (Arraystar, #AS-NR-001–1) and Arraystar SYBR Green qPCR Master Mix(ROX+) (Arraystar; #AS-MR-006–5). The data were analyzed by the  $\Delta C_t$  method. The  $\Delta C_t$  values were calculated for each tRNA using the following formula:  $\Delta C_t(\text{tRNA}_x) = C_t(\text{tRNA}_x) - C_t(\text{housekeeping genes})$ .

### Analysis of tRNA aminoacylation

Total RNA was extracted under acidic conditions using TRI Reagent® (Sigma-Aldrich) and 10 mM NaOAc (pH 5) at 4°C (65) and stored in acidic conditions (10 mM NaOAc at pH 5 and 1 mM EDTA at pH 8). As a control, a sample of the total RNA was incubated at pH 8.9 for 60 min at 37°C to achieve complete deacylation. 5 µg of each of the RNA samples

were separated in an acidic (pH 5) 10% polyacrylamide 7 M urea gel at 4°C for 20 h at 170 V. After transfer to a Hybond-N membrane (Cytiva), selected tRNAs were detected by northern blotting using probes listed in Supplementary Table S4.

#### Native PAGE analysis of tRNA structure

200 ng of total RNA extracted using TRI Reagent® (Sigma-Aldrich), resuspended in 10 mM Tris pH 7.4 and 50% glycerol, supplemented with bromophenol blue, were separated on 20% non-denaturing polyacrylamide gels. RNAs were transferred to Hybond-N membranes (Cytiva) and detected by northern blotting using probes listed in Supplementary Table S4.

#### Polysome profiling by sucrose gradient

10 × 10<sup>6</sup> cells were grown for 24 h in 15 cm plates prior to incubation with cycloheximide (100 µg/ml final concentration; Sigma-Aldrich; #C7698) for 10 min at 37°C. Cells were washed twice with 10 ml ice cold PBS supplemented with cycloheximide (100 µg/ml final concentration). 5 ml ice cold PBS supplemented with cycloheximide (100 µg/ml final concentration) were added and the cells were harvested by scraping. Cells were pelleted by centrifugation at 300 × g for 5 min at 4°C and resuspended in 400 µl of hypotonic buffer (5 mM Tris-HCl pH 7.5, 1.5 mM KCl, 2.5 mM MgCl<sub>2</sub> supplemented with 1 × PIC). After addition of cycloheximide (100 µg/ml final concentration), DTT (2 mM final concentration) and RNasin (100 U), cells were vortexed for 5 s. Next, Triton-X100 and Na-deoxycholate (0.5% final concentration each) were added and vortexing repeated. After 10 min incubation at 4°C, samples were centrifuged for 5 min at 4°C and 12 000 × g. The supernatant was transferred to a pre-chilled tube and the optical density (OD) at 260 nm was measured using a Nanodrop ND-1000 spectrophotometer (Labtech). Then, 300 µl of sample at OD<sub>260</sub> of 7.5 were loaded on a 15–45% sucrose gradient (prepared with a BioComp Gradient Master 107ip in 20 mM Tris-HCl pH 7.5, 100 mM KCl, 10 mM MgCl<sub>2</sub>, 1mM DTT, 100 µg/ml cycloheximide). After ultracentrifugation (200 000 × g for 90 min at 4°C), the fractions were collected using a Foxy Jr fraction collector (Teledyne ISCO) and the absorbance at 254 nm was measured with a UA-6 device (Teledyne ISCO).

#### AHA-click labeling assay

7 × 10<sup>5</sup> cells were seeded into 6-well plates and after overnight incubation, the medium was replaced by DMEM medium lacking methionine (Gibco; #21013-024) for 1 h. 100 µM L-azidohomoalanine (AHA; Sigma Aldrich; #900892) was then added for 3 h. Cells were washed twice with cold PBS and lysed in lysis buffer (50 mM Tris-HCl pH 8.0 and 1% SDS) followed by sonication for 2 min. The soluble cell extract was separated by centrifugation at 20 000 × g and 4°C for 10 min and transferred into a new tube. Protein concentration was determined by absorbance measurements at 280 nm using a Nanodrop ND-1000 spectrophotometer (Labtech). Equal amounts of cell extracts (100 µl at 1.5 µg/µl) were prepared for all samples. 100 µM Biotin-Alkyne (Sigma-Aldrich; #764213), 1 mM TCEP (Sigma-Aldrich; #C4706), 100 µM THPTA (Sigma-Aldrich; #762342) and 1 mM CuSO<sub>4</sub> (Sigma-Aldrich; #C8027) were sequentially added to the protein samples, followed by a short vortex for several seconds. The

mixture was incubated at room temperature for 1 h. After that, 4 volumes of methanol were used to precipitate proteins. The samples were stored overnight at –20°C and then, the protein pellet was collected by centrifugation (20 000 × g, 4°C, 30 min) and dried at room temperature. The protein pellet was next denatured by adding 2 × loading buffer (40 µl per 100 µl mixture) followed by heating at 95°C for 5 min. The samples were separated by SDS-PAGE (12%) and analyzed by western blotting using HRP-conjugated streptavidin at 4°C for overnight. Newly-synthesized proteins were visualized by ECL incubation and exposure. After stripping, the membranes were incubated with anti-GAPDH or anti-α-tubulin for normalization (Supplementary Table S3).

#### Stably transfected HEK293 Flp-in cell lines

To generate stably transfected cell lines for the tetracycline-inducible expression of N-terminally 2 × Flag-His<sub>6</sub> or C-terminally His<sub>6</sub>-2 × Flag tagged THUMPD2, the HEK293 Flp-In T-Rex system (Invitrogen) was used. HEK293 Flp-In cells seeded 24 h prior were transfected with appropriate pcDNA5-based plasmids as well as a plasmid for expression of the Flp recombinase using X-tremeGENE 9 DNA transfection reagent (Roche) and transfected cells were selected with hygromycin B and blasticidin S according to the manufacturer's instructions. Expression of the tagged proteins was induced by treatment with 1 µg/µl tetracycline for 24 h before harvesting.

#### RNA-IP after cross-linking

To detect RNAs directly bound by THUMPD2, HEK293 cells expressing THUMPD2-His<sub>6</sub>-2 × Flag, 2 × Flag-His<sub>6</sub>-THUMPD2 or the His<sub>6</sub>-2 × FLAG tag were irradiated with UV light at 254 nm 3 × at 800 mJ/cm<sup>2</sup> to covalently crosslink proteins to their associated RNAs. Cells were lysed in a buffer containing 50 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.1% NP-40, 5 mM 2-mercaptoethanol and protease inhibitors (Roche) by sonication. Protein-RNA complexes were affinity purified on anti-Flag magnetic beads (Sigma-Aldrich) then eluted overnight with 3 × Flag peptide (Sigma-Aldrich). RNA-protein complexes were then immobilized on Ni-NTA (Qiagen) in denaturing conditions (6 M guanidium-HCl) before elution of RNAs using Proteinase K. RNAs were extracted from input and eluate samples using phenol-chloroform, separated by denaturing PAGE and analyzed by northern blotting using probes listed in Supplementary Table S4.

#### Purification of the U6 snRNA

Parental and THUMPD2 KO1 HCT116 cells were lysed in a buffer containing 10 mM Tris-HCl pH 8.4, 140 mM NaCl, 1.5 mM MgCl<sub>2</sub>, 0.5 mM EDTA, 0.5 mM DTT, 0.5% NP40 on ice for 3 min. Nuclei were pelleted by centrifugation at 5900 × g for 5 min at 4°C and the supernatant discarded. The nuclear pellet was resuspended in lysis buffer supplemented with 1.5 mM CaCl<sub>2</sub> and the centrifugation step was repeated. The supernatant was removed and nuclear RNA extracted from the pellet using TRI reagent (Sigma-Aldrich) according to the manufacturer's instructions. Nuclear small RNAs (<200 nt) were then enriched using the mirVANA kit (ThermoFischer Scientific) according to the manufacturer's guidelines. To specifically enrich the U6 snRNA, hydrophilic streptavidin magnetic beads (NEB) were

washed in 10 mM Tris-HCl pH 7.4 and incubated with 200  $\mu$ M biotinylated antisense DNA oligonucleotide (5'-Biotin-TTTTAGTATATGTGCTGCCGAAGCGAGCAC-3') for 90 min at room temperature. The beads were washed five times with 10 mM Tris-HCl pH 7.4 and resuspended in 6  $\times$  NTE buffer (120 mM Tris-HCl pH 7.4, 1.2 M NaCl, 15 mM EDTA). The extracted small nuclear RNAs were added to the oligonucleotide-conjugated beads and incubated at 70°C for 30 min before slowly cooling to room temperature. The supernatant was removed and the beads washed sequentially once with 1  $\times$  NTE buffer, once with 3  $\times$  NTE buffer and twice with 1  $\times$  NTE. RNAs were then eluted using 0.1  $\times$  NTE buffer at 70°C for 15 min. RNAs were precipitated using ethanol and resuspended in RNase-free water. The purified U6 snRNA was digested into nucleosides as described above. For m<sup>2</sup>G detection by LC-HRMS, the same protocol as described above was performed.

### mRNA-seq

Total RNA was extracted from wild-type (WT) HCT116, THUMP2 KO1 and KO2 cells using TRI reagent (Sigma-Aldrich) according to the manufacturer's instructions. A DNase digest was performed using TURBO DNase (Ambion) for 15 min at 37°C and RNAs were purified using the RNA clean and concentrator kit (Zymo) following the manufacturer's protocol. Polyadenylated RNAs were enriched and a cDNA library prepared using the TrueSeq Stranded Total RNA kit (Illumina). 50 nucleotide, unstranded, single-end sequencing was performed using the TruSeq SBS Kit v3 on a HiSeq4000. Library preparation, quality control and sequencing were performed at the NGS Integrative Genomics (NIG) Core Unit of the University Medical Center Göttingen (UMG).

### mRNA-seq mapping and expression quantification

FASTQC from the FASTX toolkit was run to ensure high quality of the data before mapping ([http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html)). For mapping, the aligner STAR v2.7.10a with the following settings was used: `-outFilterMultimapNmax 10 -outSAMattributes All -outSAMtype BAM SortedByCoordinate -outReadsUnmapped Fastx -chimSegmentMin 20 -chimOutType WithinBAM Junctions -quantMode TranscriptomeSAM GeneCounts -outWigType bedGraph -outWigNorm RPM -outWigStrand Unstranded -outFilterMismatchNmax 2` (66,67). UCSC Genome Browser tools `bedSort` and `bedGraphToBigWig` were used to convert from `bedgraph` format to `bigwig` format for visualization of the data in the IGV genome browser (68,69). Genome assembly version GRCh38.p13 and the GENCODE gene set from release 39 were used and gene, exon, intron and transcript annotations from `ensembl`, `entrez` and `hgnc` were obtained with the Bioconductor package `biomaRt` (70,71). SnoRNA annotations were retrieved from the GENCODE annotation `gtf` file. Genes containing introns spliced by the minor spliceosomes were retrieved from the Minor Intron Database (MIDB; (72); Olthof *et al.*; <https://doi.org/10.1101/2022.10.12.511939>).

To assess the presence of the single nucleotide insertion in all replicates of THUMP2 KO1, we used `sam-` and `bedtools` to extract all reads overlapping `chr2:39761363-39761365` (66,67). The fraction of reads with an insertion and

overlapping the region of interest is given in Supplementary Figure S7C. To ease analysis soft-clipped reads were excluded from the quantification in WT and KO1.

Gene expression was quantified using the `DESeqDataSetFromHTSeqCount`, `DESeq` and results functions from the `DESeq2` package in Bioconductor (73). Read counts per gene were obtained during the STAR mapping step. Only genes with 50 or more reads were kept for downstream analysis and plotting. Genomic regions with an FDR < 0.05 and a log<sub>2</sub>-fold change greater or smaller than 0.5 were used for the differential expression comparison in Figure 5A.

### Alternative splicing analysis with rMATS and FRASER

For exon-centric analysis of alternative splicing changes, `rMATS.4.0.2` was used with default settings for unstranded single-end RNA-seq data and at least a median of 5 reads per splicing event across replicates was required to consider events for further analysis (Supplementary Table S10). Intron-focused alternative splicing was performed in the software package `FRASER` (74,75) (see Supplementary data and Supplementary Tables S11 and S12). The significance of splice site usage determined by `FRASER` between the WT, THUMP2 KO1 and KO2 were assessed using the two-tailed Student's t-Test for splice sites that had coverage of 20 or more reads in all replicates.

For the analysis of intron retention, only introns that had only one reported 5' and 3' splice site in our data were considered. Splice site scores were calculated with the `MaxEnt` software (76). Polypyrimidine tract length was computed in an interval of 50 nt upstream of 3' splice sites. To this end, the fraction of C or U nucleotides (PY-fraction) in sliding 5 nt windows across the 50 nt interval was calculated. The distance in nt from the nucleotide position where the PY-fraction increased to 0.5 or higher and stayed high to the end of the intron was taken as PY-tract length.

The modified Z-score was calculated to transform intron feature data onto a uniform scale for visualization, using the following formula (77):

$$z_i = \frac{x_i - \bar{x}}{MAD} \quad (1)$$

with  $x_i$  being the median of the intron feature in group  $i$ ,  $\bar{x}$  being the intron feature median for all introns and `MAD` the median of absolute deviation.

$$MAD = \text{median}(|x_i - m|) \quad (2)$$

with  $m$  being all intron feature values.

To test for statistical significance, we performed the Wilcoxon-rank sum test between groups that showed increased (up) or decreased (down) intron retention (based on `theta3`, Figure 5G, column 3 in the *P*-value heatmap)/alternative splice site usage (based on `psi3`, Supplementary Figure S9E, column 3 in the *P*-value heatmap). We also tested for significance between the not significantly different group of splice sites and the up or down group, respectively (columns 1 & 2 in the *P*-value heatmaps in Figure 5G and Supplementary Figure S9E).

All mRNA-seq data analysis was done using bash scripts, software as cited above and `Rstudio` (2022.02.0 + 443). The R packages `pheatmap` and `ggplot2` were used for data visualization. For `sashimi` plot visualization, the command-



line tool ggsashimi was used (78). The high-performance cluster SCC maintained by the GWDG at the University of Göttingen was used for mapping, rMATS analysis and file format conversions.

### Validation of mRNA-seq data by RT-PCR

250 ng of purified, DNase-treated total RNA reverse transcribed with 50 pmol oligo-dT primer (5'-TTTTTTTTTTTTTTTTTTTTTTVN-3'; V = G/A/C, N = G/A/C/T) and Superscript III reverse transcriptase according to the manufacturer's instructions. The resultant cDNA (diluted 1:2.5) was used as a template for PCR with Phusion polymerase and dedicated oligonucleotide (Supplementary Table S4). The number of cycles was titrated to ensure exponential amplification of both products for quantification and 25 cycles yielded the best results for the three tested genes. PCR products were separated by agarose gel electrophoresis and detected using SybrGold Nucleic Acid Gel Stain (ThermoFisher Scientific). Intensity quantification of individual bands was done with Image Studio Lite.

