THE UNIVERSITY OF CHICAGO

LOCATION-DEPENDENT EFFECTS OF N^6-METHYLADENOSINE ON mRNA TRANSLATION AND STABILITY

A DISSERTATION SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES

AND THE PRITZKER SCHOOL OF MEDICINE

IN CANDIDACY FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

INTERDISCIPLINARY SCIENTIST TRAINING PROGRAM:

BIOCHEMISTRY AND MOLECULAR BIOPHYSICS

BY

CARALINE LEE SEPICH

CHICAGO, ILLINOIS

DECEMBER 2022

To all of my family,
especially
Dave, Jayann, Katie, and A.J. Sepich,
and Greg Sepich-Poore.

TABLE OF CONTENTS

LIST OF FIGURES……..vi

LIST OF TABLES……..viii

ACKNOWLEDGMENTS……ix

ABSTRACT……xiv

INTRODUCTION……1

RNA modifications and the importance of their location in gene expression regulation……..1
Influences of chemical structure on functional consequences of RNA modifications..……...2
Modifications in various RNA types and the functional effects of their distribution patterns……....10
The impact of cellular location: nuclear functions of mRNA N^6-methyladenosine...……...15
The impact of cellular location: cytosolic roles of mRNA N^6-methyladenosine in translation……17
The impact of cellular location: cytosolic roles of mRNA N^6-methyladenosine in mRNA stability and decay……..22
Summary, outlook, and dissertation outline...……..….28

CHAPTER 2—THE METTL5-TRMT112 N^6-METHYLADENOSINE METHYLTRANSFERASE COMPLEX REGULATES mRNA TRANSLATION VIA 18S rRNA METHYLATION…….30

Abstract…..……31
Introduction……32
Results……35
METTL5 is an m^6A methyltransferase stabilized by TRMT112…….……35
18S rRNA is a major substrate of the METTL5-TRMT112 complex…….……..…....40
METTL5 regulates translation of a subset of transcripts…….……46
Mettl5$^{-/-}$ mice demonstrate growth and metabolic changes…….……....50
Discussion……..55
Supplemental Note..…..60
Materials and Methods..……..64

CHAPTER 3—LOCATION-SPECIFIC EFFECTS OF N^6-METHYLADENOSINE ON mRNA TRANSCRIPT STABILITY……94

Introduction……94
Results……98

Investigating the relationship between YTHDF proteins and regional m^{6}A location in regulating transcript stability……..…..98
P body dysregulation mediates mRNA stabilization upon triple knockdown of YTHDFs……...110
Discussion……116
Materials and Methods..……122

CHAPTER 4— SUMMARY AND PERSPECTIVES..……...……126

Advancements in understanding the role of m^{6}A in gene expression regulation...…….126
Contributions of this thesis to m^{6}A biology and future directions……...127
Characterization of the METTL5-TRMT112 complex, its role in translation, and future directions…….……127
Investigation of the impact of location on stabilization and decay of methylated transcripts…….…...133
Broader implications of this thesis to the field of m^{6}A biology……..138
Perspectives for investigations of RNA modifications…….……..…..140

APPENDIX: PROGRESS IN SINGLE-BASE RESOLUTION MAPPING OF *N*^{3}-METHYLCYTIDINE ON mRNA AND caRNA…….……..……..143

Introduction...……...143
Results...…….....146
Discussion...…….....157
Materials and Methods...……..161

REFERENCES……...167

LIST OF FIGURES

Figure 1.1: Common chemical modifications in mRNA…….3

Figure 1.2: The location of methylation influences functional consequences of modification..…….10

Figure 1.3: Distributions of chemical modifications along an mRNA transcript..……..15

Figure 1.4: The critical role of cellular and molecular location in mRNA N^6-methyladenosine (m^6A) function…….……...27

Figure 2.1: METTL5 is an m^6A methyltransferase stabilized by TRMT112……..36

Figure 2.2: Identification and investigation of METTL5-TRMT112 interactions……..38

Figure 2.3: METTL5 is predominantly cytoplasmic and methylates 18S rRNA..……..42

Figure 2.4: Investigation of METTL5 substrates in human cells……44

Figure 2.5: METTL5 affects cell proliferation and metabolism through differential translation of specific transcripts…….……..…..47

Figure 2.6: The effects of METTL5 knockout on translation and cell state……...49

Figure 2.7: *Mettl5* knockout mice have developmental and metabolic defects…….….51

Figure 2.8: Characterizations of *Mettl5*$^{-/-}$ mice…….……..53

Figure 2.9: Biochemical fractionation to identify METTL5……...63

Figure 3.1: m^6A in the 5'UTR correlates with increased stability…….….99

Figure 3.2: The effects of YTHDF loss on transcript stability by m^6A region…….….101

Figure 3.3: The effects of YTHDF loss on translation efficiency by m^6A region……..…103

Figure 3.4: The effects of YTHDF2 depletion mRNA lifetime by m^6A region……104

Figure 3.5: YTHDF interaction partners from public proximity-dependent biotin identification data.…108

Figure 3.6: Triple knockdown of YTHDF proteins causes transcriptomic dysregulation……..…113

Figure 3.7: Increased P body formation after YTHDF triple knockdown……..…115

Figure 4.1: Potential models of stabilization of 5'UTR-methylated transcripts by YTHDF2 ……..…135

Figure 4.2: Aspects of location that impact m^6A function……..…140

Figure A.1: N^3-methylcytidine in human RNA……..……147

Figure A.2: Method development for sequencing m^3C in human mRNA and caRNA……..…148

Figure A.3: Detection of mutations at m^3C sites in probes and tRNA from libraries……150

Figure A.4: Putative m^3C sites detected in HeLa cells…….……...……..153

Figure A.5: Effects of METTL8 depletion on m^3C sites…….……...…...155

Figure A.6: Changes in the transcriptome with METTL8 depletion…….…..…..157

LIST OF TABLES

Table 1.1. Chemical modifications in major cellular RNA types……...…..4

Table 2.1. METTL5 guide sequences…….…….....65

Table 2.2. siRNAs used to investigate METTL5…….…….....68

Table 2.3. Table of primers and oligonucleotides for METTL5 investigation….…….....90

Table 2.4. Antibodies used for METTL5 investigation….……..…….....91

Table 3.1. Datasets analyzed to study the effects of mRNA m^6A location…….…..……..…..122

ACKNOWLEDGEMENTS

I have many people to thank who have helped me through this journey and who had a hand in the work presented here.

Firstly, I would like to express my sincere gratitude to my advisor, Professor Chuan He. From the beginning, I have been inspired by his enthusiasm and curiosity about new ideas, as well as the talent with which he leads in bringing them to fruition. He has been a supportive mentor who cares about both science and his students, and always provides encouragement along with constructive feedback. He has fostered a great lab environment in which I have gained a good scientific foundation and training in how to ask important questions, design controlled experiments, and communicate results clearly.

Professor Sigrid Nachtergaele also mentored me starting with my rotation in the He lab when she was a postdoctoral fellow, but has continued to be a mentor to me even after she left to start her own lab at Yale. She dedicated a lot of her time over the first two years of my PhD to teaching me many fundamental wet lab techniques, as well as less concrete skills involved in designing experiments, making protocols, and staying organized. I sincerely appreciate her dedication to working with me to get the METTL5 project finished, even after she had many new responsibilities. I am also thankful for her kindness, resilience, and candor.

Dr. Zhong Zheng contributed significantly to our METTL5 project as well, helping directly with many aspects of the mouse work, and teaching me almost everything I know about working with mice in research. I am thankful to him for the time he spent teaching me and working on this project.

There are many other members of the He lab and collaborators who have been essential contributors to the projects presented in this thesis. Linda Zhang, Kailong Wen, and Prof. Xiaoxi

Zhuang made important contributions to the *METTL5* mouse experiments. Zhongyu Zou has been a hard-working, thoughtful, and helpful collaborator for the mRNA m^6A-related projects. Finally, Dr. Qing Dai, Kinga Pajdzik, Dr. Xiaoyang Dou, and Dr. Hui-Lung Sun laid a strong foundation for a project on the development of an m^3C sequencing method, which I had the opportunity to join.

Emily Schmitt spent over two years working with me as an undergraduate researcher. It was an incredible experience to watch her grow in her skills and competence as a researcher, and I am thankful for the ways she contributed to this work, as well as for her friendship. I wish her the best of luck as she pursues a career in medicine.

I would especially like to thank my labmates Dr. Allen Zhu, Dr. Yuru Wang, and Dr. Cody He for their friendship and camaraderie over the years, as well as my E307 labmates Dr. Lulu Hu, Xinran Feng, Dr. Bei Liu, and Dr. Chang Liu, for being kind and helpful and making E307 a great working environment.

I am sure that every He lab member has helped me at some point in this journey. I would like to thank Dr. Chang Ye, Dr. Xiaoyang Dou, Dr. Yong Peng, Dr. Shun Liu, Dr. Scott Zhang, Dr. Xiao-Long Cui, Dr. Bei Liu, and Dr. Allen Zhu for answering many questions about bioinformatics. Dr. Bryan Harada, Dr. Allen Zhu, Dr. Cody He, Zhongyu Zou, Dr. Hui-Lung Sun, Dr. Huiqing Zhou, Dr. Chang Liu, Dr. Hailing Shi, Dr. Yu Xiao, Dr. Jiangbo Wei, Boyang Gao, Yuhao Zhong, Dr. Phil Hsu, Dr. Diana West-Szymanski, Kinga Pajdzik, and Ruiqi Ge have provided reagents and advice for experiments. In particular, Dr. Lisheng Zhang is always willing to help when I have a question in either area. I have learned so much from my labmates and collaborators over the years and will miss working with them.

Many thanks to Dr. Jordi Tauler and Krissana Kowitwanich for keeping our lab running, and to the staff of the Chemistry Department and GCIS for keeping the department and building functional. I would also like to express my gratitude to the staff members of the core facilities, particularly the UCCC Sequencing Facility and Genomics Facility, as well as to the patient donors for the cell line samples used in my research.

I am grateful for all of my past research mentors who inspired my interest in science, taught me foundational research skills, and contributed to getting me here, including Charlin Vasquez and the staff at the Center of Excellence for Hazardous Materials Management, who gave me my first opportunity to work in a lab, Dr. Winnie Liang and Lori Phillips (TGen), Dr. Yongchang Chang, Dr. Qinhui Zhang, and Dr. Xiaojun Xu (Barrow Neurological Institute), and Prof. Petra Fromme, Dr. Jesse Coe, and Dr. Michael Vaughn (Arizona State University).

I have been fortunate to have a knowledgeable and supportive thesis committee including Prof. Tao Pan, Prof. Bryan Dickinson, Prof. Lucy Godley, and Prof. Alex Ruthenburg. I am thankful to them for taking the time to meet with me and for providing insight, guidance, and perspective on my projects and results over the years, as well as for being supportive through the turns my projects have taken.

The administrative leadership of the MSTP and BMB department have done many helpful things to assist me in my PhD. Prof. Marcus Clark, Alison Anastasio, Kristin McCann, Shay McAllister, Sarah Laloggia, Hafsah Mohammed, and Marisa Davis have all played important roles in my MSTP experience.

I am grateful for all of my friends who have been supportive along the way, from my friends since I was very young (including Brie Jones, Brittany Purvis, Sandya Athigaman, and Zayne Marquez), to my friends from Chicago. Thank you to Reba Abraham and Lakshmi Kirkire, my

roommates for my first four years in Chicago, who provided me with tea and baked goods, as well as a listening ear to discuss scientific ideas, experimental troubleshooting, or life in general. Thank you also to my other MSTP cohort members and their significant others, especially Saara-Anne Azizi and Jeffrey Dewey, as well as to my friends from the Graduate Christian Fellowship and New Community Covenant Church.

Most importantly, I am so grateful to all of my extended family. I have a large and wonderful family and every single one of them has been nothing but supportive and encouraging. Thank you to my in-laws, the Poores, for their welcoming kindness and encouragement. Thank you to my grandparents, Leonard and Olivia Salcido, for taking care of me as a child and encouraging me all the way through my PhD, as well as Frank and Mary Ann Sepich, who were also a major part of my childhood and instilled in me the value of pursuing higher education, even though it was never an opportunity they had. Watching my grandfathers battle cancer was also one of the original inspirations for my pursuit of a career in medicine.

To my parents, Dave and Jayann Sepich, thank you for all of your selfless love in raising me and in supporting my decision to pursue an MD/PhD. Thank you for providing me with examples of passion for your vocation and steadfast hard work. Finally, thank you for constantly reminding me to have fun and enjoy life.

To my brother, A.J., thank you for always bringing humor into my life, even during tough times, as well as for the unwavering love which is paired with it.

I am thankful for the time I had with my sister, Katie, and that she was so inspiring and loved me so deeply that I carry her memory with me every day, even years after she died. Katie and what she went through are major reasons why I became interested in science and its power to help people.

To my wonderful husband, Greg, thank you for everything. You have helped me through this PhD journey more than anyone else with your daily supportive words and actions, including editing everything from my F30 application to my thesis as well as answering innumerable questions about bioinformatics. You keep me inspired about science through your own research endeavors and our exciting conversations, and inspired about life through your faithfulness, love, and adventurous spirit. I am grateful for a spouse who treats me with such kindness.

Finally, I am in awe of the beautiful molecular systems I have had the opportunity to study, which I believe were created by a brilliant and loving God. I thank Him for creating such marvelous work, for providing me the opportunity to study it, and for helping me along the way.

ABSTRACT

Chemical modifications on RNA play critical roles in gene expression regulation. Although recent advancements have allowed for significant elucidation of their functions, challenges remain in the understanding of their functional effects and the mechanisms involved, and it is becoming increasingly apparent that it is essential to consider the role of context. Using N^6-methyladenosine (m^6A) as an example, I explore as a fundamental aspect of context the role of the modification's location in shaping its function, including the type of RNA on which it is located, its position on the RNA molecule, and the location of the methylated RNA within the cell. First, I present a study characterizing the METTL5-TRMT112 complex which deposits m^6A on 18S rRNA. I show that TRMT112 is required for METTL5 stability and that loss of METTL5 regulates gene expression at the translational level, and discuss how the position of the m^6A site on 18S rRNA near the decoding center influences its functional roles. Then, I examine the relationship between m^6A location and mRNA transcript stability. I show through analysis of public data that m^6A in the 5'UTR is associated with increased stability, and that this effect is reduced by depletion of YTHDF2. I further investigate the role of all three YTHDF proteins in transcript stability, revealing that triple knockdown causes extensive transcriptome dysregulation, potentially explained by increased P body formation. In aggregate, this work underscores the importance of location in m^6A function, highlighting the necessity of considering a modification's context to elucidate its cellular function.

CHAPTER 1: INTRODUCTION

RNA modifications and the importance of their location in gene expression regulation

Cellular ribonucleic acids (RNAs) have long been characterized to include chemically modified bases, and over the past decade, numerous investigations have highlighted the roles of these modifications in gene expression regulation. Technological advances in mass spectrometry and transcriptome-wide sequencing technologies have permitted detailed study of the location and dynamics of modifications in diverse cellular contexts. In parallel, discoveries and characterizations of related proteins that add, read, and/or remove RNA modifications have enabled mechanistic understanding of their functional consequences. In particular, such breakthroughs have allowed for the extensive study of N^6-methyladenosine (m^6A)—the most prevalent internal messenger RNA (mRNA) modification—revealing its dynamic nature and key role(s) in almost every aspect of the mRNA lifecycle, as well as in cellular, developmental, and disease processes. Abundant noncoding RNAs (ncRNAs) are also chemically modified with m^6A and other modifications, and repurposing tools and discoveries from mRNA studies have revealed their importance and potential dynamic nature as well. However, as more information has been gathered about RNA modifications, more questions have arisen, particularly when the same modification appears to play multiple, even opposing roles under different cellular contexts.

Employing the framework of structural and cellular *location* can be helpful for understanding the context-dependent functional consequences of RNA modifications. Firstly, the location of the modified region on the nucleoside is critical in determining how it changes its physical structure, non-covalent bonding, and other chemical properties. Secondly, the type of RNA molecule modified and the location of the modification on that molecule determine the

protein binding contexture and potential functional impact. Thirdly, the cellular compartment location of the modified RNA influences function by framing the biochemical milieu in which the modified RNA acts. Notably, m^6A is an exemplary modification for exploring the impact of location on RNA modification function, because it: (i) has been widely studied, especially in mRNA; (ii) exists in multiple RNA types, including mRNA and ribosomal RNA (rRNA); and (iii) because transcriptome-wide mapping technologies exist that permit studying the dynamics of its location in diverse cellular contexts. Thus, I will focus on the impact of m^6A's structural and cellular location on its function as well as describe remaining controversies and gaps in the literature.

Influences of chemical structure on functional consequences of RNA modifications

Chemical modifications alter the structure of a nucleoside by changing its size, shape, and/or charge. These nucleoside-specific structural alterations can have direct consequences on the base pairing potential of the nucleoside, the molecule's secondary structure, and RNA–protein interactions. In turn, these transcript-modifying effects can influence gene expression through impacting RNA structure, processing, localization, stability, and translation. Many of the differences in functional outcomes between different RNA modifications trace back to the chemical structure of the modification: the base being modified and how it is being changed. Modifications exist on all four RNA bases and all major RNA types, including rRNA, transfer RNA (tRNA), mRNA, long noncoding RNA (lncRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), micro RNA (miRNA), circular RNA (circRNA), and chromatin-associated RNA (caRNA). They range from base and ribose methylations to the addition of

elaborate ring structures. The most common modifications are shown in **Figure 1.1a** and listed in **Table 1.1**, with a focus on those prevalent in mRNA.