## RESULTS

### Identification of the human TRMT112 interaction network by proximity labelling

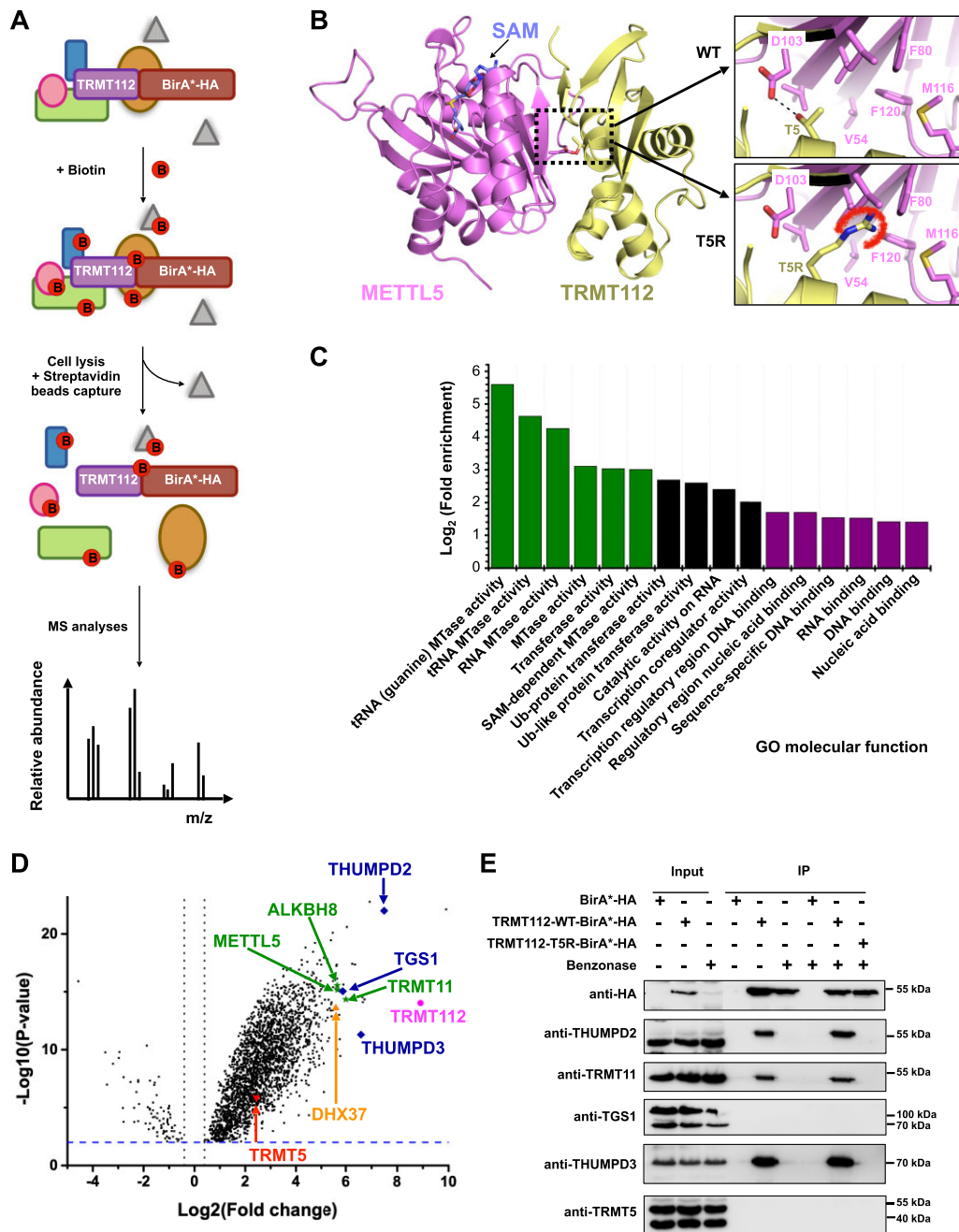
Recently, an immunoprecipitation approach coupled to stable isotope labelling by amino acids in cell culture (SILAC) was used to probe the MTase interactome of human TRMT112 (51). It is possible that MTase interaction partners with low expression levels and other transient or weak interactors are not captured by such affinity-based techniques. Therefore, to identify proteins interacting with human TRMT112 in an unbiased way in intact cells, we used the BioID (proximity-dependent biotin identification) approach coupled with mass spectrometry (MS) (Figure 1A; (79)). As the crystal structures of the human TRMT112-METTL5 (46) and TRMT112-HEMK2 (30) complexes show that the N-terminus of TRMT112 is directly engaged in the interaction with its MTase partners, the BirA\* domain endowed with biotin ligase activity was fused to the C-terminus of WT TRMT112 to generate the TRMT112-WT-BirA\*-HA construct. Importantly, BirA\*-HA alone and eGFP-BirA\*-HA were included as negative controls to eliminate background in the list of identified proteins. In addition, a TRMT112-T5R-BirA\*-HA construct, designed based on the direct involvement of threonine 5 side chain in interaction with HEK2 or METTL5 (Figure 1B), was included. As all available crystal structures of TRMT112-MTase complexes show that the MTases interact in a very similar manner with TRMT112 (17,20), the substitution of T5 by the longer and bulkier arginine side chain (T5R) was anticipated to generate steric hindrance and prevent TRMT112 interaction with MTase partners. The effectiveness of this amino acid substitution in preventing association of TRMT112 with METTL5 was confirmed by co-IP (Supplementary Figure S1A), confirming that the T5R mutant can serve as an ideal negative control for the identification of TRMT112-associated proteins.

Human HEK293T cells were transfected with appropriate amounts of each of the plasmids expressing either BirA\*-HA, eGFP-BirA\*-HA, TRMT112-WT-BirA\*-HA or the mutant TRMT112-T5R-BirA\*-HA to transiently express

these four protein constructs at approximately the same level (Supplementary Figure S1B). Each condition was performed in five replicates to allow robust statistical analyses, and proteins enriched by an affinity purification step using streptavidin beads were identified by MS. A total of 2281 proteins (Fold change (FC) > 2 and *P*-value < 0.001) were specifically enriched with the TRMT112-WT-BirA\*-HA protein compared to the three control conditions (Supplementary Table S5). Gene ontology 'molecular functions' analysis for the 100 proteins with the highest FCs revealed strong enrichment of proteins annotated as SAM-dependent MTases, in particular RNA MTases as well as proteins involved in nucleic acid binding (Figure 1C). This is in line with the known function of TRMT112 as a central activator of RNA-modifying MTases. Six MTases were identified among the 100 most enriched proteins, including ALKBH8 and METTL5 that are established TRMT112 partners (Figure 1D; (36,37,46)). TRMT11, the orthologue of yeast and archaeal Trm112-interacting Trm11 proteins, was identified, suggesting a conserved interaction (21,31,40,44). THUMP3, a recently identified tRNA m<sup>2</sup>G MTase and an uncharacterized putative MTase, THUMP2, were also recovered. Interestingly, TGS1 (Trimethyl guanosine synthetase 1), an MTase modifying the 5' end cap of short non-coding RNAs involved in pre-mRNA splicing (U snRNAs) or ribosome biogenesis (snoRNAs; (80)) was also found to be enriched. Notably, two well-established TRMT112 MTase partners (HEMK2 and BUD23) were not identified by this BioID experiment (30,81). However, DHX37, a DEAH helicase involved in the release of the U3 snoRNP from pre-ribosomal particles and that is orthologous to yeast Dhr1, a physical partner of the Bud23-Trm112 complex (23,82-84), was found among the most enriched proteins suggesting that BUD23 and TRMT112 likely interact in these conditions but that BUD23 is resistant to biotinylation for some yet unknown reasons.

To confirm these results, TRMT112-WT-BirA\*-HA or TRMT112-T5R-BirA\*-HA were expressed in human cells in the absence of biotin and complexes immunoprecipitated via the TRMT112 HA tag in the presence or absence of the benzonase nuclease. The presence of the MTases identified by proximity labelling in the immunoprecipitated fraction was determined by western blotting (Figure 1E). With the exception of TGS1 as well as TRMT5, an MTase present in the enriched proteins but not among the most enriched, all the MTases tested (TRMT11, THUMP3 and THUMP2) co-precipitated with WT TRMT112 but not with the T5R mutant, indicating a specific interaction. These interactions were not affected by nuclease treatment indicating that they are not bridged by nucleic acids (Figure 1E). In parallel, we performed a reverse co-IP experiment using 3 × Flag-THUMP2 or -THUMP3 as bait and detected TRMT112 in the IP fractions, confirming the interaction of TRMT112 with THUMP2 and THUMP3 (Supplementary Figure S1C).

Consistent with a study published during the preparation of this article (51), these results suggest that the MTase interactome of TRMT112 comprises at least seven RNA MTases, including the recently characterized tRNA MTase THUMP3 (52) and the currently uncharacterized THUMP2. Our proximity labelling data are also a useful resource for learning more about the functions and cellular context of TRMT112-MTase complexes and their regulation (see Discussion).



**Figure 1.** Human TRMT112 is the hub of an MTases network. **(A)** Schematic representation of the BiID-MS technique used to characterize the interaction network of human TRMT112 protein. Proteins directly or indirectly interacting with TRMT112 are depicted as rectangles and spheres, respectively. Neighboring proteins are depicted as triangles. **(B)** Representation of the 3D structure of the human METTL5-TRMT112 complex (PDB code: 6H2V; (46)). The environment of the Thr5 from TRMT112 in the complex with METTL5 is shown in the upper inset. The black dotted line represents the hydrogen bond formed by the hydroxyl group of the side chain of Thr5 and the carboxylate group of the side chain of Asp103 from METTL5. Lower inset: Model of the T5R TRMT112 mutant highlighting the steric clashes (red curved line) likely to be generated by the Arg side chain with METTL5 amino acids. **(C)** Analysis of the ‘Molecular function’ Gene Ontology (GO) terms of the 100 most enriched proteins identified in our BiID experiment. GO analysis was performed using the Panther classification system as implemented in the Gene Ontology resource (<http://geneontology.org/>; (116–118)). Statistical analysis was performed using the Fisher’s exact test and the false discovery rate (FDR). Only GO terms with fold enrichments greater than 2.5 ( $\log_2$  value higher than 1.4) are shown. MTases and nucleic acid binding terms are colored green and purple, respectively. **(D)** Volcano plot of proteins differentially purified by TRMT112-WT-BirA\*-HA vs controls (TRMT112-T5R-BirA\*-HA, eGFP-BirA\*-HA and BirA\*-HA; adjusted  $P$ -value < 0.05). TRMT112 is shown in pink. Previously known partners of human TRMT112 protein (based on experimental evidence or similarities with yeast proteins) are indicated in green. Other highly enriched MTases are shown in blue. TRMT5, a SAM-dependent class I MTase not found among the most enriched proteins, is shown in red. The blue dotted line indicates a  $-\log_{10}(P\text{-value})$  of 2. The black dotted lines correspond to fold changes higher than  $-1.32$  or  $1.32$ . Statistics were calculated from five biologically independent experiments. **(E)** Co-immunoprecipitation experiments confirming that TRMT11, THUMPD2 and THUMPD3 interact with TRMT112. The TRMT112-WT-BirA\*-HA, TRMT112-T5R-BirA\*-HA and BirA\*-HA (control) constructs were expressed in HEK293T cells in the absence of biotin and purified using anti-HA beads. Co-immunoprecipitated proteins were detected by western blot using the indicated antibodies. Co-IP experiments were performed in the absence or presence of benzonase to exclude a role of nucleic acids in these interactions. Three biologically independent experiments were performed and a representative image is shown.

## The TRMT11-TRMT112 and THUMP3-TRMT112 complexes are m<sup>2</sup>G tRNA MTases

TRMT11 and THUMP3 both contain an N-terminal THUMP domain and a C-terminal SAM-dependent MTase domain (Figure 2A; (42,44,85)). In tRNA modifying enzymes, the THUMP domain is predicted to bind to the tRNA 'CCA' tail via a large positively charged region (86), while the catalytic domain modifies the base of a nucleotide located in the tRNA amino acyl acceptor stem loop. The MTase domains of these enzymes contain a highly conserved signature (Figure 2A) characteristic of enzymes methylating planar amino groups on nucleic acids, such as m<sup>2</sup>G (87,88). Based on its high similarity to *S. cerevisiae* Trm11 (34% and 54% sequence identity and similarity, respectively), human TRMT11 is anticipated to be responsible for m<sup>2</sup>G formation at position 10, but this has not yet been demonstrated (89). During the course of this study, THUMP3, which has no yeast orthologue, was identified as the enzyme installing m<sup>2</sup>G at positions 6/7 of many metazoan tRNAs (52).

To analyze the functions of these two complexes, we co-expressed TRMT112 with TRMT11 or THUMP3 and purified the stable heterodimeric complexes (Supplementary Figure S2; Supplementary Table S6). In agreement with Yang and colleagues (52), we confirmed that the THUMP3-TRMT112 complex indeed acts as an m<sup>2</sup>G tRNA MTase targeting position 6 of an *E. coli* tRNA<sub>i</sub><sup>Met</sup> substrate using a combination of *in vitro* biochemical assays and mass spectrometry approaches (LC/MS and MALDI-MS/MS; see Supplementary text for details about the MALDI-MS/MS mapping of m<sup>2</sup>G on tRNA; Supplementary Figures S3A–D and S4A–C). Using the same approaches, we have characterized for the first time the enzymatic activity of the human TRMT11–TRMT112 complex. Similarly to its yeast and archaeal orthologues (31,40,44), LC/MS and MALDI-MS/MS data indicate that this complex catalyzes the formation of m<sup>2</sup>G at position 10 of tRNAs (Supplementary Figures S3A–D and S4D–F and Supplementary text).

To further explore the molecular functions of these m<sup>2</sup>G tRNA MTases in human cells, we generated HCT116 cell lines lacking TRMT11 or THUMP3 using the CRISPR-Cas9 gene knockout technique. For each cell line, two different cell populations in which TRMT11 or THUMP3 could not be detected in the total cell extract by western blot were obtained (KO1 and KO2; Figure 2B). In the TRMT11 KO2 cell line, the *THUMP3* gene was further edited to create the TRMT11/THUMP3 double KO1 and KO2 cell lines (Figure 2B). Sequencing of the targeted loci in these cell lines revealed single nucleotide insertions in these genes, resulting in translation frameshifts (Supplementary Figure S5A–C and Supplementary text for more details). We then selected one of the each of the KO cell lines, purified total tRNAs, and quantified the levels of m<sup>2</sup>G and m<sup>2,2</sup>G by LC-HRMS after digestion into nucleosides (Figure 2C). Compared to parental cell lines, lack of TRMT11 or THUMP3 resulted in ~40% and ~20% decreases in the m<sup>2</sup>G level in total tRNAs, respectively. However, concomitant inactivation of TRMT11 and THUMP3 resulted in an 85% decrease in m<sup>2</sup>G levels. These data thus confirm that human TRMT11 and THUMP3 are the major m<sup>2</sup>G tRNA MTases. The residual m<sup>2</sup>G observed in the TRMT11/THUMP3 double KO1 cell line may be explained by the presence of m<sup>2</sup>G<sub>26</sub> in tRNA<sup>Val</sup><sub>CAC/Met</sub><sup>CAU</sup>, m<sup>2</sup>G in mitochondrial tRNAs and/or

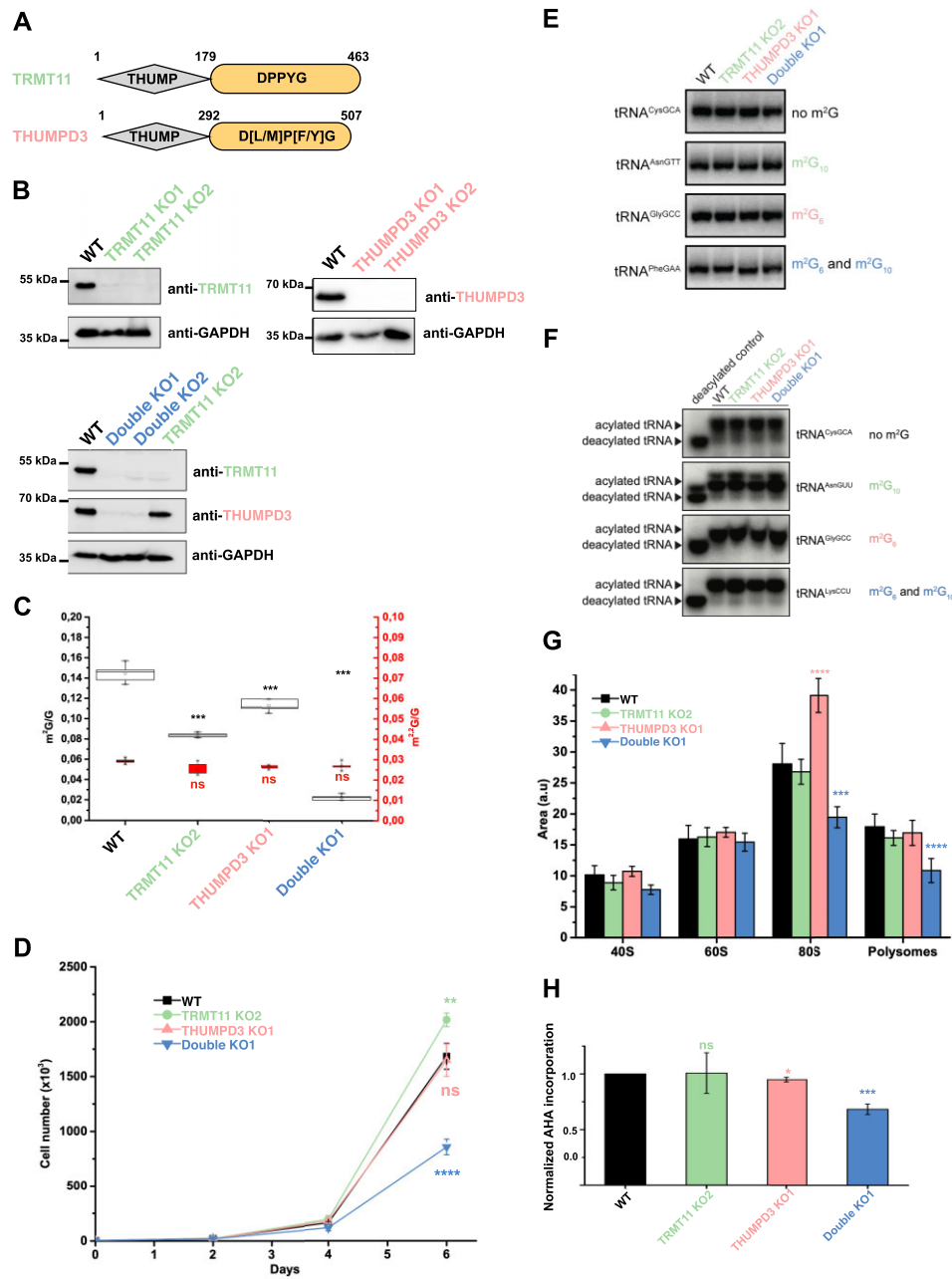
intermediates generated during the formation of m<sup>2,2</sup>G at position 26 of some tRNAs by TRMT1. Notably, m<sup>2,2</sup>G levels are not affected across these different HCT116 cell lines (Figure 2C). This confirms that, 1) TRMT11 and THUMP3 do not catalyze the formation of m<sup>2,2</sup>G on tRNAs and 2) lack of m<sup>2</sup>G<sub>6/7/10</sub> does not influence the formation of m<sup>2,2</sup>G at position 26 in tRNAs.

## The concomitant deletion of both m<sup>2</sup>G tRNA MTases affects cell proliferation and protein synthesis without significantly influencing tRNA stability, folding or aminoacylation

The molecular functions of m<sup>2</sup>G modifications in tRNAs remain unknown. Comparing the growth rate of the TRMT11, THUMP3 and double KO cell lines demonstrated that lack of THUMP3 or TRMT11 did not, or only mildly, affected the growth of HCT116 cells, respectively (Figure 2D). However, lack of both TRMT11 and THUMP3 strongly reduced cell proliferation, revealing that the concomitant presence of m<sup>2</sup>G modifications at tRNA positions 6, 7 and 10 is biologically important. Consistent with this, the double KO cell line formed markedly smaller colonies on soft agar, confirming the profound proliferation defects of this cell line devoid of m<sup>2</sup>G modifications in the acceptor stem/D-arm (Supplementary Figure S5D).

As presence of modified nucleotides can influence tRNA stability or folding (90,91), tRNA levels in the KO cell lines were first comprehensively determined using the nrStar™ human/mouse tRNA PCR array service from Arraystar Inc. (Rockville, MD 20850, USA; Supplementary Figure S6A–C and Supplementary Table S7). No significant alterations in the levels of tRNAs lacking m<sup>2</sup>G<sub>6/7/10</sub> modifications were observed, indicating that they are not targeted for degradation. tRNA modifications can also influence folding (92) so RNAs extracted from the WT and KO cell lines were separated by native polyacrylamide gel electrophoresis under conditions that allow modification-dependent alterations in tRNA structure to be detected (Supplementary Figure S6D; (92)) and selected tRNAs detected by northern blotting. However, the migration patterns of m<sup>2</sup>G<sub>6-</sub>, m<sup>2</sup>G<sub>10-</sub>, m<sup>2</sup>G<sub>6+10-</sub> and non-m<sup>2</sup>G-containing tRNAs did not differ, suggesting that m<sup>2</sup>G modifications do not strongly influence the structure of these tRNAs (Figure 2E). To fulfill their role as adaptors between mRNAs and nascent peptides in the context of translation, tRNAs are aminoacylated and, in some cases, tRNA modifications can influence aminoacylation (93,94). Aminoacylated tRNAs were purified from WT and KO cells by RNA extraction under acidic conditions, which do not disrupt covalently bound amino acids. RNAs were then separated by acidic polyacrylamide gel electrophoresis alongside control samples that had been deacylated by alkaline treatment. Detection of m<sup>2</sup>G<sub>6-</sub>, m<sup>2</sup>G<sub>10-</sub>, m<sup>2</sup>G<sub>6+10-</sub> and non-m<sup>2</sup>G-containing tRNAs (tRNA<sup>Gly</sup><sub>GCC</sub>, tRNA<sup>Asn</sup><sub>GUU</sub>, tRNA<sup>Lys</sup><sub>CCU</sub>, tRNA<sup>Cys</sup><sub>GCA</sub>, respectively) revealed that these tRNAs are efficiently aminoacylated in the presence and absence of m<sup>2</sup>G<sub>6/10</sub> modifications (Figure 2F).

To explore further whether m<sup>2</sup>G tRNA modifications are required for protein synthesis, the effect of lack of the relevant MTases on ribosome and polysome levels was determined. As



**Figure 2.** tRNA m<sup>2</sup>G modifications are important for proliferation of HCT116 colorectal cancer cells and optimal translation. **(A)** Schematic representation of TRMT11 and THUMP3 showing their domain organization. The predicted amino acid boundaries of the different domains are shown above each diagram. The MTase domains are colored in orange. The strongly conserved signature found in the MTase domains, which is assumed to coordinate the substrate into the MTase active site, is indicated. **(B)** Western blot analyses of HCT116 TRMT11 (top left panel), THUMP3 (top right panel) or double (lower panel) knockout cell lines. Proteins of interest were detected by western blot using the indicated antibodies. Three biologically independent experiments were performed and a representative image is shown. **(C)** Quantification by LC–HRMS of m<sup>2</sup>G (left black axis) and m<sup>2,2</sup>G (right red axis) levels in total tRNAs purified from the indicated cell lines. Mean values calculated from five replicates are shown and error bars represent standard deviation. **(D)** Growth analysis of WT, TRMT11, THUMP3 or TRMT11/THUMP3 KO HCT116 cell lines. Cell numbers were monitored every 24 h for 6 days. Three biologically independent experiments were performed and error bars represent standard deviations. **(E)** Total RNAs extracted from the indicated cell lines were separated by native polyacrylamide gel electrophoresis and analyzed by northern blotting using probes hybridizing to the tRNAs indicated on the left. The presence and position of m<sup>2</sup>G in the detected tRNAs is noted on the right. Three biologically independent experiments were performed and a representative image is shown. **(F)** Total RNAs were extracted from the indicated cell lines under acidic conditions. A sample subjected to alkaline treatment served as a deacylated control. Samples were separated by denaturing polyacrylamide gel electrophoresis under acidic conditions before analysis by northern blotting using probes hybridizing to the tRNAs indicated on the right. The presence and position of m<sup>2</sup>G in the detected tRNAs is also given. Three biologically independent experiments were performed and a representative image is shown. **(G)** Polysome profile analyses using the indicated cell lines were performed. For each cell line, the mean areas of each of the ribosomal complexes (small subunit – 40S, large subunit – 60S, monosomes – 80S and polysomes) observed in the polysome profile analyses for each cell line is shown. Non-statistically significant differences are not indicated for the sake of clarity. Five biologically independent experiments were performed and error bars represent standard deviations. **(H)** Effect of TRMT11 and/or THUMP3 depletion on overall protein synthesis. For TRMT11 KO2 and TRMT11/THUMP3 KO1 cell lines, the GAPDH signal was used to normalize the data, whereas for the THUMP3 KO1 cell line, we used the alpha-tubulin signal as GAPDH levels are affected in this cell line. Three biologically independent experiments were performed and error bars represent standard deviations.

anticipated for tRNA MTases, polysome profile analyses using sucrose gradients revealed no changes in the levels of 40S and 60S subunits in any of the cell lines (Figure 2G; Supplementary Figure S6E). Although no changes in monosome (80S) or polysome levels were detected in cells lacking TRMT11, the absence of THUMP3 caused a significant accumulation of monosomes (Figure 2G; Supplementary Figure S6E), potentially indicating stalling of translation. Strikingly, the numbers of monosomes (80S) and polysomes were reduced in the TRMT11/THUMP3 double KO1 cell line (Figure 2G; Supplementary Figure S6E), suggesting that translation is affected in this double KO cell line. Impaired translation will manifest in reduced protein synthesis so to analyze nascent protein production in cells lacking m<sup>2</sup>G<sub>6/7/10</sub>, the incorporation of azidohomoalanine (or AHA), a methionine analogue that can be labeled by biotin using click-chemistry, into newly synthesized proteins was monitored. This confirmed that protein synthesis is significantly reduced (~30% decrease) in the TRMT11/THUMP3 double KO1 cell line, while it is unaffected in the TRMT11 KO2 cell line and only slightly reduced in the THUMP3 KO1 cell line (Figure 2H).

Altogether, these data reveal that individually, lack of either the TRMT11 and THUMP3 m<sup>2</sup>G tRNA MTases does not strongly affect either cell proliferation or mRNA translation. However, their concomitant absence significantly reduces both cell growth and protein synthesis, demonstrating the combined importance of these two tRNA modifications for eukaryotic cells.

### THUMP2 is an active m<sup>2</sup>G MTase with a small RNA substrate other than tRNAs

The remaining MTase found associated with TRMT112, THUMP2, has so far remained uncharacterized. Similar to TRMT11 and THUMP3, THUMP2 contains an N-terminal THUMP domain and a C-terminal SAM-dependent MTase domain with the signature characteristic of enzymes methylating planar amino groups on nucleic acids (Figure 3A; (87)). To gain insight into its function, THUMP2 knockout HCT116 cell lines were generated by CRISPR-Cas9 as previously done for the tRNA m<sup>2</sup>G MTases. Two different cell lines in which THUMP2 could not be detected in the total cell extract by western blot were obtained (THUMP2 KO1 and KO2; Figure 3B). Sequencing of the targeted locus in these cell lines revealed a single nucleotide insertion resulting in a translation frameshift and premature stop codon in the case of the KO1 cell line and a three-nucleotide insertion encoding a premature stop codon in the KO2 cell line (Supplementary Figure S7A and see Supplementary text for more details).

In parallel, despite many efforts, the recombinant human THUMP2–TRMT112 complex could not be purified because THUMP2 was not soluble under the different experimental conditions tested, either alone or when co-expressed with TRMT112. As other protein partners may be important for THUMP2 protein solubility, IP experiments of transiently expressed 3 × Flag-THUMP2 protein in HEK293T cells were performed, followed by MS analysis to identify associated proteins. However, the only protein significantly enriched with THUMP2 was TRMT112, implying that the latter protein is the only stable direct partner of THUMP2 (Supplementary Figure S8A and Supplementary Table S8). Therefore, we purified

the heterodimeric *Bos taurus* complex (BtTHUMP2–TRMT112) after co-expression in *E. coli* cells (Supplementary Figure S8B and Supplementary Table S6).