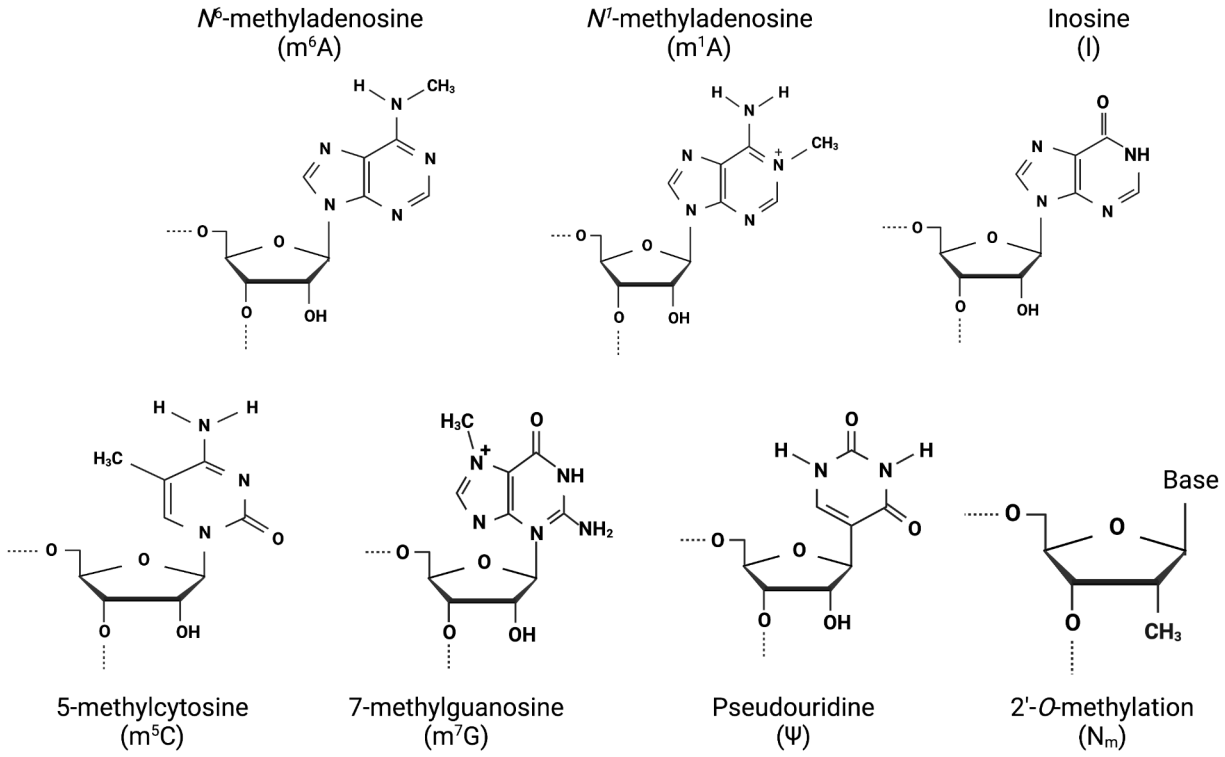


Figure 1.1. Common chemical modifications in mRNA. The chemical structures of seven of the most common modified nucleosides found in human messenger RNA (mRNA) are shown.

Table 1.1. Chemical modifications in the major cellular RNA types

Base	Modification	RNA type (with abundance on mRNA if known)	Known mRNA Effector Proteins	Notes	References
A	m^6A (N^6-methyladenosine)	**mRNA** (0.15-0.6% of A) **rRNA**	**Writers:** METTL3/METTL14, PCIF1/CAPAM (methyltransferase for cap m^6A_m) **Erasers:** FTO, ALKBH5 **Readers:** YTHDF1-3, YTHDC1-2, IGF2BP1-3, HNRNPAC, HNRNPA2B1, PRRC2A, FMRP	Most prevalent internal mRNA modification, well-studied with roles throughout mRNA life cycle. mRNA m^6A is enriched in the coding sequence (CDS) near stop codons and in the 3' untranslated region (3'UTR).	(He et al., 2021; Natchiar et al., 2017; Roundtree et al., 2017a; Shi et al., 2019)
	m^1A (N^1-methyladenosine)	**mRNA** (0.02-0.05% of A) **tRNA** **rRNA**	**Writers:** TRMT6, TRMT61A **Erasers:** ALKBH3 **Readers:** YTHDF1-3, YTHDC1	Methylation blocks Watson-Crick base-pairing, and adds a positive charge.	(Dai et al., 2018; Dominissini et al., 2016; Li et al.; Wilkinson et al., 2021)
	I (inosine)	**mRNA** **tRNA**	**Writers:** adenosine deaminases acting on RNA (ADARs)	Inosine on tRNA occurs at the wobble position (pos. 34), whereas inosine in mRNA mostly occurs in noncoding regions that contain repetitive elements. A-to-I editing often alters secondary structure. ADARs bind and modify dsRNA substrates.	(Alseth et al., 2014; Walkley and Li, 2017)

Table 1.1, continued.

Base	Modification	RNA type (with abundance on mRNA if known)	Known mRNA Effector Proteins	Notes	References
A (continued)	Other A-derived modifications: m^6A_m*, $m^{6,6}A$, hm^6A, f^6A, t^6A, m^6t^6A, ms^2t^6A, m^2A, m^1I (m^6A_m is often near the mRNA cap; the rest are mostly in tRNA and/or rRNA)				(Mathlin et al., 2020; Natchiar et al., 2017; Roundtree et al., 2017a)
C	m^5C (5-methylcytosine)	**mRNA** (0.1-0.3% of A) **tRNA** **rRNA**	**Writers:** NSUN2, DNMT2 **Erasers:** Tet enzymes **Readers:** YBX1, ALYREF	Distribution favors untranslated regions (UTRs), promotes mRNA nuclear-cytoplasmic shuttling through ALYREF binding, can stabilize mRNA through YBX1 binding	(Legrand et al., 2017; Liu et al., 2022; Roundtree et al., 2017a; Xue et al., 2020)
	Other C-derived modifications: ac4c*, m^3C*, hm^5C*, f^5C (those marked with an asterisk have been studied in mRNA but are generally lowly abundant; overall these are mostly in tRNA and/or rRNA)				(Mathlin et al., 2020; Natchiar et al., 2017; Roundtree et al., 2017a)

Table 1.1, continued.

Base	Modification	RNA type (with abundance on mRNA if known)	Known mRNA Effector Proteins	Notes	References
G	m^7G (7-methylguanosine)	**mRNA** (0.4-0.5% of G without decapping, ~0.05% with decapping) **tRNA** **rRNA**	**Writers:** METTL1/WDR4, RNMT (cap) **Erasers:** DCP2 (cap), NUDT16 (cap)	Modifies and protects mRNA at cap, internal m^7G promotes translation	(Esteve-Puig et al., 2020; Zhang et al., 2019)
	Other G-derived modifications: m^1G, m^2G, $m^{2,2}G$, Q, galQ, manQ, yW, o^2yW (mostly in tRNA)				(Mathlin et al., 2020; Roundtree et al., 2017a)
U	Ψ (Pseudouridine)	**mRNA** **tRNA** **rRNA**	**Writers:** Pus family of enzymes, DKC1	Dynamically regulated in response to stress, and likely not reversible due to installation of a C-C bond between the base and sugar	(Roundtree et al., 2017a; Schwartz et al., 2014)
	Other U-derived modifications: D, s2U, acp^3U, m^3U, m^5U, $m^5Ψ$, mcm^5U, mcm^5s^2U, $m^1Ψ$, $m^1acp^3Ψ$ (mostly in tRNA and/or rRNA)				(Mathlin et al., 2020; Natchiar et al., 2017; Roundtree et al., 2017a)

Table 1.1, continued.

Base	Modification	RNA type (with abundance on mRNA if known)	Known mRNA Effector Proteins	Notes	References
Any	2'-O-Me (2'-O-methylation)	**mRNA** **tRNA** **rRNA**	**Writers:** FBL	Modification on ribose, common at second or third base from 5' cap to increase stability, also common in CDS	(Elliott et al., 2019; Natchiar et al., 2017; Roundtree et al., 2017a)

The most prevalent internal modification on mRNA is *N*^{6}-methyladenosine (m^{6}A). The methylation in m^{6}A is installed on an exocyclic amine that participates in Watson-Crick base pairing (Roundtree et al., 2017a). Base-pairing of m^{6}A with uracil (U) forces rotation of the methyl group into the *anti* conformation, leading to a steric clash with the base, causing disfavoring of such base pairing and secondary structure (**Figure 1.2a**) (Liu et al., 2018a; Roundtree et al., 2017a). However, when m^{6}A is at a dangling end or apical loop, it can conversely stabilize the local duplex (Lee et al., 2020).

Although local changes in structure caused by m^{6}A can have consequences in RNA metabolism and gene expression, greater effects can be mediated by the modification's influence on RNA-protein interactions. Certain proteins, termed "readers", favor binding to methylated over unmethylated RNA. For example, the structural changes induced by m^{6}A, as described above, can cause the region of RNA near the methylated site to be unstructured and predispose the RNA

molecule for binding by the proteins HNRNPC/G and HNRNPA2B1 (Liu et al., 2015, 2017b; Zhou et al., 2016). More directly, m^6A-modified RNA is a preferred substrate over unmodified RNA of the same sequence for the YTH domain-containing family of proteins, including YTHDF1-3 and YTHDC1-2 (Edupuganti et al., 2017). In these cases, the YTH domain itself contains a binding pocket that specifically recognizes m^6A (Li et al., 2014; Roundtree et al., 2017a; Xu et al., 2014, 2015; Zhu et al., 2014). Other RNA binding proteins, such as IGF2BP1-3 and FMR1, prefer m^6A-modified RNAs through their tandem common RNA binding domains, though these mechanisms are still not fully understood (Edupuganti et al., 2017; Huang et al., 2018; Shi et al., 2019). These readers, in turn, act as effector proteins, interacting with other cellular factors to regulate RNA fate with consequences for overall RNA processing, stability, translation, and decay.

Another modification involving methylation of adenosine, N^1-methyladenosine (m^1A), provides an interesting comparison to m^6A. Although both modifications involve addition of a simple methyl group to adenosine, methylation at the N^1 position introduces a positive charge on the Watson-Crick interface (**Figure 1.2b**). Unlike m^6A, which can still base pair with U, though with energetic consequences that disfavor base pairing and RNA duplex formation, the presence of the methyl group and positive charge at the N^1 position in m^1A prevent hydrogen bonding and Watson-Crick base-pairing with U. This effect on base-pairing leads to blocks and/or mutations in reverse transcription, depending on the specific reverse transcriptase (Zhou et al., 2019a). The presence of the positive charge can also dramatically affect interactions of the RNA with itself and other molecules through electrostatic effects, including stabilizing interactions with the phosphate backbone (Roundtree et al., 2017a). Thus, although the m^1A modification is approximately 10-fold less abundant than m^6A in mouse and human tissues, its potency is greater (Dominissini et

al., 2016; Roundtree et al., 2017a). Interestingly, some members of the YTH family are also putative m^1A readers, with conserved residues in the hydrophobic pocket of the YTH domain important for m^6A binding also shown to be important for m^1A binding (Dai et al., 2018). Therefore, m^1A may also have functional consequences through direct structural effects as well as indirectly through reader protein binding.

Other modifications also modulate RNA function through both direct, local effects—changes in chemical properties—and effects mediated by changes in protein-RNA interactions. Like m^1A, adenosine-to-inosine (A-to-I) editing changes base-pairing properties of the target nucleoside, with effects on secondary structure, reverse transcription, and translation (Mathlin et al., 2020). A-to-I editing can also influence binding of miRNAs or RNA-binding proteins (RBPs), affecting target gene expression. Pseudouridine (Ψ) can lead to rigidification of local RNA structure by coordinating an extra water molecule compared to uracil, which can hydrogen bond with the phosphate backbone (Mathlin et al., 2020; Roundtree et al., 2017a). Additionally, 2'-O-methylation (Nm) of the ribose sugar can stabilize RNA through increasing base stacking and by masking the hydrophilic 2' hydroxyl of RNA (Roundtree et al., 2017a; Sloan et al., 2017; Yildirim et al., 2014). These effects are particularly important in the role of Nm at the 5' cap of mRNA, where it helps to stabilize mRNA. Moreover, Nm and N^7-methylguanosine (m^7G), which is also present at the cap, have specific interactions with cap-binding proteins that further regulate transcript stability, decapping, and translation (Picard-Jean et al., 2018; Topisirovic et al., 2011). Thus, the location of each modification on the base fundamentally shapes its functional impact.

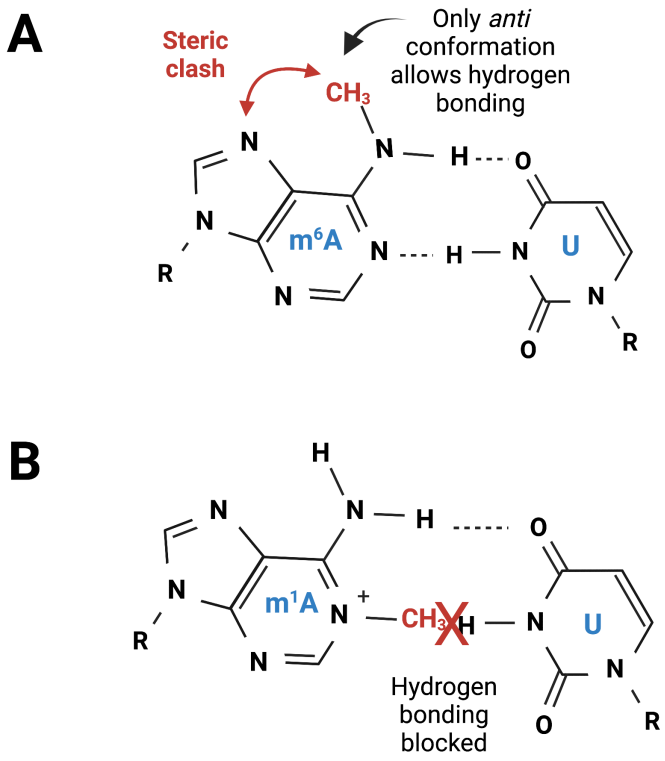


Figure 1.2. **The location of methylation influences functional consequences of modification. (A)** Watson-Crick base-pairing between N^6-methyladenosine (m^6A) and Uracil (U) is shown, illustrating the necessity for rotation around the carbon-nitrogen bond to display the methyl group in the *anti* conformation in order to allow for hydrogen bonding. This positioning leads to steric clash with the base. **(B)** Representation of interactions between N^1-methyladenosine (m^1A) and Uracil (U) needed for potential base-pairing, showing that the methylation at the N1 position of m^1A blocks one of the potential hydrogen bonds.

Modifications in various RNA types and the functional effects of their distribution patterns

The most well-known classes of RNA are those involved in protein synthesis (e.g., rRNA, tRNA, mRNA), and all three contain modifications critical for function. Although other types of RNA are also modified, comparing modifications among these three classes provides a straightforward model for understanding the role and importance of their localization within the larger structure of the molecule on functional impact. Hence, this section will overview

modifications present on these three RNA types and the functional impacts of their location, with a focus on m^6A.

The most heavily modified RNA type by number, density, and diversity of modifications is tRNA, which has 13 modifications per molecule on average (Roundtree et al., 2017a). These modifications cluster in functional areas, such as the anticodon loop, with consequences for decoding, or in other locales important for stabilizing secondary structure. The anticodon loop involves the region of the tRNA that base pairs with an mRNA codon (i.e., a three nucleotide segment), resulting in selection of the correct tRNA and addition of the correct corresponding tRNA-bound amino acid to the growing polypeptide chain during translation. Accordingly, modifications of bases in the anticodon can increase or decrease selectivity and fidelity of decoding, for instance by strengthening the codon-anticodon interaction, or by enabling non-canonical base-pairing. Notably, many tRNAs are chemically modified at position 34, corresponding to the wobble position. In fact, the original Wobble Hypothesis, which asserts some tRNAs can recognize more than one codon, leading to partial degeneracy of the genetic code, was originally inspired by common A-to-I editing at position 34 which expands base-pairing properties (Agris et al., 2018). Position 37 is another hotspot for tRNA modification, corresponding to the nucleotide at the 3' end of the anticodon. Modifications here help maintain translational accuracy and fidelity and may prevent frameshifting (Roundtree et al., 2017a). Intriguingly, although tRNA has the highest diversity of modifications, m^6A in tRNA is rare, with its presence, location, and effector proteins mostly characterized only in bacteria (Jeong et al., 2022; Roundtree et al., 2017a).

Approximately 2% of nucleotides in ribosomal RNA (rRNA) are modified (Roundtree et al., 2017a). Like tRNA, there is a general enrichment of modifications on rRNA in functional regions, such as the decoding and peptidyl transferase centers of the ribosome and the intersubunit

interface (Decatur and Fournier, 2002; Natchiar et al., 2017; Sloan et al., 2017). In the decoding and peptidyl transferase centers, chemical modifications help ensure the accuracy and efficiency of translation (Sloan et al., 2017). More generally, modifications throughout rRNA are important for rRNA biogenesis and processing, and they play a role in stabilizing ribosome structure. The most common modifications in rRNA are Nm and Ψ, which are installed by either snoRNA-guided or stand-alone enzymes, and have functional consequences in ribosome production and cell growth. Other less common rRNA modifications include the base methylations m^6A, m^1A, m^5C, and m^7G (Natchiar et al., 2017). Currently, Nm and Ψ modifications are thought to be introduced mainly during early stages of ribosomal subunit maturation; in contrast, many base modifications are often introduced later to aid subsequent ribosome assembly and translational activities (Sloan et al., 2017). As stated above, the modification's impact on ribosome function traces back to how it influences the topological potential of the corresponding nucleotide. Specifically, Nm and Ψ are thought to stabilize the rRNA structure through enhanced base stacking (Nm) and hydrogen bonding (Ψ). These and other modifications also may serve to stabilize the ribosome's higher-order structure through base-pairing, duplex or hairpin formation (i.e., secondary structure stabilization), or charge-charge interactions.