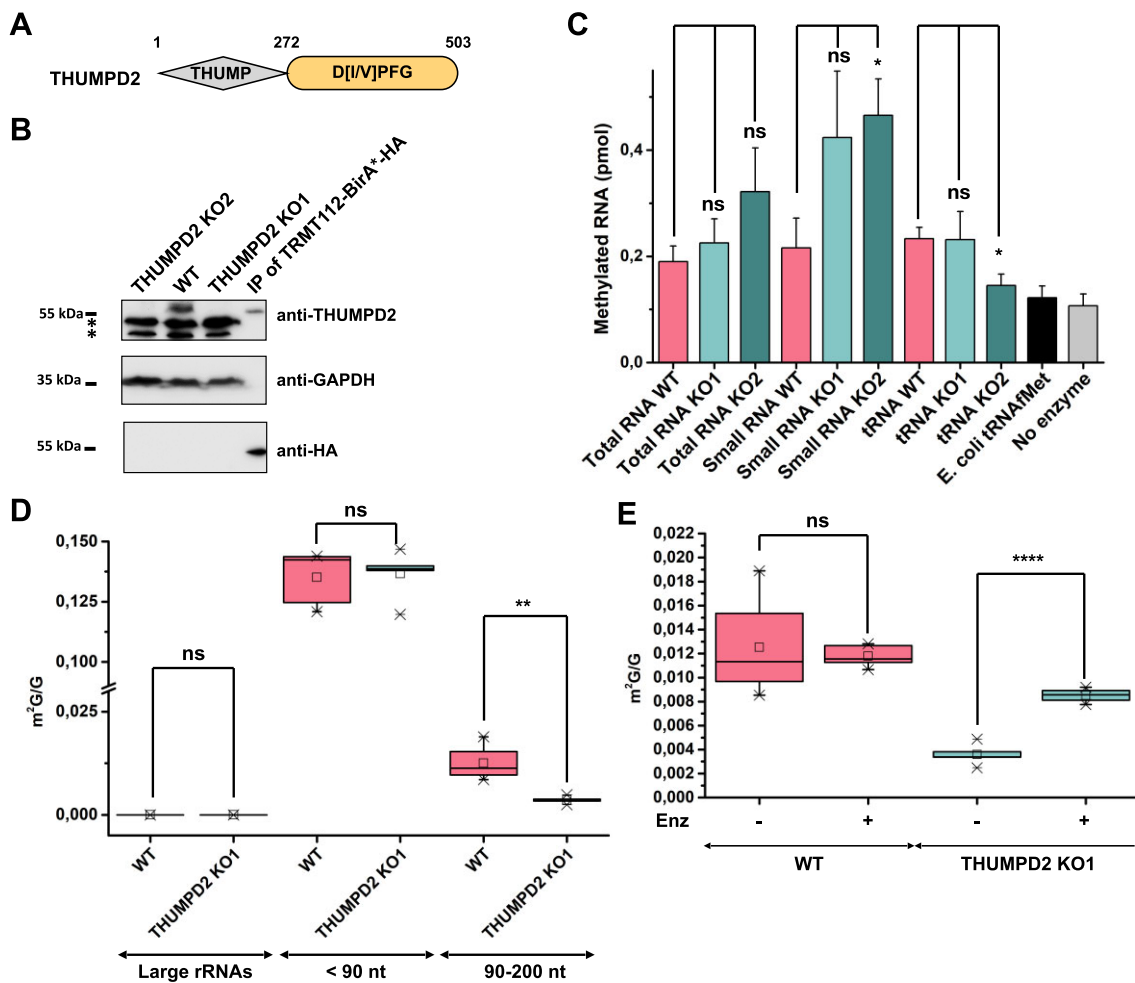
To determine whether THUMP2 is an active MTase and, if so, identify its targets, different RNA populations (total RNAs, small RNAs, or tRNAs; see Materials and Methods for preparation details) were purified from the WT and THUMP2 KO1 or KO2 cell lines. Assuming THUMP2 is an active MTase, RNAs from WT cells will already contain endogenously installed modifications while those from the THUMP2 KO cell lines will lack these modifications and therefore represent potential substrates for the recombinant BtTHUMP2–TRMT112 complex in *in vitro* enzymatic assays. Only minimal changes in the extent of methylation of total RNAs purified from the different cell lines were observed after *in vitro* methylation and no increases in tRNA methylation were detected in the THUMP2 KO cell lines compared to the WT (Figure 3C). However, significant MTase activity of the recombinant BtTHUMP2–TRMT112 complex was observed on the enriched population of small RNAs purified from the THUMP2 KO2 cell line compared to the WT, and methylation of small RNAs from the THUMP2 KO1 cell line were also notably increased (Figure 3C). This indicates that the THUMP2–TRMT112 complex has RNA MTase activity most likely targeting small RNAs that are not tRNAs.

In parallel, small RNAs enriched from the WT and THUMP2 KO1 cell lines were subjected to size exclusion chromatography to more precisely separate different RNA species based on their lengths: large rRNAs present as contaminants in the small RNA fraction, RNAs shorter than 90 nt (mostly tRNAs) and RNAs of intermediate size (90–200 nt; Supplementary Figure S8C, D). These different RNA pools were digested into nucleosides to quantify m<sup>2</sup>G levels by LC–HRMS. As expected, m<sup>2</sup>G was not detected in the fraction corresponding to large rRNAs (Figure 3D). Depletion of THUMP2 also did not affect m<sup>2</sup>G levels in the fraction comprising RNAs shorter than 90 nt compared with the parental cell line, confirming that THUMP2 is not an m<sup>2</sup>G tRNA MTase (52). However, while m<sup>2</sup>G was detected in the 90–200 nt long RNAs extracted from parental cell lines, it was almost undetectable in the equivalent RNA pool purified from the THUMP2 KO1 cell line (Figure 3D). Furthermore, incubation of the 90–200 nt long RNAs purified from the THUMP2 KO1 or parental cell lines with the recombinant BtTHUMP2–TRMT112 complex resulted in a specific increase in m<sup>2</sup>G levels only for RNAs extracted from the THUMP2 KO1 cell line (Figure 3E).

Taken together, these results reveal that THUMP2 is an active m<sup>2</sup>G MTase with a small RNA substrate of 90–200 nt.

### THUMP2 is the elusive m<sup>2</sup>G MTase targeting the U6 snRNA

The finding that THUMP2 targets small RNAs other than tRNAs is particularly interesting because an m<sup>2</sup>G modification was described in 1980 in mouse and rat U6 snRNAs, which are 108 nt in size (53,54). The U6 snRNA m<sup>2</sup>G modification was mapped at position 72, however, the enzyme responsible for introducing this modification is still unknown (5,53,54). The U6 snRNA plays a critical role in pre-mRNA splicing by dynamically interacting with other



**Figure 3.** THUMP2 is a small RNA  $m^2G$  methyltransferase. **(A)** Schematic representation of THUMP2 showing its domain organization. The predicted amino acid boundaries of the different domains are indicated. The MTase domain is colored in orange. The strongly conserved signature found in the MTase domain, which is assumed to coordinate the substrate into the MTase active site is indicated. **(B)** Western blot analysis of proteins from WT HCT116 and the THUMP2 KO1 and KO2 cell lines. Proteins of interest were detected using indicated antibodies. The GAPDH signal was used as a loading control. Due to the presence of two major contaminant bands (\*) detected by the anti-THUMP2 antibody around the THUMP2 expected molecular weight, the elution of a co-immunoprecipitation experiment performed with HA-tagged TRMT112 (right lane: IP of TRMT112-BirA\*-HA) was analyzed alongside to confirm that the band absent in the THUMP2 KO1 and KO2 samples corresponds to endogenous THUMP2. **(C)** *In vitro* enzymatic assays performed using the recombinant bovine THUMP2-TRMT112 complex and specific RNA populations enriched from WT HCT116 and the THUMP2 KO1 and KO2 cell lines as substrate. Error bars represent standard deviation calculated from three biological replicates. **(D)** Quantification by LC-HRMS of  $m^2G$  levels in different RNA species purified from the parental (WT) or THUMP2 KO1 HCT116 cell lines. Error bars represent standard deviation calculated from five biological replicates. **(E)** Quantification by LC-HRMS of  $m^2G$  formation upon *in vitro* incubation of small RNAs (100–150 nucleotides) purified from the parental (WT) or THUMP2 KO1 HCT116 cell lines, with recombinant bovine THUMP2-TRMT112 complex. Error bars represent standard deviation calculated from five biological replicates.

snRNAs and protein partners to form the catalytic center of the spliceosome (95,96). In mouse, the U6 snRNA  $m^2G_{72}$  modification is located two nucleotides upstream of  $U_{74}$ , which is essential for the formation of the catalytic U6 snRNA triplex and the binding of two  $Mg^{2+}$  ions, which are directly involved in splicing catalysis (Figure 4A-B; (97)). THUMP2 localizes to the nucleus (51) and we therefore investigated whether THUMP2 could be responsible for  $m^2G$  formation on the snRNA U6 by first ascertaining if this MTase physically associates with the U6 snRNA. Stably transfected cell lines expressing N-terminally 2xFlag-His<sub>6</sub>- or C-terminally His<sub>6</sub>-2xFlag-tagged versions of THUMP2 were generated. These cell lines were subjected to UV cross-linking to covalently attach proteins to bound RNAs and extracts were then used for RNA immunoprecipitation experiments under strongly denaturing conditions to identify RNAs directly bound by

this MTase. Northern blotting of input and eluate samples revealed that the U6 snRNA specifically interacts with THUMP2 while other snRNAs and the 5.8S rRNA do not (Figure 4C). Notably, the U6atac snRNA of the minor spliceosome, which also contains a G at an equivalent position but has not been reported to contain an  $m^2G$ , was not recovered in the RNA-IP eluate (Figure 4C), implying that THUMP2 specifically associates with the U6 snRNA of the major spliceosome.

To determine if  $m^2G$  is present in the human U6 snRNA and, if so, whether THUMP2 is necessary for this methylation, the U6 snRNA was purified from WT and THUMP2 KO1 and KO2 cell lines using a biotinylated complementary DNA oligonucleotide. The recovered RNA was digested into nucleosides, and analyzed for the presence of  $m^2G$  by LC-HRMS. The presence of  $m^2G$  in the

human U6 snRNA was confirmed in WT cells demonstrating the evolutionary conservation of this modification from mouse/rat to human (Figure 4D). However, in the U6 snRNA extracted from the THUMP2 KO1 and KO2 cell lines, we observed a specific loss of the m<sup>2</sup>G nucleoside (Figure 4D). These data demonstrate that the human THUMP2 MTase protein is necessary for m<sup>2</sup>G formation on U6 snRNA, an 'orphan' modification identified over four decades ago.

To explore the function of THUMP2 and the m<sup>2</sup>G modification for which it is required on the cellular level, growth of the THUMP2 KO1 and KO2 cell lines was analyzed. This revealed that lack of THUMP2 in the KO1 and KO2 cell lines significantly affects proliferation compared to the WT HCT116 cell line (Figure 4E), implying an important role for this protein and probably the modification it contributes to install. We also analyzed whether the absence of THUMP2 and the m<sup>2</sup>G on the U6 snRNA, affected the stability of this snRNA or others, but we did not detect a significant differences in the amounts of U6, U6atac, U2 or U4 snRNAs in THUMP2 KO1 or KO2 cell lines compared with the parental cell line (Figure 4F), implying that the function of this modification is not to stabilize the U6 snRNA.

### THUMP2 is important for alternative pre-mRNA splicing

As the U6 snRNA is a core component of the catalytic spliceosome (Figure 4A, B), we explored how the lack of m<sup>2</sup>G in U6 affects the transcriptome. To determine mRNA levels and detect alterations in splicing in a transcriptome-wide manner, mRNA-seq analysis was performed on WT and THUMP2 KO1 and KO2 cells. Triplicate samples were analyzed. Individual biological replicates showed good reproducibility and confirmed the genomic disruptions to the THUMP2 locus in the two KO cell lines (Supplementary Figure S7B–D; Supplementary Figure S9A). Global analysis of transcript levels showed several hundred transcripts with significant expression changes (Figure 5A, Supplementary Table S9). A significant portion of transcripts that showed downregulation in the THUMP2 KO1 cell line were also downregulated in the THUMP2 KO2 cell line, suggesting these changes represent a common phenotype arising due to lack of the m<sup>2</sup>G modification in the U6 snRNA (Chi-squared test, p-Value downregulated genes  $4 \times 10^{-17}$ , 47% of downregulated genes in THUMP2 KO2 common to THUMP2 KO1). Notably, the THUMP2 KO2 cell line displayed fewer gene expression changes than KO1 (Figure 5A). Taken together with the observation that the THUMP2 KO1 cell line has more affected proliferation than KO2 cell line (Figure 4E), this could suggest that the THUMP2 KO2 cell line buffers transcriptome alterations to some degree, thereby partially recovering proliferation of this cell line.

Alternative splicing, which can manifest as skipped exons, the use of alternative 5' or 3' splice sites, the incorporation of mutually exclusive exons or retained introns (Figure 5B), can lead to changes in transcript levels. Within the RNA-seq datasets, a number of alternative splicing events were detected as significantly different between the WT cells and those lacking THUMP2 (KO1 and KO2; Figure 5C, Supplementary Figure S9B–C, G–H, Supplementary Tables S10 and S11). The majority of alternative splicing events significant in one of the two knock-outs changed in the same direction in the second knock-out. Intron retention

was the most prevalent type of alternative splicing affected by the loss of THUMP2 when considering the relative proportion of significantly changed events compared to all events that passed the minimal read cutoff criteria. Two different ways to quantify alternative splice site usage and intron retention supported these conclusions (rMATS and FRASER, Supplementary text, Supplementary Figures S9B–C, G–H, Supplementary Tables S10–S12).

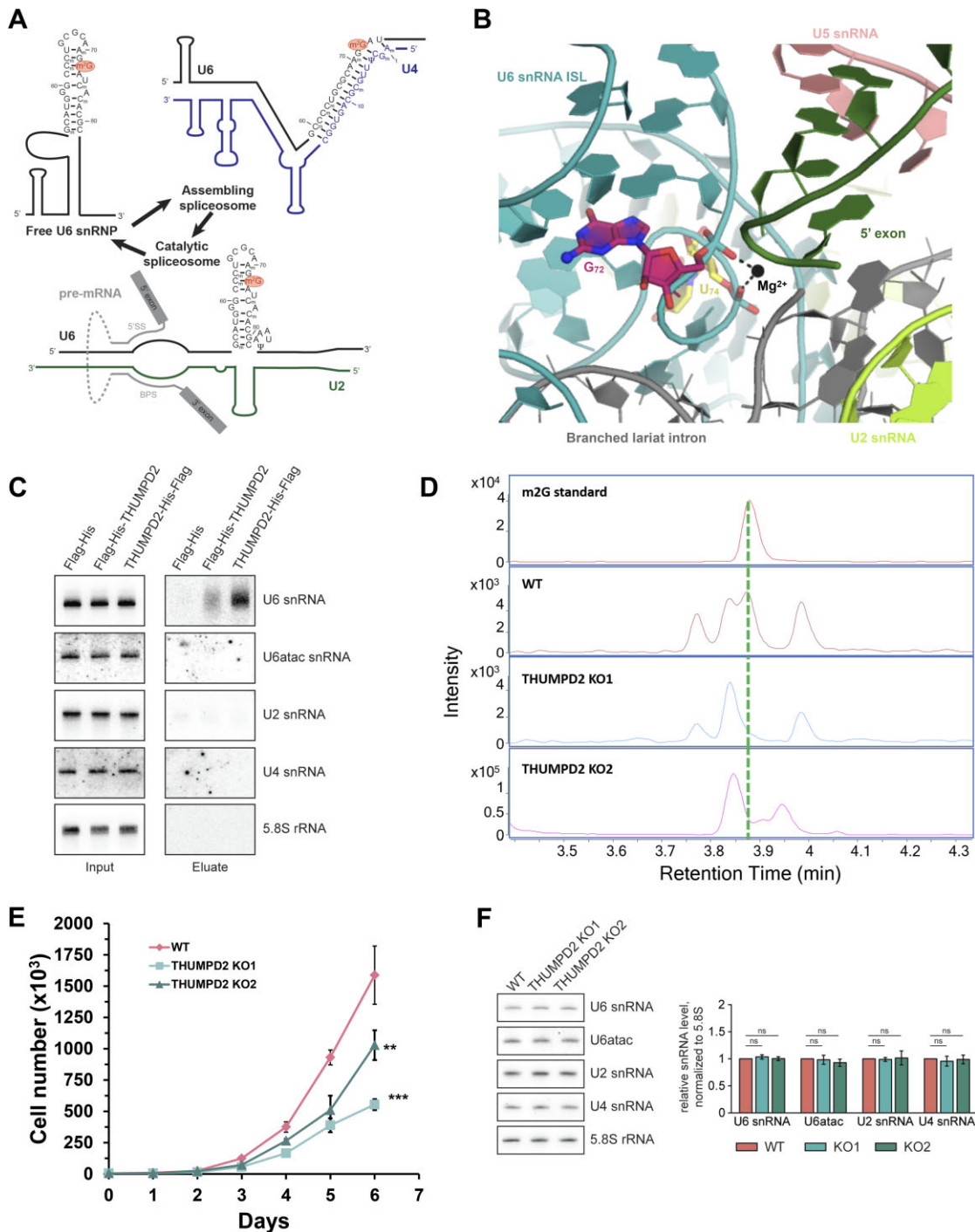
To verify the mRNA-seq based alternative splicing analysis further, selected alternative splicing events were monitored by RT-PCR. Total RNA extracted from WT HCT116 cells and the THUMP2 KO1 and KO2 cell lines was reverse transcribed from the poly(A) tail and the resultant cDNA served as a template for linear-range PCR with primers specific to detect the alternative 3' splice site usage for intron 29 of *TTC37*, retention of the second to last intron of *ANO9* and skipping of exon 3 of *RTF1*. In line with the mRNA-seq data (Figure 5D–F; sashimi plots, left panels; Supplementary Table S13), this confirmed the anticipated changes in transcript isoform levels in cells lacking THUMP2 (Figure 5D–F; right panels; Supplementary Table S13).

To gain further insights into how the absence of m<sup>2</sup>G in the U6 snRNAs in cells lacking THUMP2 might influence alternative pre-mRNA splicing, key features of the introns retained in THUMP2 KO cells were analyzed. Collectively, intron feature analysis revealed that introns retained in cells lacking THUMP2 had lower splice site quality, shorter polypyrimidine tracts and were generally shorter than introns that did not change significantly in intron retention (Figure 5G, Supplementary Figure S9F). Various small non-coding RNAs are encoded within pre-mRNA introns and another distinguishing feature of introns retained in cells lacking THUMP2 was the overrepresentation of those containing small nucleolar (sno)RNAs (Supplementary Figure S9I). Furthermore, gene expression was significantly reduced for genes that had two or more introns with increased intron retention compared to the group of genes with better intron removal (Figure 5H, Supplementary Figure S9F). Other feature analyses, e.g. for the preferential inclusion of 3' splice sites, did not yield strong trends (Supplementary Figure S9D, E).

Altogether, these data suggest that the lack of THUMP2-mediated m<sup>2</sup>G methylation of the U6 snRNA affects the constitutive splicing efficiency of introns that have suboptimal splice sites and can impact final mRNA levels.

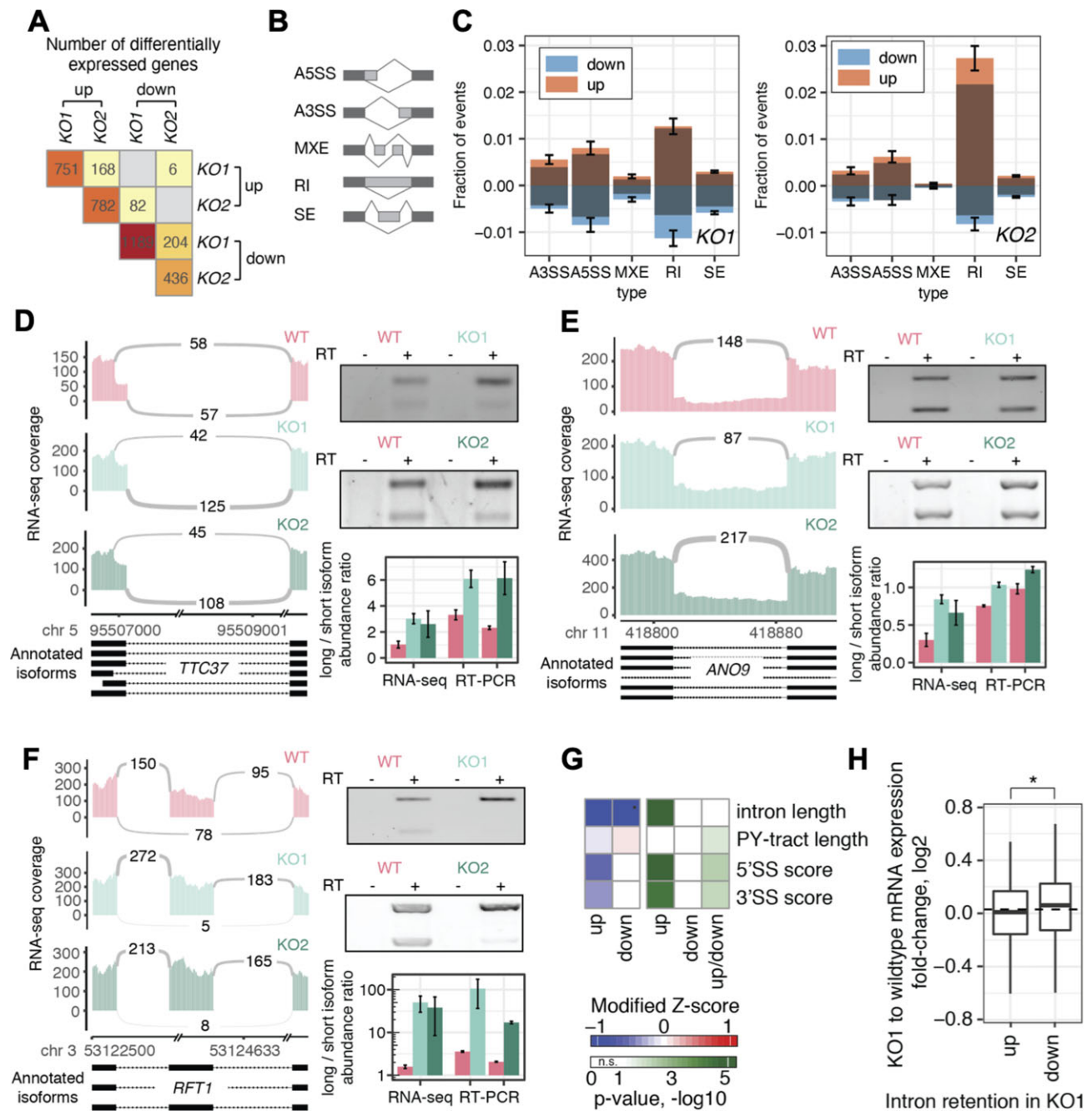
## DISCUSSION

Around 150 different types of RNA modifications decorate transcripts from the three domains of life and additional modified nucleotides are still being identified (15,98–100). The roles of many of these modifications have been investigated first in bacteria and unicellular eukaryotes, such as the yeast *S. cerevisiae*, and are now being unraveled in metazoa, particularly in humans. The first human RNA modification enzymes to be characterized were largely orthologous to yeast proteins. However, studies performed since the early 1980s on different human and/or metazoan RNA species revealed the presence of some specific epitranscriptomic marks absent from the corresponding yeast RNAs. Consequently, the metazoan enzymes responsible for the formation of many of these 'orphan' RNA modifications are progressively being identified.

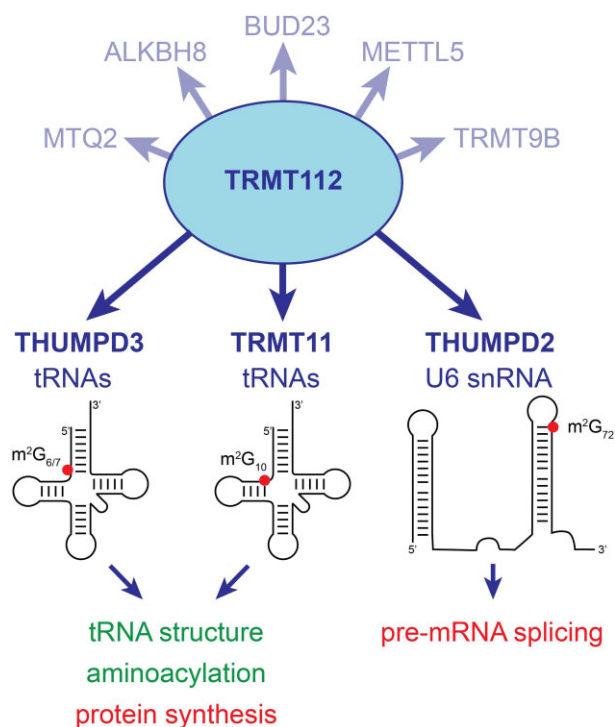


**Figure 4.** The human U6 snRNA contains an  $m^2G$  dependent on THUMP2. **(A)** Schematic secondary structure models of the U6 snRNA in different conformations. The position of  $m^2G_{72}$  is highlighted in red and circled. **(B)** Cryo-EM structure of the RNA components found in the active site of the C complex of the human spliceosome (PDB code: 6ZYM; (110)). The  $G_{72}$  and  $U_{74}$  positions from U6 snRNA are shown as sticks and the catalytic metal ion ( $Mg^{2+}$ ) coordinated (black dashed lines) by their phosphate groups is shown as a black sphere. The  $N^2$  atom from  $G_{72}$ , which is methylated by THUMP2-TRMT112 complex to form  $m^2G$  is shown as a sphere. **(C)** RNA-immunoprecipitation after UV cross-linking (254 nm) revealing a specific interaction of human THUMP2 with U6 snRNA. Input and eluate samples were separated by denaturing polyacrylamide gel electrophoresis and the indicated RNAs were detected by northern blotting using specific oligonucleotides. Two independent experiments were performed and a representative image is shown. **(D)** LC-HRMS elution profiles of  $m^2G$  standard (upper panel) or nucleosides obtained upon digestion of U6 snRNA purified from WT (second panel) and THUMP2 KO1 (third panel) or THUMP2 KO2 (bottom panel) HCT116 cell lines. The absolute intensity of  $m/z$  values of  $298^+$  (corresponding to  $m^2G$  protonated ion) is plotted. The dashed green line indicates the elution volume of  $m^2G$ . Three independent experiments were performed and a representative image is shown. **(E)** Growth analysis of WT and THUMP2 KO1 and KO2 HCT116 cell lines. Cell numbers were monitored every 24 h for 6 days. Error bars represent standard deviation calculated from three biological replicates. **(F)** U6 stability is not affected by the absence of THUMP2 protein. Northern blot showing the abundance of the U6, U6atac, U2 and U4 snRNAs and the 5.8S rRNA (loading control) in the WT and THUMP2 KO1 and KO2 cell lines. Three independent experiments were performed and a representative image is shown. The relative, normalized level of the different snRNAs in the THUMP2 KO1 and KO2 cell lines compared to WT is quantified. Error bars represent standard deviation calculated from three biological replicates.





**Figure 5.** THUMP2 is important for efficient pre-mRNA splicing. **(A)** Number of genes with significant expression changes and a  $\log_2$ -fold change of 0.5 up or down in the THUMP2 KO1, KO2 cell lines or both. **(B)** Schematic views of different types of alternative splicing are given on the left: Alternative 5' splice site (A5SS), alternative 3' splice site (A3SS), mutually exclusive exon (MXE), retained intron (RI) and skipped exon (SE). **(C)** Relative fractions of significant alternative splicing events identified with rMATS for THUMP2 KO1 (left panel) and KO2 (right panel). 95% confidence intervals are given as error bars (bootstrapping,  $n = 100$ ). Dark bars reflect the overlap with splicing changes in the same direction for the other THUMP2 KO cell line. **(D–F)** Visualization, validation and quantification of individual alternative splicing events by sashimi plots and RT-PCR. Sashimi plots show mRNA-seq coverage and junction counts of the pooled three replicates. Error bars reflect the standard deviation of three biological replicates. Change in A3SS in the gene *TTC37* (D), RI in *ANO9* (E), reduced SE/exon inclusion in *RFT1* (F) in THUMP2 KO1 and KO2 compared to WT) HCT116 cell lines. **(G)** Intron features associated with intron retention. Heatmaps show 1) the change of the median in more (up) or less (down) retained introns in THUMP2 KO1 identified with FRASER (plotted as modified Z-score) relative to introns without significant changes and 2) the significance of the changes plotted as  $-\log_{10}(P\text{-value})$  from the Wilcoxon rank-sum test. Supplementary Figure S7D shows the full distributions of all intron features without the modified Z-score transformation. PY – polypyrimidine, SS – splice site. **(H)** Expression changes associated with intron retention. To enhance the visibility of the global differences, outliers are not plotted here, but given in Supplementary Figure S9F. Significance was assessed with the Wilcoxon rank-sum test ( $*$   $P < 0.05$ ). The dashed line reflects the median expression change of all other genes.