As for m^6A, there are two known sites on rRNA, m^6A4220 on 28S rRNA and m^6A1832 on 18S rRNA (Ma et al., 2019b; Natchiar et al., 2017; Sepich-Poore et al., 2022). These two m^6A sites are in distinct local environments. The m^6A4220 modification is in the 60S ribosomal subunit, close to the interface with the 40S subunit (Natchiar et al., 2017). Structural analysis suggests that the presence of the m^6A increases stacking interactions with neighboring bases, potentially stabilizing ribosome structure and subunit assembly (Ma et al., 2019b). In contrast, the m^6A1832 site is on the 40S subunit, at the top of helix 44, near the decoding center and near the binding site

of several initiation factors (Ma et al., 2019b; Rong et al., 2020; Sepich-Poore et al., 2022; van Tran et al., 2019). Thus, it may have potential consequences in influencing decoding and translation fidelity, as well as responding to cell signaling pathways that modulate translation initiation. The recent characterization of the methyltransferases responsible for these sites by me (Chapter 2) and others has allowed for the study of their functional consequences, with ZCCHC4 revealed as the methyltransferase for 28S m^6A4220 and METTL5 as the writer of 18S m^6A1832 (Ignatova et al., 2020a; Ma et al., 2019b; Sepich-Poore et al., 2022; van Tran et al., 2019). These studies have suggested the ability of both m^6A sites to regulate different aspects of translation, with location-specific functional effects of each. For example, ZCCHC4 was shown to contribute to the control of relative ribosome subunit levels and METTL5 was found to be involved in promoting translation initiation in certain contexts (Ma et al., 2019b; Rong et al., 2020).

Recently, mRNA modifications have comprised a major focus of research, with extensive studies on mRNA m^6A, which revealed its impact on gene expression regulation, paving the way for studying other modifications. Specifically, although the presence of mRNA m^6A and other modifications have been known for decades, recent breakthroughs with highly sensitive mass spectrometry methods, transcriptome-wide base-specific sequencing methods, and characterization of effector proteins have elucidated their diverse roles and importance. Notably, the discovery that mRNA m^6A could be actively demethylated by fat mass and obesity-associated protein (FTO) led to the proposal that m^6A may be dynamically regulated; such regulation would be analogous to dynamic DNA or protein modifications, with the potential for profound effects on the control of RNA fate and related cellular processes (Jia et al., 2011). Other protein effectors that enable dynamic regulation were similarly uncovered, such as the METTL3/METTL14

methyltransferase complex that installs m^6A on mRNA, and several "reader" proteins that specifically recognize the modified RNA, as further discussed in later sections.

The development of transcriptome-wide base-specific mapping technologies allowed for the study of m^6A localization along the transcript and the dynamic effects of concomitant effector proteins. Studies employing these technologies uncovered a general pattern for m^6A: enriched near stop codons and in 3'UTRs (Jia et al., 2011; Meyer et al., 2012). Similar technologies later mapped other modifications, uncovering their unique distributions on mRNA as well (**Figure 1.3**). For example, m^1A is enriched in the 5'UTR and m^7G in the 3'UTR (Dominissini et al., 2016; Zhang et al., 2019).

Based on their interactions with other parts of the transcript and cellular machinery, the localizations of modifications have functional consequences on mRNA fate, including determination of transcript processing, localization, stability, and translation (Roundtree et al., 2017a; Shi et al., 2019). Yet as the impact of the modification's location on the particular base could not be fully decoupled from its location on the transcript, there exists another layer of location-based regulation: cellular localization. As I will explore in the following sections, cellular localization (i.e., whether the transcript is in the nucleus, cytosol, another organelle, or a granule) is key for determining the context and consequences of the modification (**Figure 1.4**). Also, as before, m^6A is used as a primary example to exemplify the influence of cellular localization on its downstream function.

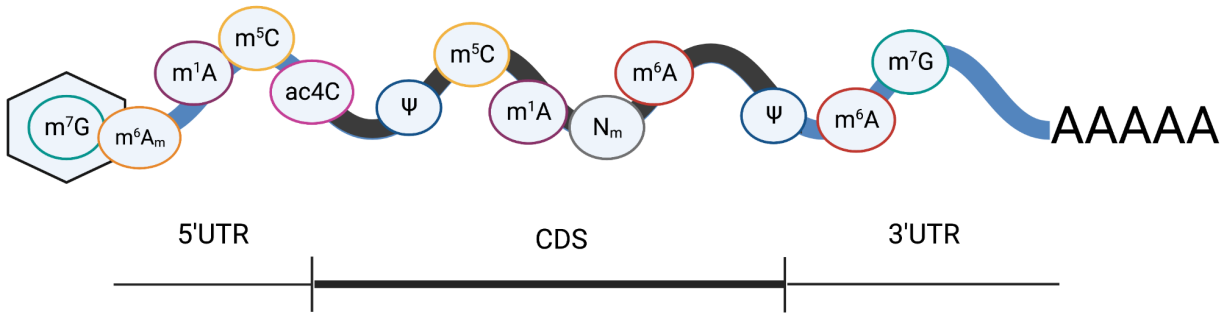


Figure 1.3. Distributions of chemical modifications along an mRNA transcript. Chemical modifications commonly found on mRNA are shown, with their location in the depiction corresponding to known locations of enrichment in human cells. The depiction is oriented by (from left-to-right) the 5' cap, 5' untranslated region (5'UTR), coding sequence (CDS), 3' untranslated region (3'UTR), and polyA tail.

The impact of cellular location: nuclear functions of mRNA *N*^{6}-methyladenosine

The deposition of m^{6}A on mRNA in mammalian cells primarily occurs co-transcriptionally in the nucleus, catalyzed by the METTL3/METTL14 methyltransferase complex, or the m^{6}A "writer" complex (Liu et al., 2014; Wang et al., 2014b). Evidence suggests that METTL3 possesses the catalytic activity for methyl transfer; in contrast, METTL14 functions largely as a scaffold involved in RNA binding (Roundtree et al., 2017a). The writer complex is associated with a regulatory complex, composed of the factors WTAP, ZC3H13, VIRMA, RBM15/15B, and HAKAI, which contributes to its activity and specificity (Lesbirel and Wilson, 2019). In particular, WTAP and ZC3H13 are important for nuclear localization of the writer complex, and WTAP is important for substrate recognition (Ping et al., 2014; Shi et al., 2019; Wen et al., 2018; Zhong et al., 2008). HAKAI plays a role in overall stabilization of the complex (Bawankar et al., 2021; Lesbirel and Wilson, 2019). VIRMA and RBM15/15B guide localization of m^{6}A deposition, with RBM15/15B binding and facilitating methylation of uracil-rich transcripts, and VIRMA guiding deposition of m^{6}A to the 3'UTR (Patil et al., 2016; Yue et al., 2018). Location of m^{6}A deposition

along a transcript is also determined by recruitment of METTL14 to histone H3 trimethylation at lysine 36 (H3K36me3), a histone modification associated with gene bodies, which results in greater installation of m^6A in coding sequences (CDSs) and 3'UTRs (Huang et al., 2019). These processes provide insight into the mechanisms that affect the distribution, but the complete picture of determinants of m^6A location, and the functional consequences of these methylated regions, are still being investigated.

Many of the functional effects of m^6A in the nucleus are mediated by nuclear readers. Of the YTH domain-containing family (m^6A readers), only one is primarily found in the nucleus of mammalian cells: YTHDC1 (also known as YT521-B) (Roundtree et al., 2017a). YTHDC1 is involved in regulating alternative splicing through interactions with the splicing-related SR-protein family, in mediating nuclear export of methylated mRNAs through interaction with the splicing factor and nuclear export adaptor SRSF3, and in accelerating transcript decay (Huang et al., 2019; Roundtree et al., 2017b; Shima et al., 2017; Xiao et al., 2016). Notably, this latter role in the decay of methylated RNA has been shown recently to have profound consequences beyond just mRNA, as the YTHDC1-mediated decay of chromatin-associated regulatory RNAs (carRNAs) was found to regulate chromatin state and transcription (Liu et al., 2020a; Roundtree et al., 2017a). Other than YTHDC1, three members of the heterogeneous nuclear ribonucleoprotein (HNRNP) family were also shown to bind to transcripts in a manner regulated by m^6A presence, based on a structural switch, with consequences affecting splicing outcomes (Liu et al., 2015, 2017b; Zhou et al., 2016). An additional reader, fragile X mental retardation protein (FMRP), was recently found to bind m^6A-modified mRNA targets in the nucleus and facilitate their nuclear export (Hsu et al., 2019). Thus, the direct functions of nuclear m^6A readers comprise transcript export, splicing, and decay modulation, with indirect effects on chromatin state.

Completing the dynamic loop, demethylases, or "erasers," in the nucleus can remove m^{6}A, mitigating the functional consequences for transcripts that rely on its presence. FTO, the first characterized mRNA m^{6}A eraser, localizes predominantly to the nucleus in some cell lines and to the cytosol in others (Hsu et al., 2019; Wei et al., 2018). In the nucleus, it predominantly mediates demethylation of m^{6}A from mRNA and chromatin-associated RNA (Wei et al., 2018, 2022). Another eraser, ALKBH5, appears to primarily be localized in nuclear speckles, supporting the idea that nascent RNAs are one of its major substrates (Qu et al., 2022).

These nuclear writers, readers, and erasers initiate and shape the journey of m^{6}A in determining transcript fate in mammalian cells. The writer complex installs m^{6}A, determining which transcripts are methylated and where. Nuclear erasers may remove it before or after its interaction with readers, which in turn, can determine how the methylated transcript is processed, and whether it experiences decay before reaching the cytosol or is fast-tracked for nuclear export. Once it reaches the cytosol, the trajectory of the methylated transcript will be shaped by the different set of effector proteins (and their differing levels) located there compared to in the nucleus. As discussed in the following sections, two of the major cytosolic processes influenced by m^{6}A processing are translation and decay, and focus is given to how the methylation's location on the transcript affects its functional consequences.

The impact of cellular location: cytosolic roles of mRNA *N*^{6}-methyladenosine in translation

A major way in which m^{6}A impacts protein expression is through its effects on translation. As seen above, the specific effect it provides and associated mechanisms by which it acts revolve around the location of the modification. As previously discussed, there are two m^{6}A sites on rRNA, and the loss of either has been shown to cause dysregulation of translation, with consequences in

cell proliferation and other cellular processes (Ignatova et al., 2020a; Ma et al., 2019b; Sepich-Poore et al., 2022). Nonetheless, the number of m^6A sites is much higher on mRNA than rRNA, with over 10,000 mRNA sites mapped in multiple cell lines (Hu et al., 2022). mRNA m^6A has varied effects on translation, which can be mediated directly or through effector proteins, and its impact is determined in part by the regional localization of the m^6A along the transcript (Meyer, 2019).

m^6A in the 5'UTR has been shown to be particularly important for regulating translation in contexts of stress, during which major changes occur in the 5'UTR to tightly control translation for a timely and regimented response to the stressor. For example, a major mechanism of translation regulation in response to stress, the Integrated Stress Response (ISR), involves the global suppression of cap-dependent translation. Cap-dependent translation generally involves the binding of the translation initiation factor eIF4E to the mRNA cap, followed by binding of several other factors, eIF4A, eIF4G, eIF4F, and the preinitiation complex (PIC) (Meyer, 2019; Uttam et al., 2018). The PIC includes the 40S subunit, additional translation factors, and a ternary complex composed of eIF2, GTP, and initiator methionine tRNA (Uttam et al., 2018). PIC recruitment is followed by scanning of the mRNA 5'UTR and joining of the 60S subunit at the start codon to allow for translation elongation. In the context of stress, however, the activation of one of four kinases (HRI, PKR, PERK, or GCN2) results in the phosphorylation of the α subunit of eIF2, reducing the formation likelihood of the ternary complex and suppressing global cap-dependent translation (Meyer, 2019).

Although global cap-dependent translation is suppressed following eIF2α phosphorylation, a surprising amount of protein production persists–an observation motivating investigations of cap-independent translation mechanisms (Coots et al., 2017; Uttam et al., 2018). A key study

implicated m^{6}A in this process, when it was found that METTL3 depletion caused greater reductions in protein synthesis upon activation of eIF2α phosphorylation than by mTORC1 inhibitor, torin1, treatment alone (i.e., 80% reduction with METTL3 depletion compared to 50% with torin1 alone) (Coots et al., 2017). Complementing this finding, another study documented how several types of cellular stresses (e.g., heat shock) cause dramatic increases in m^{6}A levels in select 5'UTRs (Dominissini et al., 2012). Additional studies also demonstrated through *in vitro* assays that 5'UTR m^{6}A could promote translation in the absence of cap-binding proteins (Meyer et al., 2015; Zhou et al., 2015). Applications in cells showed that increased methylation of the 5'UTR of *HSP70*, a transcript important in heat shock response, in turn, increased its translation (Meyer et al., 2015; Zhou et al., 2015). The initiation factor eIF3 was found to bind preferentially to m^{6}A-containing transcripts, and to bind 5'UTR m^{6}A methylation sites specifically in cells, providing insights into a mechanism by which m^{6}A promotes cap-independent translation (Meyer et al., 2015). Notably, this mechanism relies on an accessible 5' end and scanning of the 5'UTR, so only m^{6}A sites within the 5'UTR appear able to utilize it, further emphasizing the functional consequences of m^{6}A's location along a transcript (Meyer et al., 2015).

Later studies revealed that eIF3-mediated translation initiation was not the only mechanism by which m^{6}A could promote cap-independent translation. The m^{6}A reader YTHDF3 was shown to mediate cap-independent translation by binding to 5'UTR m^{6}A in mRNA transcripts, including the transcript for YTHDF3 itself in an auto-regulation mechanism (Chang et al., 2020). Cap-independent translation through YTHDF3 was also shown to drive extensive translation of circRNAs, which lack a 5' end, by interaction of YTHDF3 with initiation factor eIF4G2 (Yang et al., 2017). YTHDF3's family member, YTHDF2, which increases in abundance in response to heat shock, was also implicated in the stress response through its nuclear translocation and

protection of 5'UTR m^{6}A sites from demethylation by FTO (Zhou et al., 2015). However, whether or how YTHDF2 may play a more direct role in translation of the methylated transcripts remains unclear.

5'UTR m^{6}A sites can also regulate the translatome through determining open reading frame (ORF) selection. Qian and colleagues demonstrated that the presence of m^{6}A in the 5'UTR of activating transcription factor 4 (ATF4), a key gene for ER stress responses, governed ribosome scanning and start site selection, highlighting it as an essential part of controlling ORF selection that can be dynamically regulated during stress (Zhou et al., 2018). They found that methylation of ATF4's uORF2 was decreased during starvation, alleviating m^{6}A-mediated ribosome retention at upstream ORFs and allowing for increased translation of the main ATF4 ORF important in the stress response (Zhou et al., 2018). Importantly, global profiling of initiating ribosomes revealed that the involvement of 5'UTR m^{6}A in initiation site selection may be widespread (Zhou et al., 2018).

m^{6}A in the coding sequence (CDS) also influences multiple aspects of translation. Increased m^{6}A deposition in the CDS was found to be associated with slower processing of RNA polymerase and was correlated with reduced protein production from reporters (Choi et al., 2016; Slobodin et al., 2017). Experiments *in vitro* and in *E. coli* translation systems also confirmed that CDS m^{6}A causes translation inhibition (Choi et al., 2016; Slobodin et al., 2017). Mechanistic studies further revealed that the presence of m^{6}A can disrupt tRNA accommodation during ribosome processing, resulting in slowed translation kinetics (Choi et al., 2016). Separately, binding of the m^{6}A reader FMRP to the CDS of target transcripts has also been shown to inhibit translation, a mechanism which is thought to be related to translational changes associated with

autistic features in patients with Fragile X Syndrome who lack functional FMRP (Darnell et al., 2011).

In contrast, m^6A in the CDS can also increase translation through other mechanisms. For instance, it was shown to relieve ribosome stalling at GAN codons (Barbieri et al., 2017; Darnell et al., 2011). Strikingly, Qian and colleagues found that, although transcripts with CDS m^6A show reduced ribosome occupancy, depletion of METTL3 resulted in a further reduction in translation efficiency (TE) (Mao et al., 2019), paradoxically suggesting that CDS m^6A presence was still necessary to support translation. A thorough investigation then revealed that CDS m^6A typically occurs on highly structured transcripts with lower TE. Mechanistically, it was found that the presence of m^6A in the CDS leads to recruitment of the helicase-containing reader YTHDC2, which can resolve secondary structures, resulting in translation promotion (Mao et al., 2019). Thus, the location of CDS m^6A relative to a transcript's secondary structure can also influence the resultant functional outcome.

In the 3'UTR, m^6A has been shown to upregulate translation through the preferential binding of effector proteins. For example, METTL3 can recognize 3'UTR m^6A sites in the cytoplasm, bind to them, and associate with ribosomes to promote translation (Lin et al., 2016; Mao et al., 2019). The mechanism for this process was found to involve mRNA looping via the interaction of METTL3 with initiation factor eIF3h (Choe et al., 2018). The m^6A reader YTHDF1 also preferentially binds 3'UTR m^6A and can promote translation through its interaction with multiple translation initiation factors, including eiF3, thereby increasing translation efficiency of its targets (Choe et al., 2018; Wang et al., 2015). Other m^6A readers, YTHDF3 and FMRP, were also shown to mediate translation promotion through interplay with YTHDF1 (Edupuganti et al.,

2017; Shi et al., 2017a). Thus, through multiple mechanisms, effector protein interactions with 3'UTR-localized m^6A can regulate the degree of translation.

It is important to note that, in addition to translation efficiency, another major determinant of the protein expression of a transcript is the abundance of the transcript. Complex pathways controlling mRNA stability and decay dictate the amount of transcript available to be translated, and these processes are often impacted by m^6A, as discussed in the next section.