**Figure 6.** Interaction network of the human TRMT112 protein with MTases, emphasizing the  $m^2G$  MTases. The substrates of each  $m^2G$  MTase are indicated and a representative view of their secondary structure is shown. The locations of the  $m^2G$  nucleotides are shown as red spheres. RNA functions affected upon depletion of the MTase are shown in red, whereas those not affected are shown in green.

### TRMT112 as a common cofactor of MTases with diverse functions in regulating gene expression

In this study, we have investigated the protein interaction network of human TRMT112 (Figure 1; Supplementary Figure S1). Probing the interactome of TRMT112 in intact cells by proximity labelling, we generated a robust inventory of TRMT112-associated proteins (Supplementary Table S5), which will serve as a valuable resource to learn more about the functions and regulation of this MTase cofactor. Intriguingly, extending the list of TRMT112 interactors beyond MTases, revealed numerous post-translational modification enzymes (predominantly kinases, protein methyltransferases and components of the (de-)ubiquitination machinery) that represent potential regulators of TRMT112 and/or the various TRMT112-MTase complexes. Furthermore, numerous RNA-binding proteins linked to various different aspects of gene expression *e.g.* transcription, ribosome assembly, tRNA biogenesis, mRNA maturation, translation and RNA decay were discovered, and verification of such interactor will likely provide important insights into the context of the modification events it is involved in. The interaction network between human TRMT112 and MTases is more extensive (eight partners) than in yeast (four partners; Figure 6). Among the TRMT112 interactors identified were three additional MTases, TRMT11, THUMPDP2 and THUMPDP3 (Figure 1D, E), confirming recent studies (51,52). Our results indicate that these three enzymes, each composed of a class I MTase domain fused to an N-terminal THUMP domain (Figures 2A and 3A), contribute to the formation of  $m^2G$  on RNAs (Supplementary Figure S3C-D, Figure 2C, 3C-

E and 4D). We demonstrate that the TRMT112-TRMT11 and TRMT112-THUMPDP3 complexes are responsible for installing  $m^2G$  modifications in tRNAs whereas TRMT112-THUMPDP2 targets the U6 snRNA (Figure 6).

These  $m^2G$  modifications were known for several decades but analysis of their biological functions was limited by lack of knowledge of the enzymes responsible for installing them (53,54,101–103). Here, we show that these three enzymes are important for the proliferation of human HCT116 colorectal cancer cells (Figures 2D and 4E, Supplementary Figure S5D). Hence, TRMT112 emerges as an essential cofactor for methylation events on diverse non-coding RNA molecules (18S rRNA, tRNAs and snRNA U6). The discovery that TRMT112 functions as the central hub of a MTase network opens the intriguing possibility of cross-regulation of different gene expression processes via coordinated regulation of RNA modification events.

### $m^2G$ s in the tRNA body function co-operatively to optimize protein synthesis for efficient cell proliferation

In this study, we functionally characterize the human TRMT11-TRMT112 complex for the first time, revealing that it catalyzes the formation of  $m^2G$  at position 10 of tRNAs (Supplementary Figure S3B-D and Figure 2C, Supplementary Figure S4D-F), similarly to yeast and *Archaeoglobus fulgidus* archaeal complexes (31,40,44). We also demonstrate that THUMPDP3 is responsible for the formation of  $m^2G$  at tRNA position 6, confirming the findings of a recent study that also shows the enzyme installs  $m^2G$  at position 7 of tRNA<sup>Trp</sup> (52), Figure 2C, Supplementary Figures S3B-D and S4A-C). Lack of either TRMT11 or THUMPDP3 have only minor phenotypic effects. However, the concomitant absence of both THUMPDP3 and TRMT11 results in an almost complete loss of  $m^2G$  in tRNAs (Figure 2C) and strongly affects the proliferation of the HCT116 cancer cell line (Figure 2D, Supplementary Figure S5D). Rationalizing this profound growth defect, we revealed that concomitant loss of THUMPDP3 and TRMT11 leads to reduced monosome and polysome levels and impairs global translation (Figure 2G,H; Supplementary Figure S6E). Collectively, these results indicate that the absence of individual  $m^2G$  modifications at positions 6, 7 or 10 of tRNAs is not detrimental for most of the investigated phenotypes, whereas the combined absence of these  $m^2G$  modifications on one or several tRNAs is important. This suggests functional interdependency of the  $m^2G$  modifications at these positions. Such synergies have also been observed between other tRNA modifications. Indeed, the deletion of individual, non-essential genes encoding yeast tRNA modification enzymes targeting regions outside the anticodon stem loop rarely leads to obvious phenotypes whereas double mutants, lacking two tRNA modification enzymes, often exhibit more pronounced growth defects (104,105). For example, *S. cerevisiae* strains lacking either the *TRM11* (yeast homolog of *TRMT11*) or *TRM1* genes grow as WT whereas the *trm11Δ/trm1Δ* double mutant exhibits a severe growth defect (40). With the methods used here, we observe no effect of the individual or concomitant absence of  $m^2G_{6/7/10}$  on tRNA stability (Supplementary Figure S6A-C; Supplementary Table S7), folding (Figure 2E) or aminoacylation (Figure 2F). Although future studies will be required to dissect the precise molecular functions of

these m<sup>2</sup>G modifications, our study highlights their combined importance for optimizing tRNA function during translation.

### THUMP2 as the elusive MTase responsible for formation of m<sup>2</sup>G in the U6 snRNA

The remaining TRMT112 MTase interactor revealed by our BioID analysis, THUMP2, has thus far remained functionally uncharacterized. Here, we demonstrate that THUMP2 is necessary for the *in cellulo* formation of m<sup>2</sup>G in RNAs with size ranging from 90 to 200 nt and that the recombinant *Bt*TRMT112–THUMP2 complex catalyzes the *in vitro* formation of m<sup>2</sup>G in human RNAs with the same size (Figure 3C–E). Interestingly, in the early 1980s, U6 snRNA extracted from rodent cells was shown to contain one m<sup>2</sup>G modification at position 72 but the enzyme responsible for this modification remained elusive (53,54). Our data reveal that THUMP2 interacts directly and specifically with the U6 snRNA and is necessary for the formation of m<sup>2</sup>G in human U6 snRNA (Figure 4C, D), implying that it is the long-sought MTase responsible for the formation of m<sup>2</sup>G at position 72. While we cannot exclude that THUMP2 also has other small RNA targets, it is likely that methylation of the U6 snRNAs represents a key function. Our data indicate no activity of *Bt*THUMP2–TRMT112 on RNAs < 90 nt, which encompasses tRNAs (Figure 3C, D) and we show that THUMP2 does not directly associate with other snRNAs, including the U6atac of the minor spliceosome (Figure 4C). A recent report suggests the presence of Trm11-dependent m<sup>2</sup>G in yeast mRNA (106), but in human cells, m<sup>2</sup>G has not been detected outside small non-coding RNAs (tRNAs and snRNAs). While it is possible that THUMP2 contributes to methylation of other RNA species, the development of novel m<sup>2</sup>G detection and mapping approaches will be required to enable characterization of other potential sites and verify the MTases responsible for installing them.

The U6 snRNA, which base pairs with U2 and U4 snRNAs as well as pre-mRNAs, undergoes many structural remodeling events during biogenesis of the U6 snRNP and assembly of functional spliceosomes (96). Analysis of the THUMP2 interactome did not reveal direct interactions with other snRNAs or strong association with known U6 biogenesis factors (Figure 4C, Supplementary Table 8, Supplementary Figure S8E) so the timing of m<sup>2</sup>G installation on U6 remains unknown. Currently, little is known about elements required for substrate recognition by the THUMP2–TRMT112 complex. In the context of various tRNA modification enzymes, the THUMP domain interacts with the tRNA acceptor stem and senses the presence of the CCA tail (31,42,86,107), thus functioning as a molecular ruler directing the associated catalytic domain to target specific nucleotides within the tRNA acceptor stem/D-arm. In the future, it will be interesting to dissect the role of the THUMP2 THUMP domain and determine if it is necessary for the U6 snRNA to adopt a (partial) tRNA-like structure to be modified by the THUMP2–TRMT112 complex.

### m<sup>2</sup>G modification of the U6 snRNA is important for pre-mRNA splicing

The U6 snRNA plays a central role in pre-mRNA splicing by forming catalytic elements of the spliceosome active site (96). In particular, the phosphate group of U6–U<sub>74</sub> plays a central

role in splicing by coordinating the catalytic metal ions. Interestingly, the m<sup>2</sup>G<sub>72</sub> nucleotide is located two nucleotides upstream of U<sub>74</sub> and its phosphate group also coordinates one of the catalytic metal ions (Figure 4B). Compared to parental HCT116 cells, the proliferation of THUMP2 KO1 and KO2 cells is strongly reduced (Figure 4E), likely due to alterations in pre-mRNA splicing. In line with our finding that the U6, but not the U6atac, snRNA is associated with THUMP2 (Figure 4C), only 6% of genes that showed differential alternative splicing contained introns known to be targeted by the minor spliceosome (Olthof et al; <https://doi.org/10.1101/2022.10.12.511939>), which is not significant different from the genome average (chi-squared test *P*-value > 0.05).

mRNA-seq analyses revealed a plethora of alternative pre-mRNA splicing events affected in both the THUMP2 KO1 and KO2 cells (Figure 5C, Supplementary Figure S9B–C, G–H). Closer analysis revealed that introns retained in cells lacking THUMP2 are generally shorter, have shorter polypyrimidine tracts and lower splice site quality than average (Figure 5G, Supplementary Figure S9F). Furthermore, the expression levels of mRNAs with increased intron retention in the THUMP2 KO1 cell line were significantly reduced compared to mRNAs with better intron removal (Figure 5H). This might be due to the presence of in-frame premature stop codons on retained introns, which then are likely to trigger the nonsense mediated mRNA decay quality-control pathway (108,109). As the absence of m<sup>2</sup>G on U6 snRNA does not affect its stability (Figure 4F), it remains to be understood precisely how this modification influences pre-mRNA splicing. It is possible that the G<sub>72</sub> modification enhances the stability of the U6 snRNA internal stem-loop (ISL), thus rendering it important for splicing at sub-optimal sites. Interestingly, the effects observed on alternative splicing upon THUMP2 depletion mirror those resulting from depletion of LARP7, which is important for several 2'-O-methylations present within U6 snRNA ISL (110,111). In human cells, depletion of TFIIP11 also affects U6 snRNA 2'-O-methylation and results in decreased pre-mRNA splicing fidelity and increase in intron retention (112). Likewise, depletion of *Arabidopsis thaliana* FIO1, the orthologue of the METTL16 MTase responsible for installing m<sup>6</sup>A<sub>43</sub> in the human U6 snRNA (113,114), affects pre-mRNA splicing efficiency and accuracy (115). The U6 snRNA epitranscriptomic marks are therefore emerging as key regulators of (alternative) pre-mRNA splicing.

## CONCLUSION

Using an unbiased approach, we have identified three additional MTases (TRMT11, THUMP2 and THUMP3) as direct partners of human TRMT112. We also demonstrate that these three TRMT112-MTase complexes act as holoenzymes to catalyze the formation of m<sup>2</sup>G in small RNAs. Both THUMP3 and TRMT11 MTase subunits are essential for the introduction of m<sup>2</sup>G in tRNA body and are necessary for optimal protein synthesis and for cell proliferation. The THUMP2 MTase introduces the m<sup>2</sup>G modification on U6 snRNA, the central RNA involved in the catalytic step of pre-mRNA splicing and is important for optimal pre-mRNA splicing. This study contributes to the characterization of the human enzymes responsible for 'orphan' RNA modifications. It also emphasizes on the central

role of TRMT112 as an intrinsic component of holoenzymes modifying various RNAs to fine-tune two essential cellular processes.

## DATA AVAILABILITY

The mRNA-seq datasets for the WT HTC116 and THUMP2 KO cells are deposited in Gene Expression Omnibus (GEO) database [<http://www.ncbi.nlm.nih.gov/geo/>] under the accession code GSE219260. The mass spectrometry proteomics data from the BioID experiments have been deposited to the ProteomeXchange Consortium via the PRIDE (60) partner repository with the dataset identifier PXD038997. The mass spectrometry proteomics data from the THUMP2 IP experiments have been deposited to the ProteomeXchange Consortium via the PRIDE (60) partner repository with the dataset identifier PXD038967 and 10.6019/PXD038967.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We are indebted to Dr Clément Chapat for his invaluable help with the BioID experiments and for his advice with human cell culture in general. We thank Philipp Hackert for technical assistance.

## FUNDING

C.W. is the recipient of a fellowship from the Chinese Scholarship Council (CSC); K.E.B. and M.T.B. acknowledge funding from the Deutsche Forschungsgemeinschaft (DFG) via SFB1565 [project number 469281184; P18 to M.T.B. and P12 to K.E.B.]; Cluster of Excellence ‘Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells’ (MBExC) [project number 390729940 to M.T.B.]; M.G. acknowledges financial supports from Fondation ARC pour la Recherche sur le Cancer; Centre National pour la Recherche Scientifique (CNRS); Agence Nationale pour la Recherche (ANR) [ANR-14-CE09-0016-02, ANR-22-CE12-0010-01]; Ecole Polytechnique; AAP AIN INSB CNRS; D.T. and V.G. acknowledge financial supports of DIM MAP (Région Ile-de-France); ANR Vapobio [ANR-20-CE29-0016]; C.C. is supported by the Agence Nationale de la Recherche [ANR-15-CE14-0013-02 for the PhD of N.P.]; Region Grand-Est (PhD fellowship of C.M.); French Proteomic Infrastructure [ProFI FR2048, ANR-10-INBS-08-03].

## Conflict of interest statement

None declared.

## REFERENCES

- Zhao,B.S., Roundtree,I.A. and He,C. (2017) Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.*, **18**, 31–42.
- Motorin,Y. and Helm,M. (2011) RNA nucleotide methylation. *Wiley Interdiscip. Rev. RNA*, **2**, 611–631.
- Agris,P.F., Narendran,A., Sarachan,K., Vare,V.Y.P. and Eruyal,E. (2017) The importance of being modified: the role of RNA modifications in translational fidelity. *Enzymes*, **41**, 1–50.
- Sloan,K.E., Warda,A.S., Sharma,S., Entian,K.D., Lafontaine,D.L.J. and Bohnsack,M.T. (2017) Tuning the ribosome: the influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol.*, **14**, 1138–1152.
- Bohnsack,M.T. and Sloan,K.E. (2018) Modifications in small nuclear RNAs and their roles in spliceosome assembly and function. *Biol. Chem.*, **399**, 1265–1276.
- Krutyholowa,R., Zakrzewski,K. and Glatt,S. (2019) Charging the code - tRNA modification complexes. *Curr. Opin. Struct. Biol.*, **55**, 138–146.
- Angelova,M.T., Dimitrova,D.G., Dinges,N., Lence,T., Worpenberg,L., Carre,C. and Roignant,J.Y. (2018) The emerging field of epitranscriptomics in neurodevelopmental and neuronal disorders. *Front. Bioeng. Biotechnol.*, **6**, 46.
- Barbieri,I. and Kouzarides,T. (2020) Role of RNA modifications in cancer. *Nat. Rev. Cancer*, **20**, 303–322.
- Bednarova,A., Hanna,M., Durham,I., VanCleave,T., England,A., Chaudhuri,A. and Krishnan,N. (2017) Lost in translation: defects in transfer RNA modifications and neurological disorders. *Front. Mol. Neurosci.*, **10**, 135.
- Delaunay,S. and Frye,M. (2019) RNA modifications regulating cell fate in cancer. *Nat. Cell Biol.*, **21**, 552–559.
- Jonkhout,N., Tran,J., Smith,M.A., Schonrock,N., Mattick,J.S. and Novoa,E.M. (2017) The RNA modification landscape in human disease. *RNA*, **23**, 1754–1769.
- Livneh,I., Moshitch-Moshkovitz,S., Amariglio,N., Rechavi,G. and Dominissini,D. (2020) The m(6)A epitranscriptome: transcriptome plasticity in brain development and function. *Nat. Rev. Neurosci.*, **21**, 36–51.
- Ramos,J. and Fu,D. (2019) The emerging impact of tRNA modifications in the brain and nervous system. *Biochim. Biophys. Acta Gene Regul. Mech.*, **1862**, 412–428.
- Roundtree,I.A., Evans,M.E., Pan,T. and He,C. (2017) Dynamic RNA modifications in gene expression regulation. *Cell*, **169**, 1187–1200.
- Suzuki,T. (2021) The expanding world of tRNA modifications and their disease relevance. *Nat. Rev. Mol. Cell Biol.*, **22**, 375–392.
- Fenwick,M.K. and Ealick,S.E. (2018) Towards the structural characterization of the human methyltransferase. *Curr. Opin. Struct. Biol.*, **53**, 12–21.
- Graille,M. (2022) Division of labor in epitranscriptomics: what have we learnt from the structures of eukaryotic and viral multimeric RNA methyltransferases? *Wiley Interdiscip. Rev. RNA*, **13**, e1673.
- Guy,M.P. and Phizicky,E.M. (2014) Two-subunit enzymes involved in eukaryotic post-transcriptional tRNA modification. *RNA Biol.*, **11**, 1608–1618.
- Garcias Morales,D. and Reyes,J.L. (2021) A birds'-eye view of the activity and specificity of the mRNA m(6)A methyltransferase complex. *Wiley Interdiscip. Rev. RNA*, **12**, e1618.
- Bourgeois,G., Letoquart,J., van Tran,N. and Graille,M. (2017) Trm112, a protein activator of methyltransferases modifying actors of the eukaryotic translational apparatus. *Biomolecules*, **7**, 7.
- van Tran,N., Muller,L., Ross,R.L., Lestini,R., Letoquart,J., Ulryck,N., Limbach,P.A., de Crecy-Lagard,V., Cianferani,S. and Graille,M. (2018) Evolutionary insights into Trm112-methyltransferase holoenzymes involved in translation between archaea and eukaryotes. *Nucleic Acids Res.*, **46**, 8483–8499.
- Figaro,S., Wacheul,L., Schillewaert,S., Graille,M., Huvelle,E., Mongeard,R., Zorbas,C., Lafontaine,D.L. and Heurgue-Hamard,V. (2012) Trm112 is required for

- Bud23-mediated methylation of the 18S rRNA at position G1575. *Mol. Cell Biol.*, **32**, 2254–2267.
23. Letoquart, J., Huvelle, E., Wacheul, L., Bourgeois, G., Zorbas, C., Graille, M., Heurgue-Hamard, V. and Lafontaine, D.L. (2014) Structural and functional studies of Bud23-Trm112 reveal 18S rRNA N7-G1575 methylation occurs on late 40S precursor ribosomes. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E5518–E5526.
  24. Heurgue-Hamard, V., Champ, S., Mora, L., Merkulova-Rainon, T., Kisselev, L.L. and Buckingham, R.H. (2005) The glutamine residue of the conserved GGQ motif in *Saccharomyces cerevisiae* release factor eRF1 is methylated by the product of the YDR140w gene. *J. Biol. Chem.*, **280**, 2439–2445.
  25. Polevoda, B., Span, L. and Sherman, F. (2006) The yeast translation release factors Mrf1p and Sup45p (eRF1) are methylated, respectively, by the methyltransferases Mtq1p and Mtq2p. *J. Biol. Chem.*, **281**, 2562–2571.
  26. Heurgue-Hamard, V., Graille, M., Scrima, N., Ulryck, N., Champ, S., van Tilbeurgh, H. and Buckingham, R.H. (2006) The zinc finger protein Ynr046w is plurifunctional and a component of the eRF1 methyltransferase in yeast. *J. Biol. Chem.*, **281**, 36140–36148.
  27. Liger, D., Mora, L., Lazar, N., Figaro, S., Henri, J., Scrima, N., Buckingham, R.H., van Tilbeurgh, H., Heurgue-Hamard, V. and Graille, M. (2011) Mechanism of activation of methyltransferases involved in translation by the Trm112 ‘hub’ protein. *Nucleic Acids Res.*, **39**, 6249–6259.
  28. Brown, A., Shao, S., Murray, J., Hegde, R.S. and Ramakrishnan, V. (2015) Structural basis for stop codon recognition in eukaryotes. *Nature*, **524**, 493–496.
  29. Lacoux, C., Wacheul, L., Saraf, K., Pythoud, N., Huvelle, E., Figaro, S., Graille, M., Carapito, C., Lafontaine, D.L.J. and Heurgue-Hamard, V. (2020) The catalytic activity of the translation termination factor methyltransferase Mtq2-Trm112 complex is required for large ribosomal subunit biogenesis. *Nucleic Acids Res.*, **48**, 12310–12325.
  30. Metzger, E., Wang, S., Urban, S., Willmann, D., Schmidt, A., Offermann, A., Allen, A., Sum, M., Obier, N., Cottard, F., et al. (2019) KMT9 monomethylates histone H4 lysine 12 and controls proliferation of prostate cancer cells. *Nat. Struct. Mol. Biol.*, **26**, 361–371.
  31. Bourgeois, G., Marcoux, J., Saliou, J.M., Cianferani, S. and Graille, M. (2017) Activation mode of the eukaryotic m2G10 tRNA methyltransferase Trm11 by its partner protein Trm112. *Nucleic Acids Res.*, **45**, 1971–1982.
  32. Letoquart, J., van Tran, N., Caroline, V., Aleksandrov, A., Lazar, N., van Tilbeurgh, H., Liger, D. and Graille, M. (2015) Insights into molecular plasticity in protein complexes from Trm9-Trm112 tRNA modifying enzyme crystal structure. *Nucleic Acids Res.*, **43**, 10989–11002.
  33. Mazaauric, M.H., Dirick, L., Purushothaman, S.K., Bjork, G.R. and Lapeyre, B. (2010) Trm112p is a 15-kDa zinc finger protein essential for the activity of two tRNA and one protein methyltransferases in yeast. *J. Biol. Chem.*, **285**, 18505–18515.
  34. Johansson, M.J., Esberg, A., Huang, B., Bjork, G.R. and Bystrom, A.S. (2008) Eukaryotic wobble uridine modifications promote a functionally redundant decoding system. *Mol. Cell Biol.*, **28**, 3301–3312.
  35. Kalhor, H.R. and Clarke, S. (2003) Novel methyltransferase for modified uridine residues at the wobble position of tRNA. *Mol. Cell Biol.*, **23**, 9283–9292.
  36. Fu, D., Brophy, J.A., Chan, C.T., Atmore, K.A., Begley, U., Paules, R.S., Dedon, P.C., Begley, T.J. and Samson, L.D. (2010) Human AlkB homolog ABH8 is a tRNA methyltransferase required for wobble uridine modification and DNA damage survival. *Mol. Cell Biol.*, **30**, 2449–2459.
  37. Songe-Moller, L., van den Born, E., Leihne, V., Vagbo, C.B., Kristoffersen, T., Krokan, H.E., Kirpekar, F., Falnes, P.O. and Klungland, A. (2010) Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding. *Mol. Cell Biol.*, **30**, 1814–1827.
  38. Begley, U., Sosa, M.S., Avivar-Valderas, A., Patil, A., Endres, L., Estrada, Y., Chan, C.T., Su, D., Dedon, P.C., Aguirre-Ghiso, J.A., et al. (2013) A human tRNA methyltransferase 9-like protein prevents tumour growth by regulating LIN9 and HIF1- $\alpha$ . *EMBO Mol. Med.*, **5**, 366–383.
  39. Gu, C., Ramos, J., Begley, U., Dedon, P.C., Fu, D. and Begley, T.J. (2018) Phosphorylation of human TRM9L integrates multiple stress-signaling pathways for tumor growth suppression. *Sci. Adv.*, **4**, eaas9184.
  40. Purushothaman, S.K., Bujnicki, J.M., Grosjean, H. and Lapeyre, B. (2005) Trm11p and Trm112p are both required for the formation of 2-methylguanosine at position 10 in yeast tRNA. *Mol. Cell Biol.*, **25**, 4359–4370.
  41. Armengaud, J., Urbonavicius, J., Fernandez, B., Chaussinand, G., Bujnicki, J.M. and Grosjean, H. (2004) N2-methylation of guanosine at position 10 in tRNA is catalyzed by a THUMP domain-containing, S-adenosylmethionine-dependent methyltransferase, conserved in Archaea and Eukaryota. *J. Biol. Chem.*, **279**, 37142–37152.
  42. Hirata, A., Nishiyama, S., Tamura, T., Yamauchi, A. and Hori, H. (2016) Structural and functional analyses of the archaeal tRNA m2G/m22G10 methyltransferase aTrm11 provide mechanistic insights into site specificity of a tRNA methyltransferase that contains common RNA-binding modules. *Nucleic Acids Res.*, **44**, 6377–6390.
  43. Hirata, A., Suzuki, T., Nagano, T., Fujii, D., Okamoto, M., Sora, M., Lowe, T.M., Kanai, T., Atomi, H., Suzuki, T., et al. (2019) Distinct modified nucleosides in tRNA(Trp) from the hyperthermophilic archaeon *Thermococcus kodakarensis* and requirement of tRNA m(2)G10/m(2) 2G10 methyltransferase (archaeal Trm11) for survival at high temperatures. *J. Bacteriol.*, **201**, e00448-19.
  44. Wang, C., van Tran, N., Jactel, V., Guerineau, V. and Graille, M. (2020) Structural and functional insights into *Archaeoglobus fulgidus* m2G10 tRNA methyltransferase Trm11 and its Trm112 activator. *Nucleic Acids Res.*, **48**, 11068–11082.
  45. Maden, B.E. (1986) Identification of the locations of the methyl groups in 18 S ribosomal RNA from *Xenopus laevis* and man. *J. Mol. Biol.*, **189**, 681–699.
  46. van Tran, N., Ernst, F.G.M., Hawley, B.R., Zorbas, C., Ulryck, N., Hackert, P., Bohnsack, K.E., Bohnsack, M.T., Jaffrey, S.R., Graille, M., et al. (2019) The human 18S rRNA m6A methyltransferase METTL5 is stabilized by TRMT112. *Nucleic Acids Res.*, **47**, 7719–7733.
  47. Ignatova, V.V., Stolz, P., Kaiser, S., Gustafsson, T.H., Lastres, P.R., Sanz-Moreno, A., Cho, Y.L., Amarie, O.V., Aguilar-Pimentel, A., Klein-Rodewald, T., et al. (2020) The rRNA m(6)A methyltransferase METTL5 is involved in pluripotency and developmental programs. *Genes Dev.*, **34**, 715–729.
  48. Leismann, J., Spagnuolo, M., Pradhan, M., Wacheul, L., Vu, M.A., Musheev, M., Mier, P., Andrade-Navarro, M.A., Graille, M., Niehrs, C., et al. (2020) The 18S ribosomal RNA m(6)A methyltransferase Mettl5 is required for normal walking behavior in *Drosophila*. *EMBO Rep.*, **21**, e49443.
  49. Rong, B., Zhang, Q., Wan, J., Xing, S., Dai, R., Li, Y., Cai, J., Xie, J., Song, Y., Chen, J., et al. (2020) Ribosome 18S m(6)A methyltransferase METTL5 promotes translation initiation and breast cancer cell growth. *Cell Rep.*, **33**, 108544.
  50. Xing, M., Liu, Q., Mao, C., Zeng, H., Zhang, X., Zhao, S., Chen, L., Liu, M., Shen, B., Guo, X., et al. (2020) The 18S rRNA m(6)A methyltransferase METTL5 promotes mouse embryonic stem cell differentiation. *EMBO Rep.*, **21**, e49863.
  51. Brumele, B., Mutso, M., Telanne, L., Ounap, K., Spunde, K., Abroi, A. and Kurg, R. (2021) Human TRMT112-methyltransferase network consists of seven partners interacting with a common co-factor. *Int. J. Mol. Sci.*, **22**, 13593.
  52. Yang, W.Q., Xiong, Q.P., Ge, J.Y., Li, H., Zhu, W.Y., Nie, Y., Lin, X., Lv, D., Li, J., Lin, H., et al. (2021) THUMP3-TRMT112 is a