The impact of cellular location: cytosolic roles of mRNA *N*^{6}-methyladenosine in mRNA stability and decay

mRNA abundance is regulated through a balance of stabilization and decay, with consequences for protein expression and transcriptome turnover. Predominant cytoplasmic mechanisms of mRNA decay involve degradation from one of the two ends (5'UTR or 3'UTR) (Łabno et al., 2016), which necessitates protection at these ends to increase stability. This idea underpins the existence of a 5'-terminal mRNA methylguanosine cap structure, which is bound by an uncommon 5'-to-5' triphosphate linkage to the rest of the mRNA molecule, wherein the second and third bases are often further modified with 2'-O-methylation or N^6,2′-O-dimethyladenosine (m^6A_m) to increase stability (Ramanathan et al., 2016; Roundtree et al., 2017a). This cap must be removed prior to 5'-to-3' degradation of the mRNA by the Xrn family of RNases (Łabno et al., 2016). At the 3' end, there is a poly(A) tail, which increases the distance from the end of the molecule to the CDS while providing a template for the binding of various cellular factors, including those that promote stability or degradation (Nicholson and Pasquinelli, 2019; Torabi et al., 2021)). Perhaps the most famous of these factors are the poly(A)-binding proteins (PABPs), which can bind to cap-binding factors, causing looping of the mRNA to facilitate translation while

preventing access to exonucleases for decay. Still, mRNA can be decayed even in these circumstances through endonucleolytic cleavage (Łabno et al., 2016).

YTHDF2 was the first mRNA m^{6}A reader shown to *promote* decay of methylated transcripts (Łabno et al., 2016; Wang et al., 2014a), operating through the recruitment of the CCR4-NOT deadenylase complex, which accelerates mRNA degradation from the 3' end (Du et al., 2016). Later work uncovered that YTHDF2 can also mediate decay through the promotion of endoribonucleolytic cleavage, as YTHDF2 interacts with RNaseP/MRP, an endonuclease complex, via heat-responsive protein 12 (HRSP12), an adaptor protein (Lee et al., 2020; Park et al., 2019). Critically, this mechanism requires the presence of an HRSP12 binding site in the methylated mRNA. Given the effect of YTHDF2 in destabilizing many of its targets, and the observation that many transcripts are decapped to allow for additional decay from the 5' end after deadenylation, it has been proposed that YTHDF2 may also interact with factors related to decapping or 5' decay (Lee et al., 2020). However, substantive evidence to support such a model has been lacking. Furthermore, analysis of data from a recent single-base-resolution m^{6}A sequencing method, m^{6}A-SAC-seq, uncovered a potential role for YTHDF2 in stabilizing transcripts with 5'UTR m^{6}A (Hu et al., 2022; Lee et al., 2020), which is further investigated in Chapter 3 of this dissertation.

YTHDF3 has also been shown to interact with YTHDF2 to mediate decay (Edupuganti et al., 2017; Shi et al., 2017a). Recently, a controversial model asserted that all three YTHDF proteins provide redundant functions in mediating mRNA decay (Lasman et al., 2020; Zaccara and Jaffrey, 2020). This proposal was largely based on the argument that YTHDF1-3 share almost identical targets, as well as the finding that triple knockdown of YTHDF1-3 leads to greatly increased mRNA stabilization compared to depletion of any one of the proteins alone, particularly for m^{6}A-

methylated transcripts. However, other lines of evidence suggest that this model is likely oversimplified. For example, although tethering of YTHDF1 to the 3' end of a reporter leads to a slight reduction in mRNA level, it leads to a much more significant increase in protein level and translation efficiency (Wang et al., 2015). Moreover, another targeting platform utilizing inactive Cas13b fused to the N-terminal domain of either YTHDF1 or YTHDF2 showed large decreases in both mRNA and protein levels when targeting with YTHDF2, but only minor mRNA decreases with enhanced translation when targeting with YTHDF1 (Rauch et al., 2018).

In addition to YTHDF proteins, other cytosolic m^6A reader proteins also regulate mRNA stability. The insulin-like growth factor 2 mRNA-binding proteins, IGF2BP1-3, stabilize target mRNA in an m^6A-dependent manner, with preferred binding in the 3'UTR, through a mechanism likely involving cofactors HuR and MATR3 (Huang et al., 2018). In the context of cancer, IGF2BPs have also been observed to stabilize oncogenic methylated mRNAs (Huang et al., 2018). Additionally, in the context of neurological disease and glial development, a novel reader Prrc2a was shown to bind and stabilize a critical m^6A-methylated transcript required for myelination (Huang et al., 2018; Wu et al., 2019). These studies among others have demonstrated the link between m^6A transcript location with mRNA abundance and pathology.

On a larger physical scale, mechanisms of mRNA stabilization often involve cellular relocalization. mRNA bound by RNA-binding proteins, or messenger ribonucleoproteins (mRNPs), can dynamically shuttle between polysomes and intracellular phase-separated compartments like processing bodies (P bodies) and stress granules. Early in its functional characterization, YTHDF2 was shown to localize to P bodies, which are also enriched for many decay-related factors, including decapping enzymes and 5'-to-3' exonucleases, and this localization was implicated in its mechanism of promoting mRNA decay (Lee et al., 2020; Wang

et al., 2014a). However, recent studies have presented accumulating evidence consistent with P bodies as storage granules containing inactive decapping enzymes and translationally repressed mRNAs, whereas others have presented evidence that, within P bodies, certain transcripts may be stabilized and other decayed (Luo et al., 2018; Wang et al., 2018). One imaging study even observed strong m^6A signals but weak polyA signals in P bodies, raising the possibility that YTHDF2 targets could be stored in P bodies in deadenylated form (Fu and Zhuang, 2020). However, more investigative work is needed to help elucidate the repertoire of mechanisms of stability and decay within P bodies, especially relating to methylated mRNAs and the role of m^6A readers, a topic discussed further in Chapter 3.

Interestingly, both m^6A methylated transcripts and YTHDF proteins have also been found to be enriched in stress granules, suggesting another mechanism by which m^6A methylation may be important in the cellular response to stress (Anders et al., 2018; Fu and Zhuang, 2020; Ries et al., 2019). The disordered N-terminal domains of YTHDF proteins were shown to be useful to drive phase separation in response to stress (Ries et al., 2019). Tethering YTHDF1 to the 3'UTR of a mRNA reporter also resulted in increased translation of the latter after recovery from sodium arsenite treatment, suggesting a model in which the YTHDF1-bound mRNA may be stored in stress granules in a conformation primed for rescued translation after recovery (Ries et al., 2019; Wang et al., 2015). Likewise, IGF2BPs have also been found to colocalize with stress granule markers during heat shock and to return to polysome fractions after recovery, suggesting a potentially similar functional mechanism for storage of target transcripts during stress (Huang et al., 2018).

These findings emphasize the mechanistic interconnectedness of translation, transcript stabilization, and mRNA decay, as well as the roles of m^6A and its protein effectors in regulating

the movement and fates of methylated transcripts. In general, less translationally active mRNAs are sequestered to P bodies or stress granules, where they can be stored and stabilized in a form poised for future translation upon recovery, or alternatively, perhaps for later decay. Translation and decay are interlinked at a more local level through PABP-mediated cap binding and transcript looping, which promotes translation while preventing decay; moreover, m^6A effectors, such as YTHDF1 and METTL3, have been shown to promote such a conformation. Another prominent model based on pulse-chase metabolic labeling also suggests that YTHDF1 and YTHDF3 may bind mRNAs before YTHDF2 to promote mRNA translation followed by degradation (Shi et al., 2017a; Wang et al., 2015). Thus, although regulation of stability and decay is a prominent function of m^6A methylation on mRNA, the understanding of it cannot be decoupled from other mRNA-related processes such as translation and cellular trafficking. Furthermore, although evidence supports that the loss of all three YTHDF proteins leads to greatly increased mRNA stabilization, especially of methylated transcripts, the roles of each YTHDF protein in stabilization or decay are likely more nuanced, involving more additional layers of regulation than simple functional redundancy. Continued investigations into the mechanisms by which m^6A methylation results in translation, stabilization, or decay under different cellular and pathologic conditions are warranted.

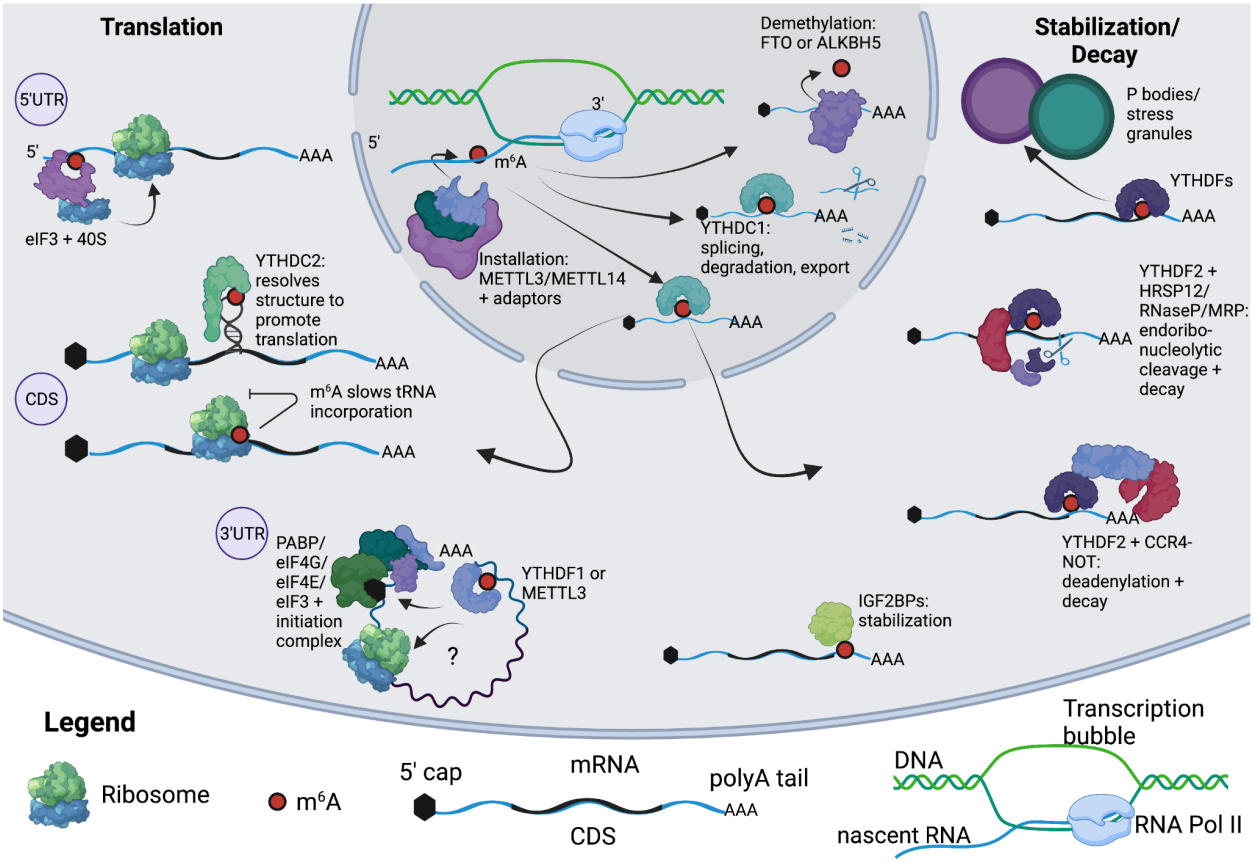


Figure 1.4. The critical role of cellular and molecular location in mRNA *N*^{6}-methyladenosine (m^{6}A) function. The functional consequences of m^{6}A (red circle) are determined in part by the localization of the methylated transcript in the cell and the location of the methylation along the transcript. In the nucleus, m^{6}A is installed co-transcriptionally by the METTL3/METTL14 complex and can be dynamically demethylated by erasers FTO or ALKBH5. Selective binding of nuclear reader YTHDC1 can influence transcript splicing, degradation, and export. Once in the cytosol, the location of the methylation on the transcript can influence its role in regulation of translation (left) or stability and decay (right). Methylation in the 5'UTR can recruit eIF3 and the 40S subunit to promote cap-independent translation, especially in contexts of stress. Methylation in the CDS can increase translation efficiency by recruiting reader YTHDC2 to resolve secondary structure or can decrease translation efficiency by directly slowing tRNA incorporation. Methylation in the 3'UTR can recruit METTL3 or YTHDF1 which interact with translation factors to promote translation. Common reader proteins mediate m^{6}A contribution to transcript stability. IGF2BPs have been shown to bind the 3'UTR to stabilize target transcripts. YTHDF2 can destabilize its targets through recruiting the CCR4-NOT deadenylase complex, which can accelerate transcript decay from the 3' end, or by recruiting the RNaseP/MRP complex via adaptor protein HRSP12, which results in endonucleolytic cleavage. Finally, YTHDF1-3 have been shown to localize to P-bodies and stress granules, sites of mRNA storage. Created with BioRender.com.

Summary, outlook, and dissertation outline

Although the presence of RNA chemical modifications has been documented for decades, studies continue to reveal the complex roles that modifications have in regulating gene expression, and the importance of their location in the involved mechanisms. Technological and methodological advances have enabled the mapping of modifications on more RNA types and for more kinds of modifications, revealing differences in function underpinned by the modification's location on the molecule. In parallel, a deeper understanding of cellular biology has informed the understanding of the role of cellular localization in the functional consequences of mRNA modifications. On multiple levels, m^6A provides a well-studied example for how molecular and physical location guide the function of the chemical modification.

Continuing the theme in this dissertation, a strong focus is placed on molecular and cellular location when characterizing the functional consequences of m^6A. Chapter 2, which has been published (Sepich-Poore et al., 2022), describes a comprehensive characterization of the previously unknown writer complex for the 18S m^6A1832 site, METTL5/TRMT112, and investigation of the functional consequences of the 18S m^6A site in regulating translation. Chapter 3, which has been submitted for publication (Zou et al., 2022), then examines the role of m^6A's location along the mRNA molecule and within the cell in regulating mRNA stability. Then, Chapter 4, which is unpublished, provides a summary discussion of the results of this work in the context of recent findings, followed by presenting key unanswered questions remaining to be explored. Finally, the Appendix, which is unpublished, includes details on progress made so far in the development of a method to comprehensively map another modification on mRNA, N^3-methylcytidine (m^3C), at single-base resolution. Overall, this dissertation advances the field's

knowledge of RNA chemical modifications, particularly m^6A, and the importance of location in elucidating their cellular functions.

CHAPTER 2: THE METTL5-TRMT112 N^6-METHYLADENOSINE METHYLTRANSFERASE COMPLEX REGULATES mRNA TRANSLATION VIA 18S rRNA METHYLATION

Preface Note:

This section (Chapter 2) is reproduced, with minor modifications, from Sepich-Poore, C., Zheng, Z., Schmitt, E., Wen, K., Zhang, Z.S., Cui, X.-L., Dai, Q., Zhu, A.C., Zhang, L., Sanchez Castillo, A., et al. (2022). The METTL5-TRMT112 N-methyladenosine methyltransferase complex regulates mRNA translation via 18S rRNA methylation. J. Biol. Chem. 298, 101590. https://doi.org/10.1016/j.jbc.2022.101590.

The study began as a collaborative effort between me and Prof. Sigrid Nachtergaele, overseen by Prof. Chuan He. Thus, some of the initial experiments for biochemical characterization of METTL5 were performed by Prof. Nachtergaele and me together, including biochemical fractionation to identify METTL5, development and proliferation analyses of HeLa METTL5 knockout lines, identification and validation of the METTL5-TRMT112 interaction, analysis of the effect of TRMT112 on METTL5 stability, *in vitro* methyltransferase assays, studies of METTL5 cellular localization, METTL5 CLIP-seq, and METTL5 knockout versus wild type RNA-seq. I performed structural analyses of human variants, and Prof. Nachtergaele performed studies of their effects on TRMT112 binding. I created the HepG2 knockout lines and monitored their growth in collaboration with Emily Schmitt. I also performed the 18S probe purification and mass spectrometry, m^6A-seq, polysome profiling, and ribosome profiling experiments, as well as all related bioinformatic analyses. Initial phases of development of the *METTL5* knockout mouse model were planned by me, Prof. Nachtergaele, Prof. He, and Dr. Zhong Zheng, and were carried

out by the staff at the Gene Targeting & Transgenic Facility at Janelia Research Campus at HHMI. The mice were later bred and cared for by me, Dr. Zheng, and the University of Chicago Animal Resources Center, and their phenotypic characterization was performed by me with assistance from Dr. Zheng, Arantxa Sanchez Castillo, and Linda Zhang. I performed all biochemical and sequencing experiments using mouse tissues, including RNA-seq, ribosome profiling, and 18S probe purification followed by mass spectrometry, as well as the related data analysis. Prof. Xiaoxi Zhuang and Kailong Wen conducted mouse behavior experiments. Dr. Zijie Scott Zhang and Dr. Xiao-Long Cui provided bioinformatic advice and support. Dr. Qing Dai synthesized probes. Dr. Haiyan Tan and Dr. Junmin Peng performed protein mass spectrometry experiments. Dr. Allen Zhu assisted with ribosome profiling experiments. Prof. Nachtergaele and I performed data visualization and wrote the manuscript, with review and editing input from Dr. Allen Zhu, Emily Schmitt, Dr. Haiyan Tan, Dr. Junmin Peng, Dr. Xiaoxi Zhuang, and Prof. Chuan He. All work was guided and overseen by Prof. Nachtergaele and Prof. He.