- m2G methyltransferase working on a broad range of tRNA substrates. *Nucleic Acids Res.*, **49**, 11900–11919.
53. Epstein, P., Reddy, R., Henning, D. and Busch, H. (1980) The nucleotide sequence of nuclear U6 (4.7 S) RNA. *J. Biol. Chem.*, **255**, 8901–8906.
  54. Harada, F., Kato, N. and Nishimura, S. (1980) The nucleotide sequence of nuclear 4.8S RNA of mouse cells. *Biochem. Biophys. Res. Commun.*, **95**, 1332–1340.
  55. Seiler, C.Y., Park, J.G., Sharma, A., Hunter, P., Surapaneni, P., Sedillo, C., Field, J., Algar, R., Price, A., Steel, J., et al. (2014) DNASU plasmid and PSI:biology-materials repositories: resources to accelerate biological research. *Nucleic Acids Res.*, **42**, D1253–D1260.
  56. Zheng, L., Baumann, U. and Reymond, J.L. (2004) An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.*, **32**, e115.
  57. Chapat, C., Jafarnejad, S.M., Matta-Camacho, E., Hesketh, G.G., Gelbart, I.A., Attig, J., Gkogkas, C.G., Alain, T., Stern-Ginossar, N., Fabian, M.R., et al. (2017) Cap-binding protein 4EHP effects translation silencing by microRNAs. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 5425–5430.
  58. Deutsch, E.W., Csordas, A., Sun, Z., Jarnuczak, A., Perez-Riverol, Y., Ternent, T., Campbell, D.S., Bernal-Llinares, M., Okuda, S., Kawano, S., et al. (2017) The ProteomeXchange consortium in 2017: supporting the cultural change in proteomics public data deposition. *Nucleic Acids Res.*, **45**, D1100–D1106.
  59. Wiczorek, S., Combes, F., Lazar, C., Giai Gianetto, Q., Gatto, L., Dorffer, A., Hesse, A.M., Coute, Y., Ferro, M., Bruley, C., et al. (2017) DAPAR & ProStaR: software to perform statistical analyses in quantitative discovery proteomics. *Bioinformatics*, **33**, 135–136.
  60. Bouyssie, D., Hesse, A.M., Mouton-Barbosa, E., Rompais, M., Macron, C., Carapito, C., Gonzalez de Peredo, A., Coute, Y., Dupierriis, V., Burel, A., et al. (2020) Proline: an efficient and user-friendly software suite for large-scale proteomics. *Bioinformatics*, **36**, 3148–3155.
  61. Perez-Riverol, Y., Bai, J., Bandla, C., Garcia-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., et al. (2022) The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.*, **50**, D543–D552.
  62. Mechulam, Y., Guillon, L., Yatime, L., Blanquet, S. and Schmitt, E. (2007) Protection-based assays to measure aminoacyl-tRNA binding to translation initiation factors. *Methods Enzymol.*, **430**, 265–281.
  63. Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. and Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.*, **8**, 2281–2308.
  64. Labun, K., Montague, T.G., Krause, M., Torres Cleuren, Y.N., Tjeldnes, H. and Valen, E. (2019) CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.*, **47**, W171–W174.
  65. Janssen, B.D., Diner, E.J. and Hayes, C.S. (2012) Analysis of aminoacyl- and peptidyl-tRNAs by gel electrophoresis. *Methods Mol. Biol.*, **905**, 291–309.
  66. 1000 Genome Project Data Processing Subgroup, Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Durbin, R. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078–2079.
  67. Quinlan, A.R. and Hall, I.M. (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, **26**, 841–842.
  68. Kent, W.J., Zweig, A.S., Barber, G., Hinrichs, A.S. and Karolchik, D. (2010) BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics*, **26**, 2204–2207.
  69. Thorvaldsdottir, H., Robinson, J.T. and Mesirov, J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform.*, **14**, 178–192.
  70. Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A. and Huber, W. (2005) BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*, **21**, 3439–3440.
  71. Durinck, S., Spellman, P.T., Birney, E. and Huber, W. (2009) Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.*, **4**, 1184–1191.
  72. Olthof, A.M., Hyatt, K.C. and Kanadia, R.N. (2019) Minor intron splicing revisited: identification of new minor intron-containing genes and tissue-dependent retention and alternative splicing of minor introns. *BMC Genomics [Electronic Resource]*, **20**, 686.
  73. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
  74. Shen, S., Park, J.W., Lu, Z.X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q. and Xing, Y. (2014) rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E5593–E5601.
  75. Mertens, C., Scheller, I.F., Yopez, V.A., Celik, M.H., Liang, Y., Kremer, L.S., Gusic, M., Prokisch, H. and Gagneur, J. (2021) Detection of aberrant splicing events in RNA-seq data using FRASER. *Nat. Commun.*, **12**, 529.
  76. Yeo, G. and Burge, C.B. (2004) Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J. Comput. Biol.*, **11**, 377–394.
  77. Ramachandran, K.M. and Tsokos, C.P. (eds). (2015) In: *Mathematical Statistics with Applications in R. Second*. Boston Academic Press, pp. 687–731.
  78. Garrido-Martin, D., Palumbo, E., Guigo, R. and Breschi, A. (2018) ggsashimi: sashimi plot revised for browser- and annotation-independent splicing visualization. *PLoS Comput. Biol.*, **14**, e1006360.
  79. Roux, K.J., Kim, D.I., Burke, B. and May, D.G. (2018) BioID: a screen for protein-protein interactions. *Curr. Protoc. Protein Sci.*, **91**, 19.23.1–19.23.15.
  80. Mouaikel, J., Verheggen, C., Bertrand, E., Tazi, J. and Bordonne, R. (2002) Hypermethylation of the cap structure of both yeast snRNAs and snoRNAs requires a conserved methyltransferase that is localized to the nucleolus. *Mol. Cell*, **9**, 891–901.
  81. Zorbas, C., Nicolas, E., Wacheul, L., Huvelle, E., Heurgue-Hamard, V. and Lafontaine, D.L. (2015) The human 18S rRNA base methyltransferases DIMT1L and WBSR22-TRMT112 but not rRNA modification are required for ribosome biogenesis. *Mol. Biol. Cell*, **26**, 2080–2095.
  82. Sardana, R., Zhu, J., Gill, M. and Johnson, A.W. (2014) Physical and functional interaction between the methyltransferase Bud23 and the essential DEAH-box RNA helicase Ecm16. *Mol. Cell Biol.*, **34**, 2208–2220.
  83. Choudhury, P., Hackert, P., Memet, I., Sloan, K.E. and Bohnsack, M.T. (2019) The human RNA helicase DHX37 is required for release of the U3 snoRNP from pre-ribosomal particles. *RNA Biol.*, **16**, 54–68.
  84. Roychowdhury, A., Joret, C., Bourgeois, G., Heurgue-Hamard, V., Lafontaine, D.L.J. and Graille, M. (2019) The DEAH-box RNA helicase Dhr1 contains a remarkable carboxyl terminal domain essential for small ribosomal subunit biogenesis. *Nucleic Acids Res.*, **47**, 7548–7563.
  85. Madeira, F., Pearce, M., Tivey, A.R.N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A. and Lopez, R. (2022) Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.*, **50**, W276–W279.
  86. Neumann, P., Lakomek, K., Naumann, P.T., Erwin, W.M., Lauhon, C.T. and Ficner, R. (2014) Crystal structure of a 4-thiouridine synthetase-RNA complex reveals specificity of tRNA U8 modification. *Nucleic Acids Res.*, **42**, 6673–6685.
  87. Bujnicki, J.M. (2000) Phylogenomic analysis of 16S rRNA:(guanine-N2) methyltransferases suggests new family members and reveals highly conserved motifs and a domain

- structure similar to other nucleic acid amino-methyltransferases. *FASEB J.*, **14**, 2365–2368.
88. Nakahigashi, K., Kubo, N., Narita, S., Shimaoka, T., Goto, S., Oshima, T., Mori, H., Maeda, M., Wada, C. and Inokuchi, H. (2002) HemK, a class of protein methyl transferase with similarity to DNA methyl transferases, methylates polypeptide chain release factors, and hemK knockout induces defects in translational termination. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 1473–1478.
  89. de Crecy-Lagard, V., Boccaletto, P., Mangleburg, C.G., Sharma, P., Lowe, T.M., Leidel, S.A. and Bujnicki, J.M. (2019) Matching tRNA modifications in humans to their known and predicted enzymes. *Nucleic Acids Res.*, **47**, 2143–2159.
  90. Kadaba, S., Krueger, A., Trice, T., Krecic, A.M., Hinnebusch, A.G. and Anderson, J. (2004) Nuclear surveillance and degradation of hypomodified initiator tRNAMet in *S. cerevisiae*. *Genes Dev.*, **18**, 1227–1240.
  91. Motorin, Y. and Helm, M. (2010) tRNA stabilization by modified nucleotides. *Biochemistry*, **49**, 4934–4944.
  92. Kleiber, N., Lemus-Diaz, N., Stiller, C., Heinrichs, M., Mai, M.M., Hackert, P., Richter-Dennerlein, R., Hobartner, C., Bohnsack, K.E. and Bohnsack, M.T. (2022) The RNA methyltransferase METTL8 installs m(3)C32 in mitochondrial tRNAs(Thr/Ser(UCN)) to optimise tRNA structure and mitochondrial translation. *Nat. Commun.*, **13**, 209.
  93. Suzuki, T., Ueda, T. and Watanabe, K. (1997) The ‘polysemous’ codon–a codon with multiple amino acid assignment caused by dual specificity of tRNA identity. *EMBO J.*, **16**, 1122–1134.
  94. Pang, Y.L., Poruri, K. and Martinis, S.A. (2014) tRNA synthetase: tRNA aminoacylation and beyond. *Wiley Interdiscip. Rev. RNA*, **5**, 461–480.
  95. Didychuk, A.L., Butcher, S.E. and Brow, D.A. (2018) The life of U6 small nuclear RNA, from cradle to grave. *RNA*, **24**, 437–460.
  96. Wilkinson, M.E., Charenton, C. and Nagai, K. (2020) RNA splicing by the spliceosome. *Annu. Rev. Biochem.*, **89**, 359–388.
  97. Fica, S.M., Tuttle, N., Novak, T., Li, N.S., Lu, J., Koodathingal, P., Dai, Q., Staley, J.P. and Piccirilli, J.A. (2013) RNA catalyses nuclear pre-mRNA splicing. *Nature*, **503**, 229–234.
  98. Dal Magro, C., Keller, P., Kotter, A., Werner, S., Duarte, V., Marchand, V., Ignarski, M., Freiwald, A., Muller, R.U., Dieterich, C., et al. (2018) A Vastly Increased Chemical Variety of RNA Modifications Containing a Thioacetal Structure. *Angew. Chem. Int. Ed Engl.*, **57**, 7893–7897.
  99. Boccaletto, P., Stefaniak, F., Ray, A., Cappannini, A., Mukherjee, S., Purta, E., Kurkowska, M., Shirvanizadeh, N., Destefanis, E., Groza, P., et al. (2022) MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res.*, **50**, D231–D235.
  100. Ohira, T., Minowa, K., Sugiyama, K., Yamashita, S., Sakaguchi, Y., Miyauchi, K., Noguchi, R., Kaneko, A., Orita, I., Fukui, T., et al. (2022) Reversible RNA phosphorylation stabilizes tRNA for cellular thermotolerance. *Nature*, **605**, 372–379.
  101. Fournier, M., Labouesse, J., Dirheimer, G., Fix, C. and Keith, G. (1978) Primary structure of bovine liver tRNATrp. *Biochim. Biophys. Acta*, **521**, 198–208.
  102. Gupta, R.C., Roe, B.A. and Randerath, K. (1979) The nucleotide sequence of human tRNAGly (anticodon GCC). *Nucleic Acids Res.*, **7**, 959–970.
  103. Harada, F., Matsubara, M. and Kato, N. (1984) Stable tRNA precursors in HeLa cells. *Nucleic Acids Res.*, **12**, 9263–9269.
  104. Hopper, A.K. and Phizicky, E.M. (2003) tRNA transfers to the limelight. *Genes Dev.*, **17**, 162–180.
  105. Alexandrov, A., Chernyakov, I., Gu, W., Hiley, S.L., Hughes, T.R., Grayhack, E.J. and Phizicky, E.M. (2006) Rapid tRNA decay can result from lack of nonessential modifications. *Mol. Cell*, **21**, 87–96.
  106. Jones, J.D., Franco, M.K., Smith, T.J., Snyder, L.R., Anders, A.G., Ruotolo, B.T., Kennedy, R.T. and Koutmou, K.S. (2023) Methylated guanosine and uridine modifications in *S. cerevisiae* mRNAs modulate translation elongation. *RSC Chemical Biol.*, **4**, 363–378.
  107. Nishida, Y., Ohmori, S., Kakizono, R., Kawai, K., Namba, M., Okada, K., Yamagami, R., Hirata, A. and Hori, H. (2022) Required elements in tRNA for methylation by the eukaryotic tRNA (guanine-N(2)-) methyltransferase (Trm11-Trm112 complex). *Int. J. Mol. Sci.*, **23**, 4046.
  108. Hamid, F.M. and Makeyev, E.V. (2014) Emerging functions of alternative splicing coupled with nonsense-mediated decay. *Biochem. Soc. Trans.*, **42**, 1168–1173.
  109. Watabe, E., Togo-Ohno, M., Ishigami, Y., Wani, S., Hirota, K., Kimura-Asami, M., Hasan, S., Takei, S., Fukamizu, A., Suzuki, Y., et al. (2021) m(6) A-mediated alternative splicing coupled with nonsense-mediated mRNA decay regulates SAM synthetase homeostasis. *EMBO J.*, **40**, e106434.
  110. Bertram, K., El Ayoubi, L., Dybkov, O., Agafonov, D.E., Will, C.L., Hartmuth, K., Urlaub, H., Kastner, B., Stark, H. and Luhrmann, R. (2020) Structural insights into the roles of metazoan-specific splicing factors in the human step 1 spliceosome. *Mol. Cell*, **80**, 127–139.
  111. Hasler, D., Meduri, R., Bak, M., Lehmann, G., Heizinger, L., Wang, X., Li, Z.T., Sement, F.M., Bruckmann, A., Dock-Bregeon, A.C., et al. (2020) The alazami syndrome-associated protein LARP7 guides U6 small nuclear RNA modification and contributes to splicing robustness. *Mol. Cell*, **77**, 1014–1031.
  112. Duchemin, A., O’Grady, T., Hanache, S., Mereau, A., Thiry, M., Wacheul, L., Michaux, C., Perpete, E., Hervouet, E., Peixoto, P., et al. (6648) (2021) DHX15-independent roles for TFIP11 in U6 snRNA modification, U4/U6.U5 tri-snRNP assembly and pre-mRNA splicing fidelity. *Nat. Commun.*, **12**, 6648.
  113. Pendleton, K.E., Chen, B., Liu, K., Hunter, O.V., Xie, Y., Tu, B.P. and Conrad, N.K. (2017) The U6 snRNA m(6)A methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell*, **169**, 824–835.
  114. Warda, A.S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., Hobartner, C., Sloan, K.E. and Bohnsack, M.T. (2017) Human METTL16 is a N(6)-methyladenosine (m(6)A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. *EMBO Rep.*, **18**, 2004–2014.
  115. Parker, M.T., Soanes, B.K., Kusakina, J., Larrieu, A., Knop, K., Joy, N., Breidenbach, F., Sherwood, A.V., Barton, G.J., Fica, S.M., et al. (2022) m(6)A modification of U6 snRNA modulates usage of two major classes of pre-mRNA 5’ splice site. *Elife*, **11**, e78808.
  116. Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.*, **25**, 25–29.
  117. Gene Ontology, C. (2021) The Gene Ontology resource: enriching a Gold mine. *Nucleic Acids Res.*, **49**, D325–D334.
  118. Mi, H., Muruganujan, A. and Thomas, P.D. (2013) PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res.*, **41**, D377–D386.