ABSTRACT

Ribosomal RNAs (rRNAs) have long been known to carry chemical modifications, including 2'*O*-methylation, pseudouridylation, N^6-methyladenosine (m^6A), and $N^{6,6}$-dimethyladenosine. Although the functions of many of these modifications are unclear, some are highly conserved and occur in regions of the ribosome critical for mRNA decoding. Both 28S rRNA and 18S rRNA carry single m^6A sites, and although the methyltransferase ZCCHC4 has been identified as the enzyme responsible for the 28S rRNA m^6A modification, the methyltransferase responsible for the 18S rRNA m^6A modification has remained unclear. Here, we show that the METTL5-TRMT112 methyltransferase complex installs the m^6A modification

at position 1832 of human 18S rRNA. Our work supports findings that TRMT112 is required for METTL5 stability and reveals that human METTL5 mutations associated with microcephaly and intellectual disability disrupt this interaction. We show that loss of METTL5 in human cancer cell lines and in mice regulates gene expression at the translational level; additionally, *Mettl5* knockout mice display reduced body size and evidence of metabolic defects. Although recent work has focused heavily on m^6A modifications in mRNA and their roles in mRNA processing and translation, we demonstrate here that deorphanizing putative methyltransferase enzymes can reveal previously unappreciated regulatory roles for m^6A in noncoding RNAs.

INTRODUCTION

Chemical modifications on RNA are a critical facet of gene expression regulation. Historically, modifications on tRNA and rRNA have been thought to have high stoichiometry and be relatively static, whereas work in the last decade on mRNA modifications suggests they are often sub-stoichiometric and more dynamic (Roundtree et al., 2017a). rRNA is heavily modified with numerous chemical marks, including pseudouridine, 2'*O*-methylation (2'*O*Me), N^7-methylguanosine (m^7G), N^1-methyladenosine (m^1A), N^6-methyladenosine (m^6A), and $N^{6,6}$-methyladenosine ($m^{6,6}A$) (Natchiar et al., 2017; Piekna-Przybylska et al., 2008; Sergiev et al., 2018). Although these modifications are thought to play critical structural roles and many of the regulatory enzymes have been identified, it is often difficult to assign specific functions to individual modifications due to the numerous interactions between the three rRNAs and more than 80 protein components that form the ribosome. Although some modifications play structural roles in ribosome assembly, others may regulate translation of specific transcripts. Disruption of rRNA modification processes has been implicated in a class of developmental disorders called

ribosomopathies (Bowman et al., 2012; Freude et al., 2004; Ramser et al., 2004; Splendore et al., 2000). Interestingly, though the ribosome is ubiquitously essential for translating protein, ribosomopathies often manifest as tissue-specific disorders, the molecular mechanisms of which we do not understand in many cases (Farley-Barnes et al., 2019).

The METTL protein family is a class of S-adenosyl-methionine-dependent methyltransferases, with over thirty family members that methylate DNA (Hao et al., 2020), RNA (Alexandrov et al., 2002; Liu et al., 2014; Pendleton et al., 2017; Xu et al., 2017), and protein (Shimazu et al., 2014; Tooley et al., 2010) substrates. Some, such as METTL3, METTL14, and METTL16, have well-characterized functions as RNA m^6A methyltransferases (Liu et al., 2014; Pendleton et al., 2017), but others remain poorly understood. Notably, mutations in many of these enzymes, including those whose functions are unknown, have been implicated in human diseases such as developmental abnormalities and cancers (Baltz et al., 2012; Hu et al., 2019; Sun et al., 2020). Revealing METTL protein substrate specificity, activity, and function are critical first steps towards understanding how mutations in these enzymes cause human disease. More specifically, mutations in METTL5 have been implicated in developmental abnormalities including microcephaly, intellectual disabilities, and attention deficit hyperactivity disorder (ADHD), but until recently very little was known about METTL5 function (Hu et al., 2019; Richard et al., 2019). We were intrigued by this connection when METTL5 appeared in proteomics experiments aimed at identifying novel proteins with methyltransferase activity.

While this work was in progress, multiple studies revealed that a complex containing METTL5 and TRMT112 m^6A-methylates 18S rRNA in flies, mice, and humans (Ignatova et al., 2020a; Leismann et al., 2020; Rong et al., 2020; van Tran et al., 2019). METTL-5 in *C. elegans* also carries out this function, but not in the context of a complex with TRMT112, as *C. elegans*

lack a TRMT112 homologue (Liberman et al., 2020; Sendinc et al., 2020). Interestingly, the role of METTL5 in protein translation differs across the different model organisms studied. Consistent with these reports, we found that METTL5 forms a complex with TRMT112 to m^6A methylate 18S rRNA in human cell lines and in mice. We further found that TRMT112 is critical for stabilizing METTL5 at the protein level, and that depletion of TRMT112 is sufficient to reduce METTL5 protein levels. Catalytically inactive METTL5 mutants can retain this critical association with TRMT112, but METTL5 mutations derived from human patients dramatically reduce this interaction. Although we did not see global changes in protein translation upon METTL5 depletion, we found evidence of dysregulated translation of specific transcripts. To complement our cellular studies, we generated *Mettl5* knockout mice (*Mettl5*$^{-/-}$), and validated the loss of 18S rRNA m^6A1832 in tissues from these mice. Consistent with Ignatova *et al.* and Wang *et al.*, we observed smaller body size in our *Mettl5*$^{-/-}$ mouse model (Ignatova et al., 2020a; Wang et al., 2022). We did not observe the significant behavioral changes that have been previously described, but RNA sequencing of mouse tissues revealed evidence of lipid metabolic defects that have not yet been reported. Ribosome profiling of mouse livers suggested that the effects of Mettl5 loss predominantly occur at the level of translation. Altogether, we propose that METTL5 may regulate the translation of transcripts associated with lipid metabolism, resulting in metabolic dysregulation that may play a role in the developmental phenotypes seen in human patients.

RESULTS

METTL5 is an m^{6}A methyltransferase stabilized by TRMT112

METTL5 is a member of the METTL family of S-adenosyl-methionine-dependent methyltransferases, which is not found in yeast but conserved in higher eukaryotes ranging from *C. elegans*, to mice, to humans. Although the activities and functions of some METTL proteins have been elucidated over the last decade (Liu et al., 2014; Pendleton et al., 2017; Xu et al., 2017), many remain poorly understood. RNA methyltransferases, particularly the METTL3-METTL14 complex, have been demonstrated to have numerous cellular functions through their methyltransferase activity, which raises the question as to how less well-characterized METTL proteins might regulate cellular processes. METTL5 drew our attention during a biochemical screen that was aimed at identifying novel RNA methyltransferase activity through biochemical fractionation followed by a liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS)-based assay (see **Supplementary Note** for more details). We surveyed commonly used cancer cell culture lines by western blot and found that HeLa, HEL, and K562 cells showed relatively higher METTL5 expression (**Figure 2.1A**). HeLa cells are the only adherent line of these three and are easily transfected, making biochemical and imaging studies straightforward. Thus, we generated METTL5 knockout (KO) HeLa lines using CRISPR-Cas9 and two sets of two guide RNAs to generate deletions across two different exons (**Figure 2.2A**). After puromycin selection, single clones were isolated, expanded, and tested for METTL5 expression by western blot (**Figure 2.1B, Figure 2.2A**). Through this process, we isolated both KO lines and clones in which METTL5 expression was unaffected (e.g., clones 4 and 6), which serve as controls throughout this work.

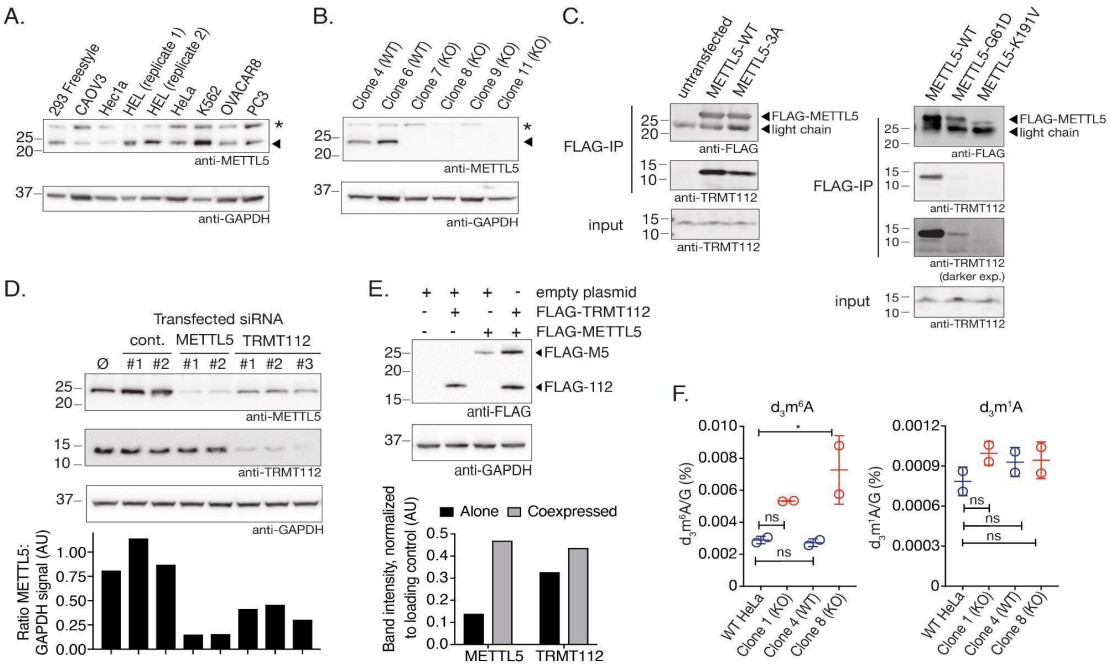
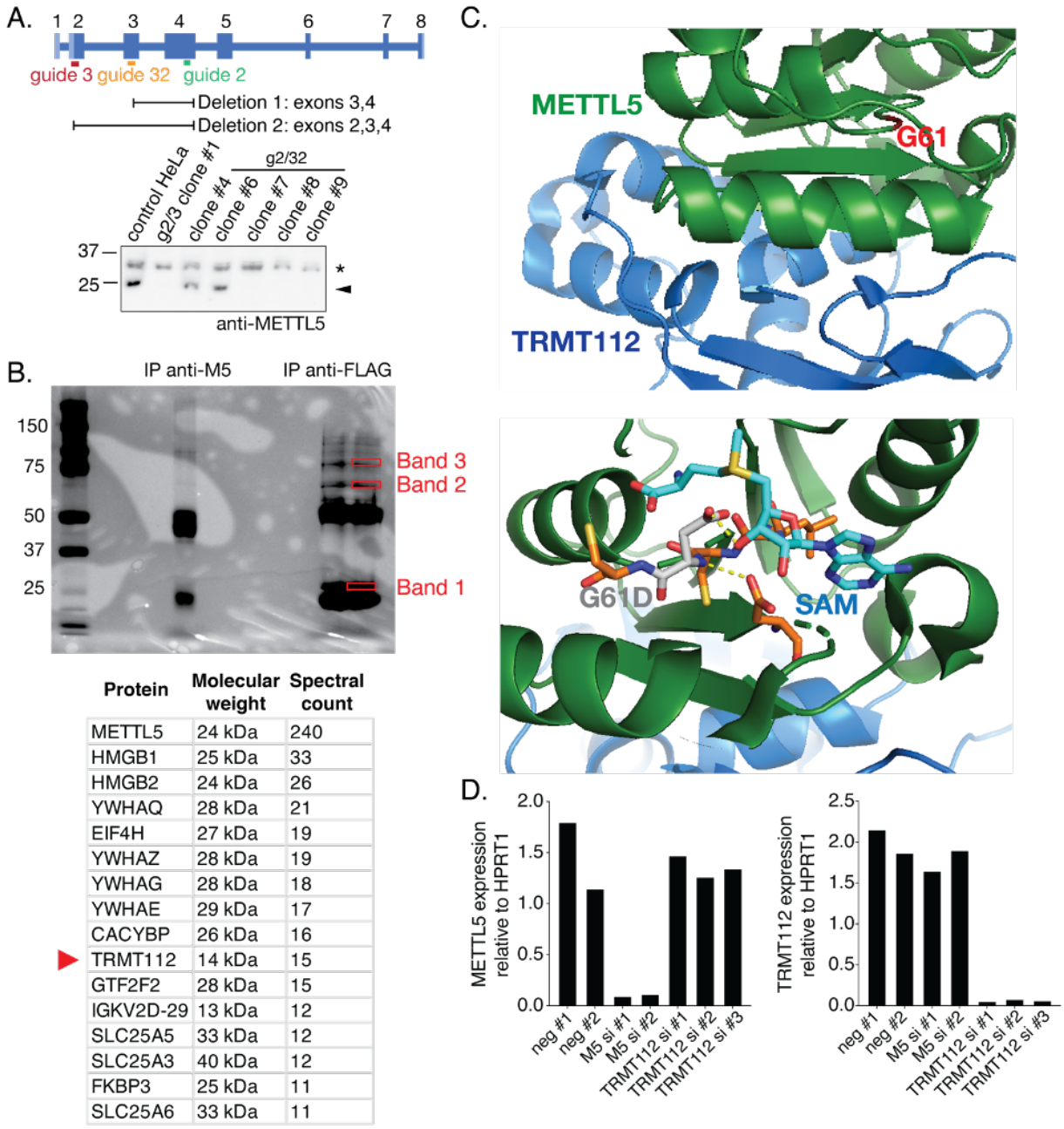


Figure 2.1. METTL5 is an m^{6}A methyltransferase stabilized by TRMT112. (A) METTL5 expression in cell lines as evaluated by western blot (top) with anti-GAPDH loading control (bottom). HEL (replicate 1) and (replicate 2) represent two different passages of HEL cells. Arrowhead signifies METTL5 band; * signifies background band. **(B)** Western blot analysis of METTL5 levels in wild type (WT) and knockout (KO) HeLa samples, each expanded from a single isolated clone (top) with anti-GAPDH loading control (bottom). Arrowhead signifies METTL5 band; * signifies background band. **(C)** Coimmunoprecipitations of TRMT112 with FLAG-METTL5. Left: Anti-FLAG (top) and anti-TRMT112 (middle) western blots in anti-FLAG immunoprecipitation samples. Input TRMT112 levels shown on bottom. Lanes from left to right are untransfected HeLa, HeLa transfected with FLAG-METTL5, and HeLa transfected with the FLAG-METTL5-3A mutant. Right: Anti-FLAG (top) and anti-TRMT112 (middle) western blots in anti-FLAG immunoprecipitation samples. Input TRMT112 levels shown on bottom. Labels: METTL5-WT, METTL5-3A, METTL5-G61D, METTL5-K191V: HeLa transfected with FLAG-METTL5-WT, FLAG-METTL5-3A, FLAG-METTL5-G61D, and FLAG-METTL5-K191Vfs*10, respectively. **(D)** Top: western blots for METTL5, TRMT112, and GAPDH (loading control) levels upon siRNA knockdown for METTL5 and TRMT112. Labels: Ø, untransfected HeLa cells; cont., HeLa transfected with two different non-targeting control siRNAs; METTL5, HeLa transfected with two different siRNAs targeting METTL5; TRMT112, HeLa transfected with three different siRNAs targeting TRMT112. Bottom: ratio of METTL5 to GAPDH signal intensity as quantified by Fiji software. **(E)** Top: western blot analysis of FLAG-METTL5 and FLAG-TRMT112 expression from HeLa transfected with labeled combinations of vectors, with GAPDH loading control. Bottom: Band intensity normalized to GAPDH band intensity, as quantified by Fiji software. **(F)** LC-MS/MS analysis of d_3m^6A (left) and d_3m^1A (right) levels normalized to guanosine from *in vitro* methyltransferase reactions performed on total RNA isolated from METTL5-WT and METTL5-KO HeLa clones. n=2 replicate reactions, mean and s.e.m. plotted, analyzed by one-way ANOVA, comparing all samples to HeLa WT with Dunnett's test for multiple comparisons. ns: not significant, * $p < 0.05$. All indicated band sizes in western blots are in kilodaltons (kDa).

Detailed studies of other m^{6}A methyltransferases have revealed specific motifs required for the catalytic activity of these enzymes, including the ^{395}DPPW^{398} motif in the m^{6}A mRNA methyltransferase METTL3 (Wang et al., 2016). METTL5 contains a chemically similar ^{127}NPPF^{130} motif, which led us to hypothesize that METTL5 may also be an m^{6}A methyltransferase. However, through our efforts to use *in vitro* methylation assays with tagged, overexpressed METTL5 protein to validate its methyltransferase activity, we noted lower-than-typical yields of purified METTL5 protein from both bacterial and mammalian expression systems, and low modification fractions in our *in vitro* assays (data not shown). Since methyltransferases such as METTL3-METTL14 function as a complex, this prompted us to perform proteomic analysis of METTL5 binding proteins using FLAG-METTL5. For these experiments, we turned to FreeStyle 293-F cells, which produce more robust overexpression of METTL5 than the HeLa cells we used previously. TRMT112 was a particularly intriguing candidate binding protein (**Figure 2.2B**), because it is known to bind and regulate other methyltransferases such as AlkBH8 (a tRNA methyltransferase) (Fu et al., 2010) and HemK2 (a protein methyltransferase) (Kusevic et al., 2016). Indeed, subsequent experiments confirmed a direct interaction between METTL5 and TRMT112 (**Figure 2.1C, left**) (van Tran et al., 2019).



Protein	Molecular weight	Spectral count
METTL5	24 kDa	240
HMGB1	25 kDa	33
HMGB2	24 kDa	26
YWHAQ	28 kDa	21
EIF4H	27 kDa	19
YWHAZ	28 kDa	19
YWHAG	28 kDa	18
YWHAE	29 kDa	17
CACYBP	26 kDa	16
▶ TRMT112	14 kDa	15
GTF2F2	28 kDa	15
IGKV2D-29	13 kDa	12
SLC25A5	33 kDa	12
SLC25A3	40 kDa	12
FKBP3	25 kDa	11
SLC25A6	33 kDa	11

Figure 2.2. Identification and investigation of METTL5-TRMT112 interactions. (A) Top: schematic of human METTL5 locus showing target locations of guide RNAs used to make METTL5 knockout HeLa and HepG2 cells in this study. Guides 2 and 3 were used to generate HeLa clone 1, whereas guides 2 and 32 were used to generate all other HeLa and HepG2 clones in this study. Bottom: METTL5 expression in cell lines as evaluated by western blot. Arrowhead signifies METTL5 band, * signifies background band. **(B)** Top: Silver-stained polyacrylamide protein gel with protein size marker, immunoprecipitate from pulldown with endogenous METTL5 antibody, and immunoprecipitate from pulldown with anti-FLAG antibody in Freestyle 293-F cells overexpressing FLAG-tagged METTL5. Boxed bands were cut out for proteomics analysis. Bottom: Table of top hits from proteomics analysis of band 1 with molecular weight and

Figure 2.2, continued. spectral count of each. Suspected common contaminant proteins were removed from the list. **(C)** METTL5-G61D human variant mutation site displayed on the METTL5-TRMT112 structure from van Tran *et al*. (PDB: 6H2U (van Tran et al., 2019)). Top: Position of G61 highlighted in an unstructured loop (red) in the context of the complex with METTL5 (green) and TRMT112 (blue). Bottom: Interactions between the mutated G>D residue (gray) and S-adenosylmethionine (cyan). Neighboring residues are highlighted in orange. Images created with PyMOL v2.4.0 (Schrodinger, 2010). **(D)** *METTL5* (top) and *TRMT112* (bottom) expression relative to *HPRT1* (housekeeping gene) in total RNA purified from HeLa cells treated with negative control siRNAs or siRNAs targeting *METTL5* or *TRMT112*, as evaluated by quantitative PCR. All indicated band sizes in western blots are in kilodaltons (kDa).

Mutation of a portion of the METTL5 putative catalytic motif ^{127}NPP^{129} to AAA (amino acids 127-129, abbreviated METTL5-3A), predicted to abolish its m^{6}A methyltransferase activity, weakened but did not completely disrupt this interaction. Since mutations in METTL5 have also been found in human patients with intellectual disability and microcephaly (Hu et al., 2019; Richard et al., 2019) we introduced two of these human variants into our FLAG-METTL5 construct, G61D and K191Vfs*10 (a 2bp deletion that results in a K>V point mutation and premature stop). In transient transfection overexpression experiments, FLAG-METTL5-G61D expressed at levels approximately 50-80% the level of expression of FLAG-METTL5-WT. Despite substantial expression of METTL5-G61D, the interaction with TRMT112 was compromised and the level of TRMT112 bound to FLAG-METTL5-G61D was approximately 5-15% relative to that bound to FLAG-METTL5-WT (**Figure 2.1C, right**). This reduced interaction could potentially be explained by structural changes introduced by disrupting flexibility of a loop region (**Figure 2.2C**). FLAG-METTL5-K191Vfs*10 expressed at much lower levels, suggesting that this truncated protein may not fold properly (**Figure 2.1C, right**; METTL5-K191V). Based on a previous report, the third variant identified in human patients, METTL5-R115Nfs*19, expresses only at very low levels, so we did not test its expression or interaction with TRMT112 (Richard et al., 2019).

The interaction with TRMT112 stabilizes METTL5 protein, as siRNA knockdown of TRMT112 reduced expression of METTL5 to approximately half relative to negative controls (**Figure 2.1D**), though less than direct siRNA knockdown of METTL5. The stabilization effect occurs at the protein level, as METTL5 transcript levels were unaffected by TRMT112 knockdown (**Figure 2.21D**). Conversely, coexpression of FLAG-METTL5 with FLAG-TRMT112 in HeLa cells significantly increased FLAG-METTL5 protein expression compared to solely expressing FLAG-METTL5 (**Figure 2.1E**). Coexpression of FLAG-TRMT112 with FLAG-METTL5 also substantially increased protein expression yields from FreeStyle 293-F cells, allowing us to purify FLAG-METTL5-FLAG-TRMT112 complex for *in vitro* methyltransferase assays. Using deuterated S-adenosyl methionine (d_3SAM), we performed *in vitro* methyltransferase assays with co-purified FLAG-METTL5-FLAG-TRMT112, using total RNA isolated from METTL5-WT and METTL5-KO HeLa cells. Analysis of deuterated, methylated nucleosides by LC-MS/MS revealed d_3m^6A in all four samples (**Figure 2.1F**). RNA from METTL5-KO cells accumulated approximately double the amount of d_3m^6A as RNA isolated from METTL5-WT cells, consistent with the idea that total RNA from METTL5-KO cells have less m^6A to begin with. Parallel measurements in these same samples showed 10-fold lower levels of d_3m^1A that remained consistent across all samples. Taken together, these results suggest that METTL5 is an m^6A RNA methyltransferase that requires TRMT112 for stability, as corroborated by van Tran *et al*. (van Tran et al., 2019).

18S rRNA is a major substrate of the METTL5-TRMT112 complex

To begin to understand the function of the METTL5-TRMT112 complex in cells, we first sought to identify its subcellular distribution. Fluorescence microscopy using antibodies targeting

endogenous METTL5 revealed nuclear puncta that colocalized with the nucleolar protein fibrillarin, but also had a more diffuse staining pattern in the cytoplasm (**Figure 2.3A**). This distribution was also verified using biochemical fractionation, which revealed that, although there was no detectable METTL5 associated with chromatin, there is a small pool of nuclear METTL5, with the majority of the protein localized to the cytoplasm (**Figure 2.3B**). This subcellular localization is particularly intriguing given the large difference observed in d_3m^6A methylation of total RNA isolated from WT and KO cells (**Figure 2.1F**), suggesting a relatively large pool of RNA to be methylated by METTL5-TRMT112. The nucleolus is a critical hub for the processing and assembly of ribosome components, and rRNAs are known to be m^6A methylated. Both 28S and 18S rRNAs have an m^6A site (Sergiev et al., 2018), and although ZCCHC4 has been identified as the methyltransferase for 28S rRNA A4220 (Ma et al., 2019c; Pinto et al., 2020), the methyltransferase for 18S rRNA was unknown when we initiated these studies.

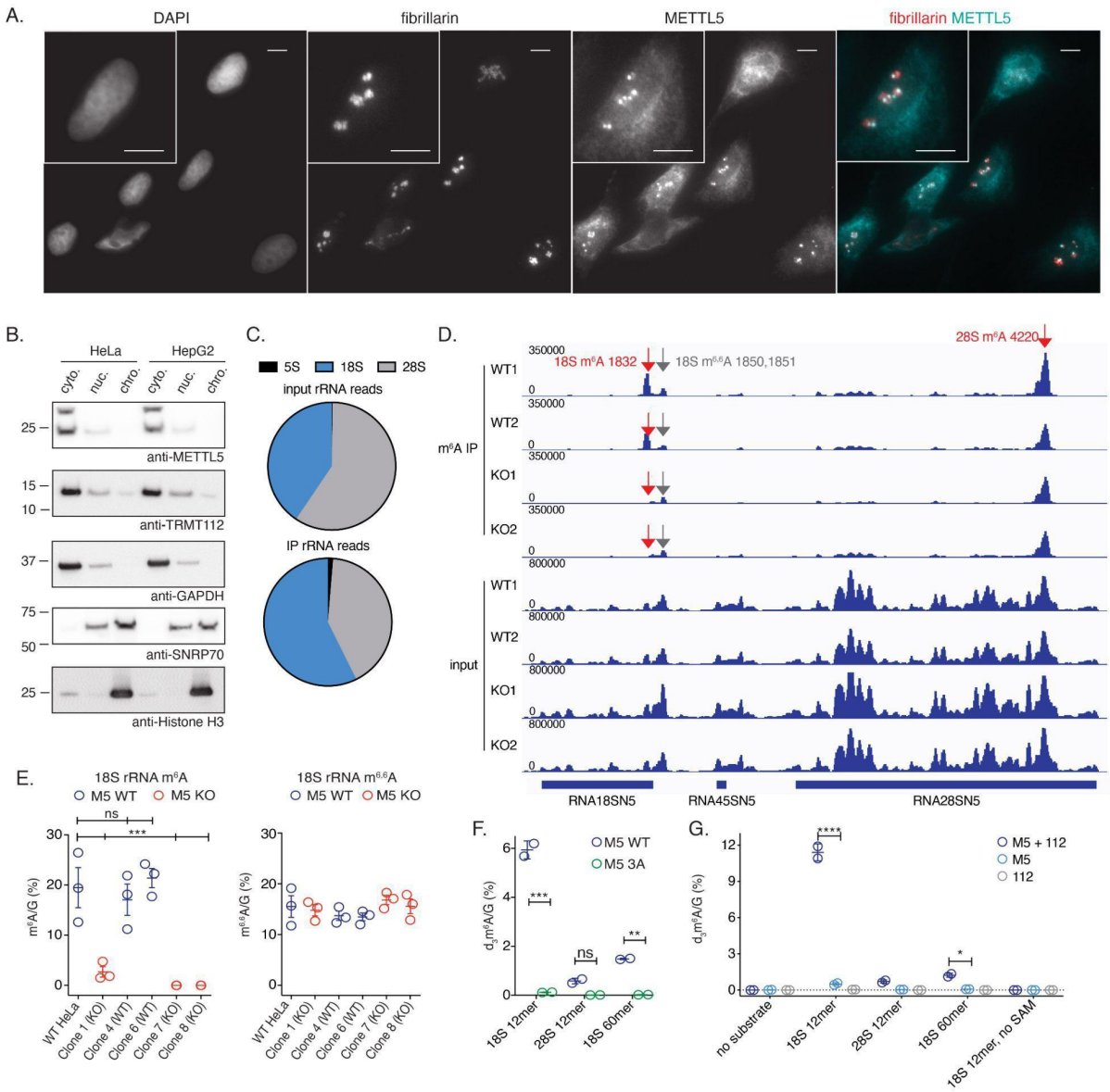


Figure 2.3. METTL5 is predominantly cytoplasmic and methylates 18S rRNA. (A) Fluorescence microscopy images of HeLa cells stained with DAPI (nuclear stain), anti-fibrillarin antibody (nucleolar marker), and anti-METTL5 antibody. Right panel shows merged image. Scale bar: 10µm. **(B)** Localization of METTL5 and TRMT112 in chromatin-associated, nuclear, and cytoplasmic cellular fractions in HeLa and HepG2 by western blot. Loading and fractionation controls: GAPDH (cytoplasmic), SNRP70 (nuclear), Histone H3 (chromatin-associated). **(C)** Pie charts of reads aligned to 5S, 18S, or 28S rDNA from HeLa input and METTL5 CLIP samples. **(D)** Visualization of reads aligned to the pre-45S locus from input and anti-m^{6}A immunoprecipitated Me-RIP-seq samples from HeLa METTL5-WT and METTL5-KO, adapted from Integrated Genomics Viewer. **(E)** Levels of m^{6}A (left) and m^{6,6}A (right), normalized to G, obtained by LC-MS/MS of 40-nt probe-purified segments of 18S rRNA surrounding the m^{6}A 1832 site from HeLa METTL5-WT and METTL5-KO. n=3 reactions per condition, mean and s.e.m. plotted. Comparisons of all samples to the WT HeLa control were tested by one-way ANOVA with Dunnett's test for multiple comparisons. **(F)** LC-MS/MS analysis of d$_3$m^{6}A levels normalized

Figure 2.3, continued. to G from *in vitro* methyltransferase reactions performed on 18S and 28S-mimicking probes with TRMT112 and either wild type METTL5 (M5 WT) or catalytic mutant METTL5-3A (M5 3A). **(G)** LC-MS/MS analysis of d_3m^6A levels normalized to G from *in vitro* methyltransferase reactions performed on 18S and 28S probes with METTL5 (M5) alone, TRMT112 alone (112), or METTL5-TRMT112 complex (M5+112). (F) and (G) were analyzed using two-way ANOVA to compare M5 WT with M5 3A (with Sidak test for multiple comparisons in (F)) and METTL5 with METTL5-TRMT112 (with Tukey test for multiple comparisons in (G)). n=2 reactions per condition, mean and s.e.m. * $p<0.05$, ** $p<0.01$, *** $p<0.005$. **** $p<0.0001$. (all other comparisons not statistically significant and labeled [ns], or not shown). All indicated band sizes in western blots are in kilodaltons (kDa).

To identify RNA substrates of METTL5, we performed crosslinking-assisted immunoprecipitation (CLIP) of FLAG-tagged METTL5 followed by high throughput sequencing of bound RNAs. To maximize the likelihood of success, we incorporated 4-thiouridine to facilitate crosslinking (Danan et al., 2016) (see **Materials and Methods** for a detailed protocol). High-throughput sequencing revealed a slight enrichment for 18S rRNA transcripts in the immunoprecipitate (IP) relative to input (**Figure 2.3C**), as well as other enriched transcripts (**Figure 2.4A**). m^6A-seq in METTL5-WT and METTL5-KO HeLa cells showed loss of an m^6A peak near adenosine 1832 in 18S rRNA whereas the adjacent $m^{6,6}A1850$ and $m^{6,6}A1851$ sites remained unchanged (**Figure 2.3D**), supporting the hypothesis that 18S rRNA is a major substrate of the METTL5-TRMT112 complex. We then applied biotinylated DNA probes complementary to the region of 18S rRNA containing both m^6A1832 and the neighboring $m^{6,6}A1850,1851$ sites to capture and purify these fragments from METTL5-WT and METTL5-KO cells for LC-MS/MS analysis (Ma et al., 2019c). Results confirmed that METTL5-TRMT112 methylates 18S rRNA A1832, as this fragment showed dramatically lower levels of m^6A in all KO lines tested relative to WT (**Figure 2.3E, left panel**). In contrast, the nearby $m^{6,6}A$ sites showed no significant change among these cell lines (**Figure 2.3E, right panel**), demonstrating that 18S rRNA processing is not dramatically altered and that these modifications are regulated independently.

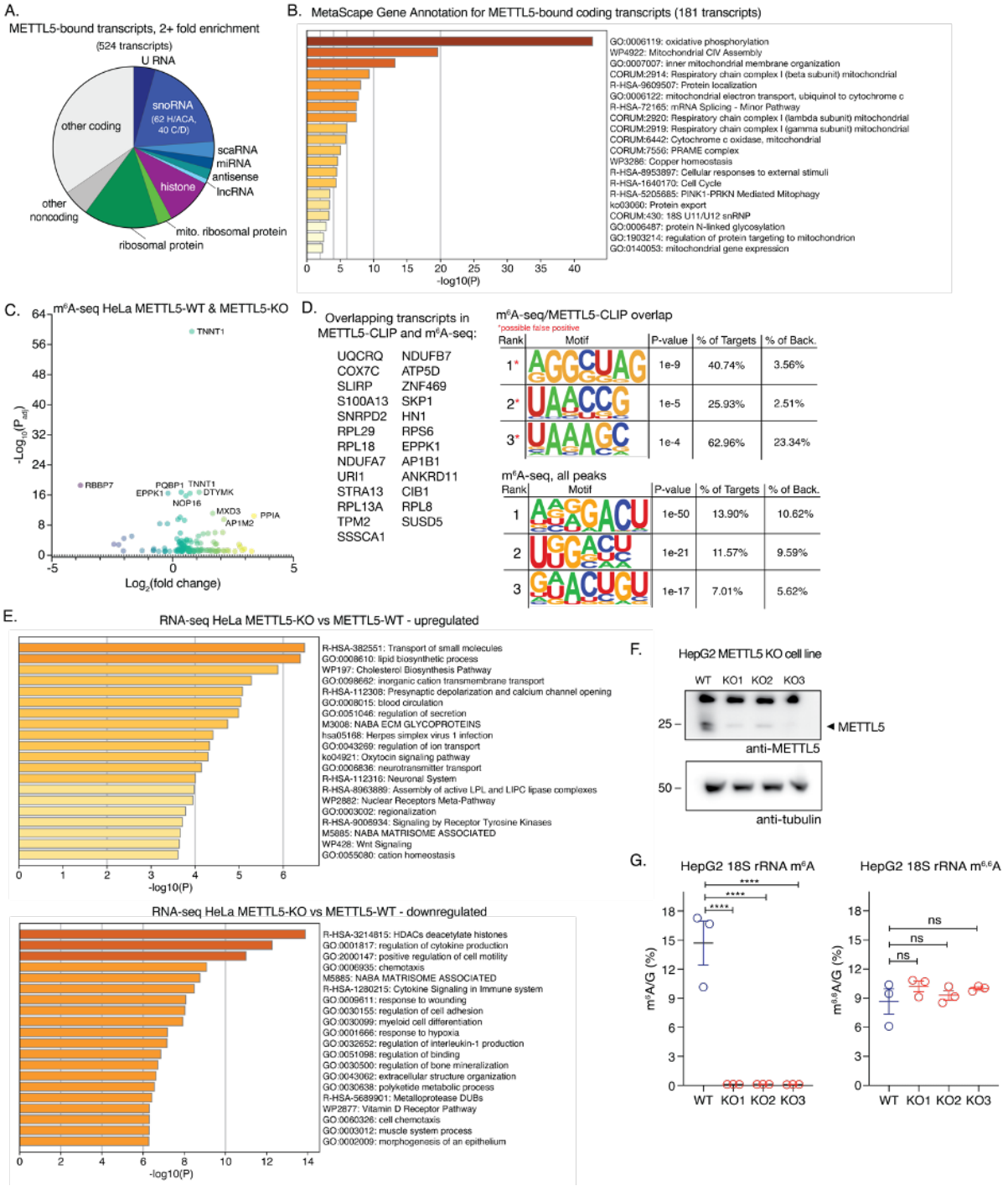


Figure 2.4. Investigation of METTL5 substrates in human cells. (A) Pie chart of METTL5-bound transcripts with ≥2-fold enrichment in METTL5 immunoprecipitate after crosslinking compared to input. **(B)** MetaScape gene annotations for METTL5-bound coding transcripts. **(C)** Volcano plot of differentially methylated m^{6}A peaks between HeLa METTL5-KO and METTL5-WT cells, shown comparing negative $\log_{10}$ of the adjusted P-value with $\log_2$ fold change. **(D)** Left: List of transcripts with both differentially methylated m^{6}A peaks from MeRIP-seq and ≥2-fold enrichment in METTL5 cross-linking and immunoprecipitation sequencing in HeLa cells. Right:

Figure 2.4, continued. Common motifs in this list of overlapping targets (top) and in all m^6A peaks from Me-RIP-seq (bottom), as predicted by HOMER (Heinz et al., 2010). Back.: background (see **Materials and Methods**) **(E)** MetaScape gene annotation terms from transcriptionally upregulated (top) and downregulated (bottom) transcripts in RNA-seq of METTL5-KO versus METTL5-WT HeLa cells. **(F)** Western blot analysis of METTL5 levels in METTL5-WT and METTL5-KO HepG2 cells, each expanded from a single isolated clone (top) with anti-tubulin loading control (bottom). **(G)** Levels of m^6A (left) and $m^{6,6}A$ (right), normalized to G, obtained by LC-MS/MS of 40-nt probe-purified segments of 18S rRNA surrounding the m^6A1832 site from METTL5-WT and METTL5-KO HepG2 cells. Analyzed by one-way ANOVA, comparing all samples to WT, with Dunnett's test for multiple comparisons. ns: not significant, * $p<0.05$, ** $p<0.01$, *** $p<0.005$, **** $p<0.0001$. Panels B and E modified from MetaScape output (Zhou et al., 2019b). All indicated band sizes in western blots are in kilodaltons (kDa).

To verify that METTL5-TRMT112 is not only necessary, but also sufficient for deposition of 18S rRNA m^6A1832, we expressed and purified FLAG-METTL5-WT and catalytically inactive FLAG-METTL5-3A with FLAG-TRMT112 from FreeStyle 293-F cells and tested their activity on RNA probes containing 18S or 28S rRNA sequences, using LC-MS/MS to measure d_3m^6A levels. Neither FLAG-METTL5-WT nor FLAG-METTL5-3A complexes could efficiently methylate a 28S rRNA 12-mer (**Figure 2.3F**), suggesting specificity for 18S rRNA sequence. Moreover, only METTL5-WT-TRMT112 complexes could effectively methylate 18S rRNA 12-mer or 60-mer probes, whereas the METTL5-3A mutant showed nearly undetectable activity. Using this 18S rRNA 12mer substrate, we then performed a more detailed assessment of the individual contributions of METTL5 and TRMT112 in methyltransferase activity. Although copurified FLAG-METTL5 and FLAG-TRMT112 could effectively methylate 18S rRNA probes, this activity was undetectable when either protein was purified individually, or when d_3SAM was left out of the reaction (**Figure 2.3G**). The loss of METTL5 activity in the absence of TRMT112 is likely the result of poorly folded METTL5 protein, as suggested by increased stability of METTL5 in the presence of TRMT112 (**Figures 2.1D, Figure 2.1E**). Taken together, our results are consistent with recent reports (Ignatova et al., 2020a; Rong et al., 2020; van Tran et al., 2019)

that the METTL5-TRMT112 complex m^6A methylates 18S rRNA and that the ^{127}NPP^{129} motif is critical for its catalytic activity.

To assess whether METTL5-TRMT112 may have other RNA substrates, we delved more deeply into the other transcripts enriched in our METTL5 CLIP experiment (**Figure 2.4A, Figure 2.4B**), which included both coding and non-coding RNAs. Cross-referencing these METTL5-bound RNAs with differentially methylated transcripts from our m^6A-seq experiment (**Figure 2.4C**) revealed only 25 overlapping transcripts (**Figure 2.4D**). Although the most prominent motif in the m^6A peaks overall was GGACU, suggesting that most m^6A peaks were METTL3/METTL14 dependent (Liu et al., 2014), the 25 transcripts in common between the two datasets showed enrichment for UAA, the motif containing the m^6A site in 18S rRNA. Thus, although it has been reported that 18S rRNA is the only METTL5-TRMT112 substrate (van Tran et al., 2019), it remains possible that other targets may exist. Notably, the coding transcripts identified in our CLIP experiment are enriched for genes involved in mitochondrial biogenesis and function (**Figure 2.4B**). RNA-seq analysis of differentially expressed transcripts in METTL5-WT and METTL5-KO HeLa cells also revealed enrichment for small molecule transport and lipid and cholesterol biosynthesis pathways (**Figure 2.4E**). These preliminary connections to metabolism and lipid biosynthesis, both liver-based functions, led us to generate HepG2 METTL5-KO cell lines, which may reflect gene expression pathways in the liver more closely (**Figure 2.4F, Figure 2.4G**).

METTL5 regulates translation of a subset of transcripts

Through the course of our experiments with both HeLa and HepG2 cells, we noted that METTL5-KO cells tended to grow more slowly than the corresponding METTL5-WT cells. To measure this difference directly, cell growth curves were monitored over the course of 96 hours

using DNA-dye-based CyQuant cell proliferation assays. In both HeLa and HepG2 cells, METTL5 KO cell lines grew more slowly than WT cells, with a small but consistent difference in growth rate (**Figure 2.5A**, **Figure 2.6A**). rRNAs are heavily modified with numerous chemical groups that regulate ribosome biogenesis and function, and play diverse roles in gene expression regulation. The slight growth defect we observed suggested that METTL5 is likely not essential for ribosome biogenesis (van Tran et al., 2019), but that it may regulate the translation of a subset of transcripts that collectively slows cell proliferation. Consistent with this, polysome profiling in both HeLa (**Figure 2.5B, Figure 2.5C**) and HepG2 cells (**Figure 2.6B**) show no notable changes in global translation. We note that the effect of METTL5 depletion on global translation has differed across recently published reports on METTL5 function, especially across different cell lines (Ignatova et al., 2020a; Leismann et al., 2020; Liberman et al., 2020; Rong et al., 2020; van Tran et al., 2019).

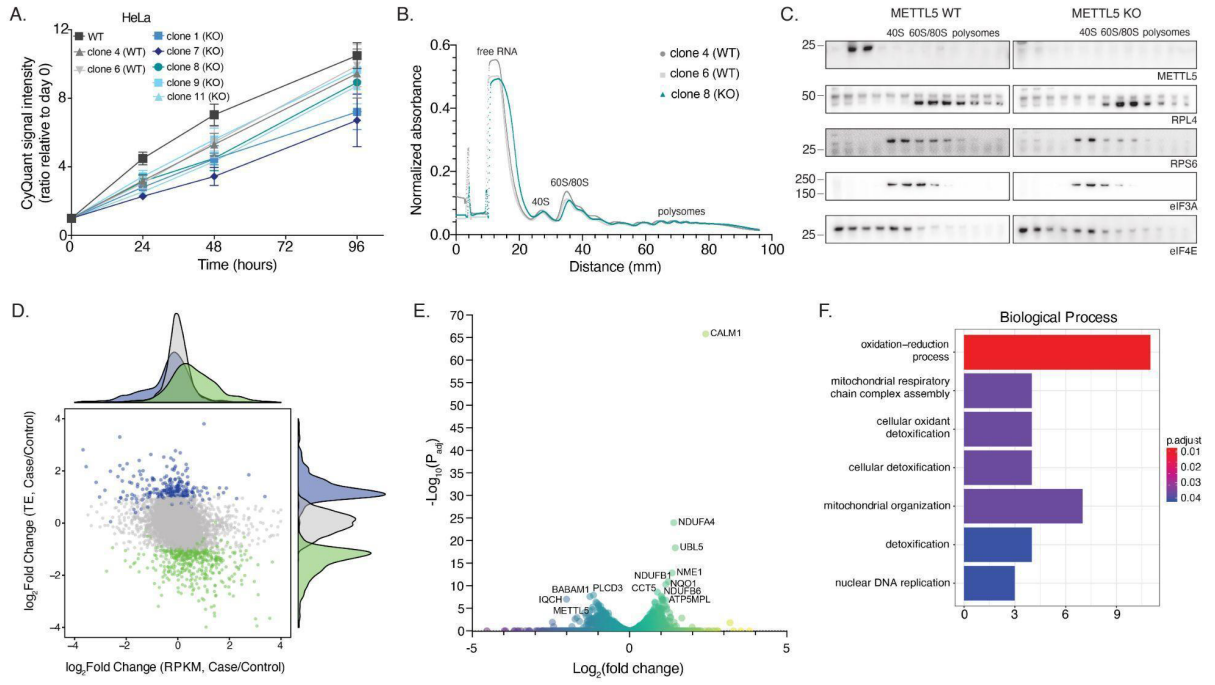


Figure 2.5. METTL5 affects cell proliferation and metabolism through differential translation of specific transcripts. (A) Cell proliferation of HeLa METTL5-WT and METTL5-KO cells over time as measured by CyQuant assay. Means (squares) and SDs (bars) are indicated

Figure 2.5, continued. for 4 replicate wells per condition. **(B)** Polysome profiles from HeLa METTL5-WT and METTL5-KO cells as measured by normalized absorbance over a 5-50% sucrose gradient. **(C)** Western blots of polysome fractions from panel B for METTL5, RPL4 (large ribosomal subunit marker), RPS6 (small ribosomal subunit marker), and eIF3A and eIF4E (translation initiation factors). **(D)** Effects of METTL5 knockout on translation and transcription as displayed by $\log_2$(fold change) of translation efficiency (TE) versus $\log_2$(fold change) of reads per kilobase of transcript per million mapped reads (RPKM) from ribosome profiling input and ribosome-protected fragment sequencing of HepG2 cells. Case = METTL5-KO, Control = METTL5-WT. **(E)** Volcano plot of the negative $\log_{10}$ of adjusted P-value versus $\log_2$(fold change) of normalized gene expression from ribosome profiling of HepG2 METTL5-KO versus METTL5-WT cells. **(F)** Biological process gene ontology (GO) functional enrichment analysis for translationally upregulated transcripts in HepG2 METTL5-KO compared to METTL5-WT cells. D and F are modified from output of RiboToolKit (Liu et al., 2020b). All indicated band sizes in western blots are in kilodaltons (kDa).

To identify specific transcripts whose regulation may be disrupted by loss of METTL5, we sequenced ribosome protected fragments. At the global level, ribosome profiling experiments in HepG2 cells revealed more significant changes at the level of transcript translation than transcript abundance. In comparing METTL5-WT versus METTL5-KO cells, translation efficiency (TE) showed greater variance than transcript abundance (measured as reads per kilobase of transcript per million mapped reads, RPKM), suggesting that gene expression changes in METTL5-KO cells may primarily occur during transcript translation rather than transcription (**Figure 2.5D**). This experiment also revealed several differentially translated transcripts in METTL5-KO cells (**Figure 2.5E**), including not only METTL5 as expected, but also CALM1, a calcium binding protein known to regulate cell proliferation and growth. Upon inspecting ribosome-protected fragment and input reads aligned to CALM1 in Integrated Genomics Viewer, we noted similar transcript levels in both METTL5-WT and METTL5-KO cells, but greatly increased ribosome occupancy in METTL5-KO cells, especially in exon 4 (**Figure 2.6C**). In addition, gene ontology analysis revealed translational upregulation of genes related to mitochondrial biogenesis and regulation, detoxification, and reduction/oxidation processes (**Figure 2.5F**), but it is not clear whether

translation upregulation of these transcripts is directly due to the loss of 18S m^6A or a compensation for broader cellular changes due to METTL5 loss. Overall, although our ribosome profiling data showed that a subset of transcripts were affected more than others, it did not point to a clear mechanism explaining the preference for certain transcripts. Additionally, although there was slight dysregulation across codons due to METTL5 depletion, no single codon or amino acid stood out as being particularly affected in terms of occupancy at the P-site, which is near 18S m^6A1832 (**Figure 2.6D**).

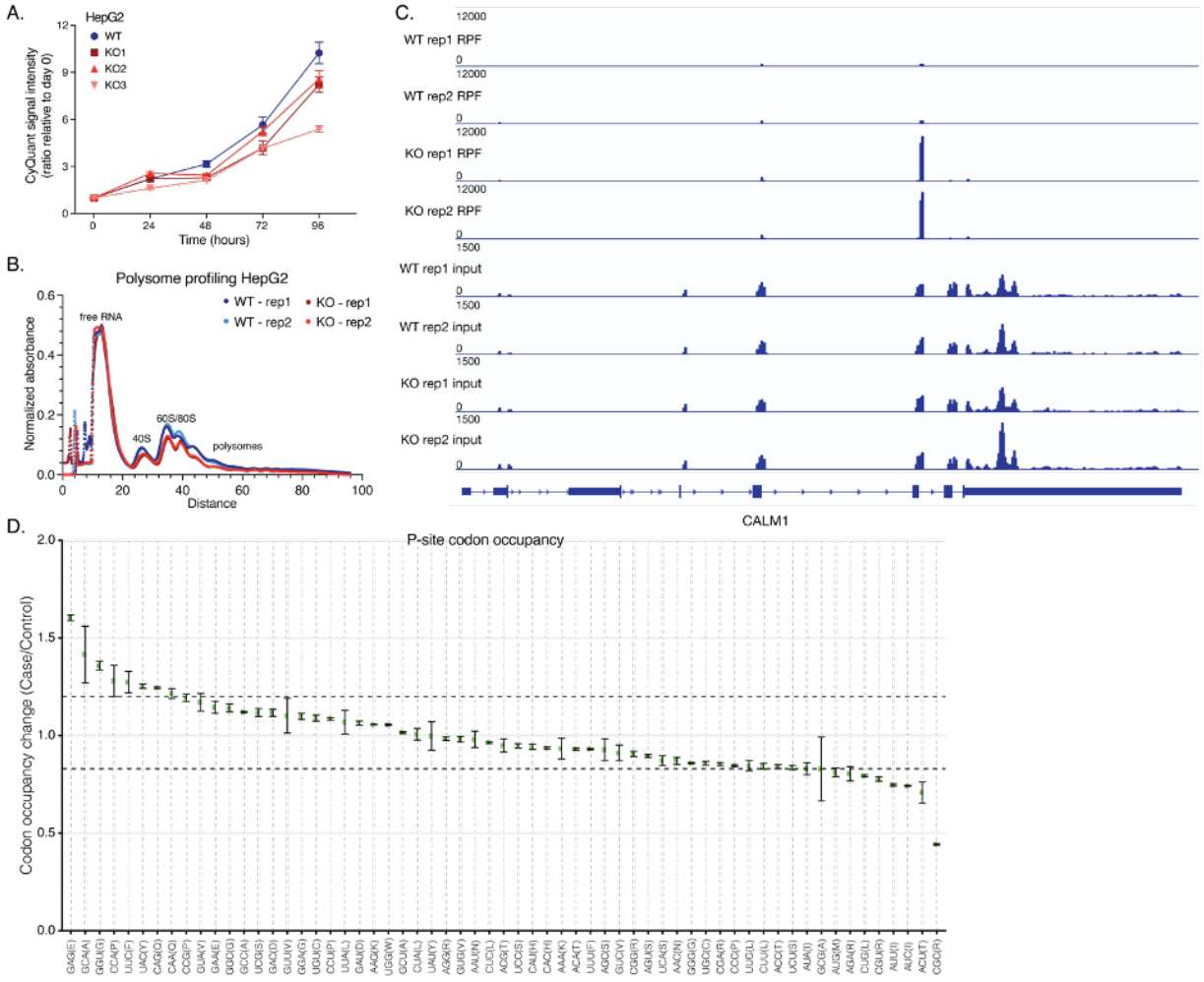


Figure 2.6. The effects of METTL5 knockout on translation and cell state. (A) Cell proliferation of HepG2 METTL5-WT and METTL5-KO cell lines over time as measured by

Figure 2.6, continued. CyQuant assay. Means (squares) and SDs (bars) are indicated for 4 replicate wells per condition. **(B)** Polysome profiles from HepG2 METTL5-WT and METTL5-KO cells as measured by normalized absorbance over a 5-50% sucrose gradient. **(C)** Visualization of reads at the *CALM1* locus from input and ribosome-protected fragment (RPF) samples from HepG2 METTL5-WT and METTL5-KO cells, adapted from Integrated Genomics Viewer. **(D)** Codon occupancy changes at the P site between METTL5-KO (case) and METTL5-WT (control) HepG2 cells. D is modified from output of RiboToolKit (Liu et al., 2020b).

Mettl5$^{-/-}$ mice demonstrate growth and metabolic changes

To assess METTL5 function at the level of a whole organism, we generated *Mettl5* knockout mice (*Mettl5*$^{-/-}$) by disrupting exon 2 of *Mettl5* with CRISPR-Cas9 (**Figure 2.8A**). *Mettl5* expression was undetectable in these mice by qPCR and greatly diminished by RNA-seq (**Figure 2.8B, Figure 2.8C**). Critically, m^{6}A levels on 18S rRNA isolated directly from the brains and livers of *Mettl5*$^{-/-}$ mice were abolished, whereas m^{6}A levels in *Mettl5*$^{+/+}$ and *Mettl5*$^{+/-}$ mice were similar and the neighboring m^{6,6}A site remained stable, consistent with our findings in METTL5-KO HeLa and HepG2 cells (**Figure 2.7A**). In agreement with Ignatova *et al.* (Ignatova et al., 2020a), we observed that *Mettl5*$^{-/-}$ mice are subviable and noted fewer than expected *Mettl5*$^{-/-}$ mice born from both HET/HET (**Figure 2.7B**) and HET/KO (**Figure 2.8D**) breeding pairs. It was also visibly apparent that *Mettl5*$^{-/-}$ mice were consistently smaller than *Mettl5*$^{+/+}$ and *Mettl5*$^{+/-}$ littermates (**Figure 2.7C**). Monitoring mouse weight across several weeks revealed that this size difference persisted over time in both male and female mice, with *Mettl5*$^{-/-}$ mice consistently weighing less than *Mettl5*$^{+/-}$ mice (**Figure 2.7D**). Though these measurements were taken from weeks 4 to 10, we noted that the difference was already present at the point of weaning (4 weeks), and persisted with similar magnitude, suggesting that this difference arose early in development, possibly even prior to birth.

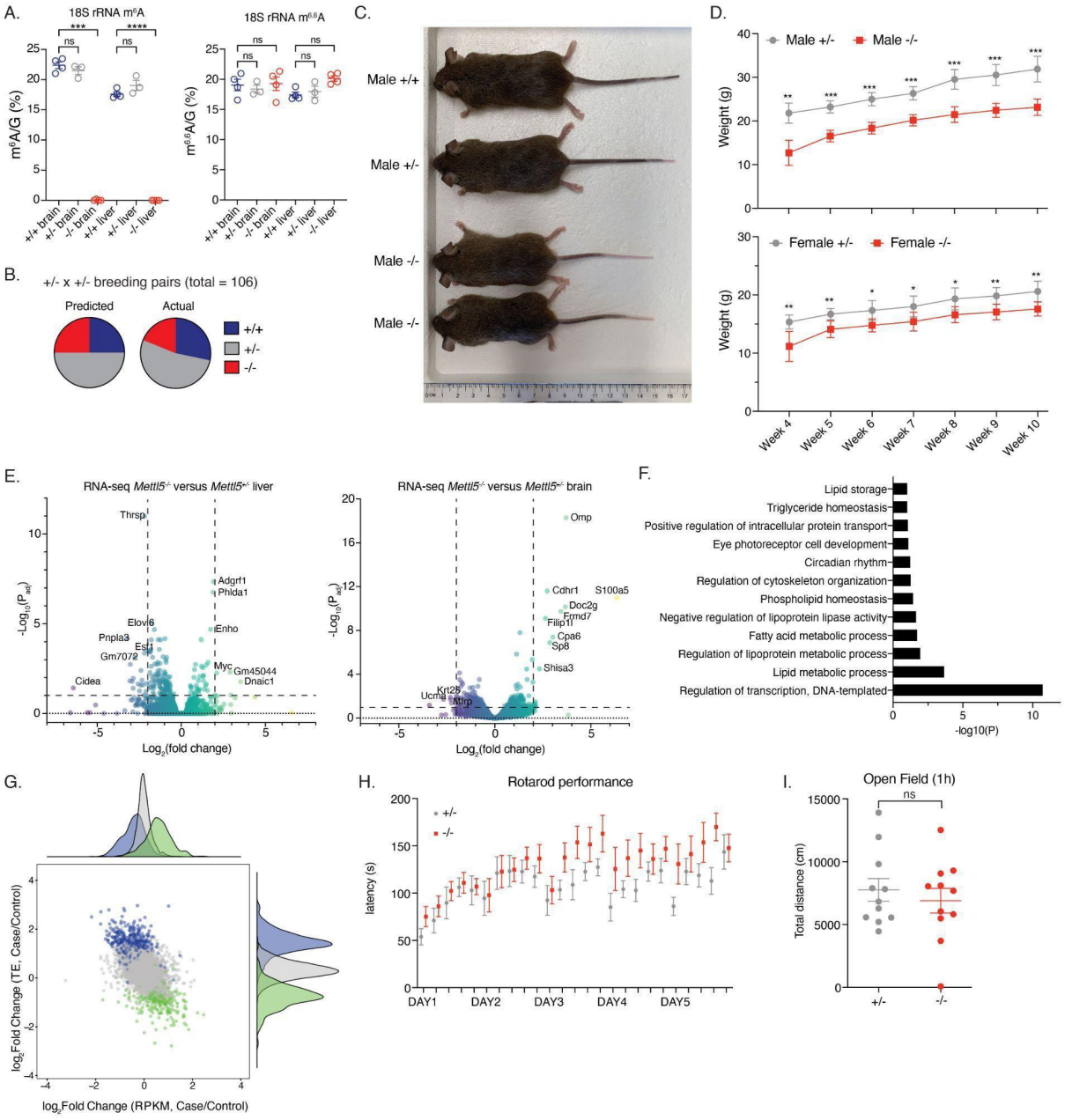


Figure 2.7. *Mettl5* knockout mice have developmental and metabolic defects. (A) Levels of m^6A (left) and $m^{6,6}A$ (right), normalized to G, obtained by LC-MS/MS of 40-nt probe-purified segments of 18S rRNA surrounding the m^6A 1832 site from brains and livers of *Mettl5*$^{+/+}$, *Mettl5*$^{+/-}$, and *Mettl5*$^{-/-}$ mice. n = 4, *Mettl5*$^{+/+}$, n = 3, *Mettl5*$^{+/-}$, and n = 4, *Mettl5*$^{-/-}$. Analyzed by one-way ANOVA comparison with Sidak test for multiple comparisons. **(B)** Pie chart of the predicted (left) and actual (right) mice born of wild type (+/+, dark blue), heterozygous (+/-, grey), and knockout (-/-, red) genotypes from heterozygous x heterozygous breeding pairs (total = 106 mice; 30 +/+, 56 +/-, 20 -/- mice). **(C)** Photograph of male littermate *Mettl5*$^{+/+}$, *Mettl5*$^{+/-}$, and *Mettl5*$^{-/-}$ mice at 8 weeks of age. **(D)** Body weights in grams of male (top) and female (bottom) *Mettl5*$^{+/-}$ and *Mettl5*$^{-/-}$ mice from 4 weeks (weaning) to 10 weeks of age. n=6 mice for male +/-, n=4 for male -/-, n=5 for female +/- and -/-; unpaired t-test was used to compare each time point. **(E)** Volcano plots from

Figure 2.7, continued. mouse liver (left) and brain (right) RNA-seq data displaying the negative $\log_{10}$ of the adjusted P-value versus the $\log_2$ fold change comparing the transcriptomes of *Mettl5*$^{-/-}$ and *Mettl5*$^{+/-}$ mice (n = 3 pairs). **(F)** Dysregulated pathways based on analysis of RNA-seq data from *Mettl5*$^{-/-}$ and *Mettl5*$^{+/-}$ mouse livers by the software DAVID (Huang et al., 2007). **(G)** Effects of METTL5 knockout on translation and transcription as displayed by $\log_2$(fold change) of translation efficiency (TE) versus $\log_2$(fold change) of reads per kilobase of transcript per million mapped reads (RPKM) from ribosome profiling input and ribosome-protected fragment sequencing of mouse livers. Case = *Mettl5*$^{-/-}$, Control = *Mettl5*$^{+/+}$. **(H)** Latency in seconds for *Mettl5*$^{+/-}$ and *Mettl5*$^{-/-}$ mice to fall off the rotarod in a rotarod performance test. **(I)** Total distance in centimeters travelled by *Mettl5*$^{+/-}$ and *Mettl5*$^{-/-}$ mice when placed in a new environment in an open field test. Unpaired t-test shown. H and I: n = 11 pairs. ns: not significant, * $p<0.05$, ** $p<0.01$, *** $p<0.005$.

In dissecting mice to harvest tissues for gene expression analysis, we noted that *Mettl5*$^{-/-}$ mice consistently had less body fat than *Mettl5*$^{+/-}$ or *Mettl5*$^{+/+}$ mice (**Figure 2.8E**). RNA-seq analysis of brain and liver tissues revealed gene expression patterns that suggest altered metabolism. In particular, gene ontology analysis of genes downregulated in *Mettl5*$^{-/-}$ mouse liver tissue revealed changes in genes involved in lipid biosynthesis and storage, consistent with the reduced body fat observed in *Mettl5*$^{-/-}$ mice (**Figure 2.7E, Figure 2.7F**). Noting the dramatic downregulation of *Thyroid Hormone Responsive Protein* (*Thrsp*) gene expression (**Figure 2.7E**), we measured T3 thyroid hormone levels in blood collected from the mice used for these RNA-seq studies and found that T3 levels may be elevated in the blood of *Mettl5*$^{-/-}$ mice relative to heterozygotes (**Figure 2.8F**). Furthermore, we performed ribosome profiling and sequenced ribosome protected fragments in *Mettl5*$^{+/+}$ and *Mettl5*$^{-/-}$ mouse liver tissues to assess changes in translation. Consistent with what we observed in METTL5-KO cell lines, although no significant changes in global translation were observed between polysome profiles from livers of *Mettl5*$^{+/+}$ and *Mettl5*$^{-/-}$ littermate pairs (**Figure 2.8G**), translation efficiency showed greater variance than transcript abundance in our ribo-seq analysis (**Figure 2.7G**). These data suggest that gene expression changes in *Mettl5*$^{-/-}$ mice occur primarily at the level of translation. However, using the

same parameters used to assess differentially translated genes in knockout cell lines (**Figure 2.5E, Figure 2.8H, Figure 2.8I**), we found there were few statistically significant differentially translated transcripts in *Mettl5*$^{-/-}$ mouse livers compared to *Mettl5*$^{+/+}$ littermate controls, potentially due to greater variability among mice (**Figure 2.8H, Figure 2.8I**).

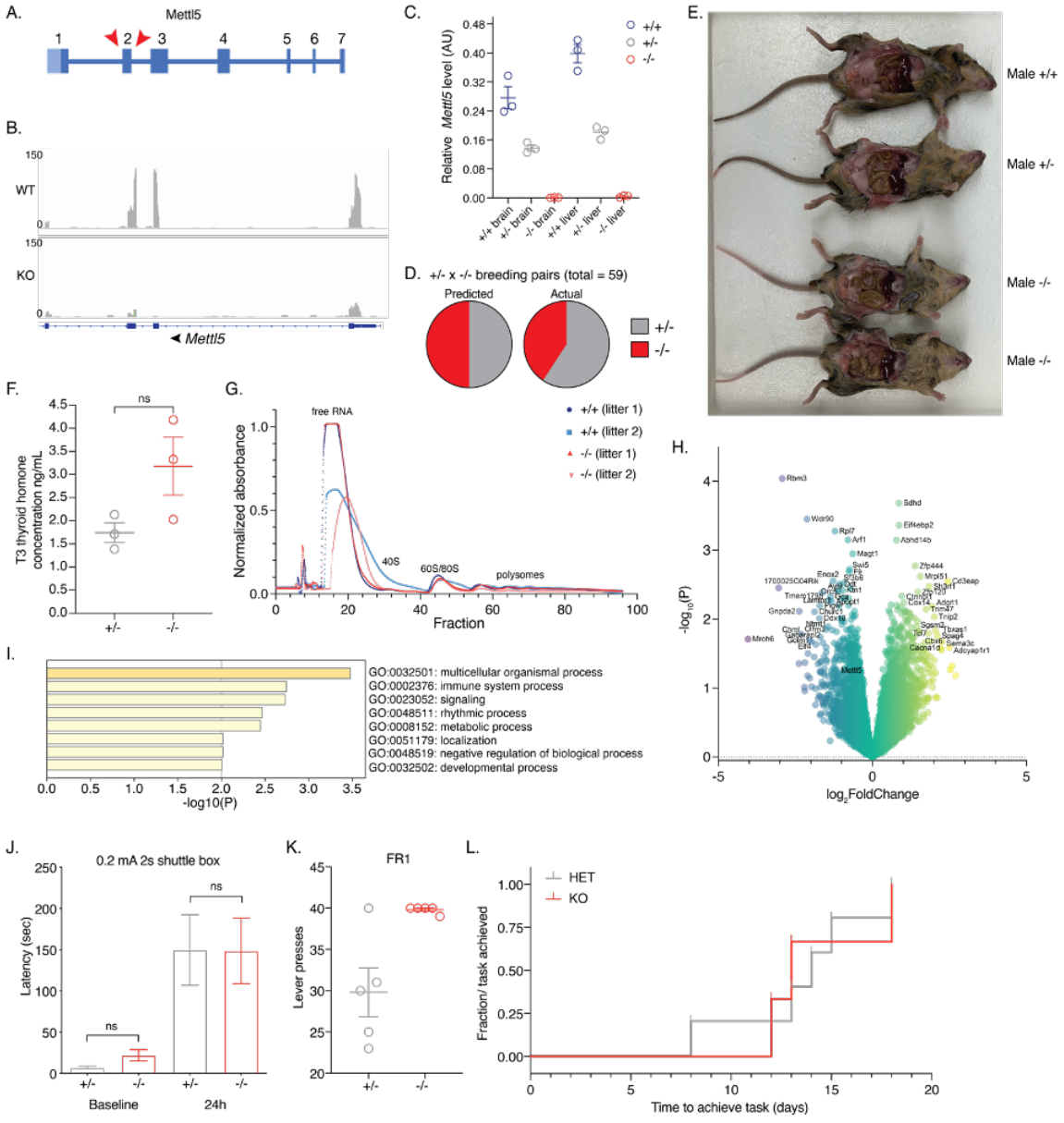


Figure 2.8. Characterizations of *Mettl5*$^{-/-}$ mice. (A) Schematic of mouse *Mettl5* locus showing region targeted for removal by CRISPR-Cas9 to create KO mice for this study. **(B)** Visualization

Figure 2.8, continued. of RNA-seq reads at mouse *Mettl5* locus from a littermate pair of one *Mettl5*$^{+/-}$ and one *Mettl5*$^{-/-}$ mouse, adapted from Integrated Genomics Viewer. Exon 1 is at far right and exon 4 is on the far left; partial view shown since *Ssb* overlaps with *Mettl5*. **(C)** Relative *Mettl5* transcript level (normalized to *Hprt*) as measured by qPCR of total RNA from brains and livers of *Mettl5*$^{+/+}$, *Mettl5*$^{+/-}$ and *Mettl5*$^{-/-}$ mice. **(D)** Pie charts of the predicted (left) and actual (right) mice born of *Mettl5*$^{+/-}$, (+/-, grey) and *Mettl5*$^{-/-}$ (-/-, red) genotypes from +/- x -/- breeding pairs (total = 59 mice; 35 heterozygous, 24 knockout). **(E)** Photo of dissected male littermate mice at 8 weeks showing abdominal fat content. **(F)** T3 thyroid hormone concentration (ng/mL) from enzyme-linked immunosorbent assay (ELISA) of thyroid hormone in serum of the same heterozygous and knockout mice used for RNA-seq (Figure 4E). n = 3 pairs, unpaired t-test, ns: not significant. **(G)** Polysome profiles from littermate pairs of *Mettl5*$^{+/+}$ and *Mettl5*$^{-/-}$ mouse liver tissue as measured by normalized absorbance over a 5-50% sucrose gradient. Two littermate pairs were used, indicated in the legend. **(H)** Volcano plot of the negative $\log_{10}$ of P-value versus $\log_2$(fold change) of normalized read counts from ribosome profiling of *Mettl5*$^{+/+}$ and *Mettl5*$^{-/-}$ mouse liver tissue as measured by normalized absorbance over a 5-50% sucrose gradient. **(I)** MetaScape gene annotation terms from translationally upregulated transcripts in ribosome profiling of *Mettl5*$^{+/+}$ and *Mettl5*$^{-/-}$ mouse liver tissue. **(J)** Latency in seconds for mice to move into the dark side of a shuttle box at baseline and 24 hours after training with shock to avoid that side. Data analyzed by one-way ANOVA with Sidak test for multiple comparisons, comparing *Mettl5*$^{+/-}$ to *Mettl5*$^{-/-}$ at baseline and at 24 hours; n=11 pairs, n.s.: not significant. **(K,L)** Number of lever presses on the last day (K) and time to learn task (L) by *Mettl5*$^{+/-}$ and *Mettl5*$^{-/-}$ mice in FR1 training to press a lever for food reward after food deprivation. n = 5 pairs (K, L). Log-rank (Mantel-Cox) test performed in (**L**) is not significant.

Since METTL5 mutations are associated with intellectual disabilities in human patients, we wanted to determine whether cognitive or behavioral changes were evident in our mouse model. Recent reports suggest that loss of *Mettl5* results in reduced locomotor activity and exploratory activity, and defects in learning and memory (Ignatova et al., 2020a; Wang et al., 2022). However, in our mouse model, we did not observe similar defects in locomotor and exploratory activity from rotarod performance and open field tests (**Figure 2.7H, Figure 2.7I**), in fear-based learning from a shuttle box test (**Figure 2.8J**), or in instrumental learning from FR1 acquisition (**Figure 2.8K, Figure 2.8L**). We note that our experiments were done comparing knockout mice with heterozygotes, not wild type mice as in other studies. However, given that

Mettl5$^{+/+}$ and *Mettl5*$^{+/-}$ mice have similar levels of 18S rRNA methylation (**Figure 2.7A**), we believe that any differences between these mice are due to something other than differences in 18S rRNA methylation status. However, combined with small variations in experimental setup, the comparison of *Mettl5*$^{-/-}$ to *Mettl5*$^{+/+}$ versus *Mettl5*$^{+/-}$ could contribute to the differences in outcomes in these behavior assays.

DISCUSSION

In our study, we identified TRMT112 as the primary binding partner of METTL5, a finding consistent with other recent reports (Ignatova et al., 2020a; Leismann et al., 2020; van Tran et al., 2019). We found that protein-protein interactions between METTL5 and TRMT112 are important for METTL5 stability (**Figure 2.1D, Figure 2.1E, Figure 2.2C, Figure 2.2D**), and methyltransferase activity (**Figure 2.3G**), though the latter is likely a consequence of the stabilizing effect of TRMT112 on METTL5. Of clinical significance, we found that the METTL5-TRMT112 interaction was severely reduced by mutations known to cause intellectual disability in humans (**Figure 2.1C**) (Hu et al., 2019; Richard et al., 2019). Structural analysis of the three major human variants reported by Richard *et al.* and Hu *et al.* (Hu et al., 2019; Richard et al., 2019) suggests that all three mutations (R115Nfs*19, K191Vfs*10, G61D) would likely disrupt proper folding (**Figure 2.2C**). Of these, the frameshift variants, R115Nfs*19 and K191Vfs*10, were demonstrated to cause lower METTL5 expression (Hu et al., 2019; Richard et al., 2019). Structural analysis suggests that the METTL5^{G61D}-TRMT112 complexes that do form may lack catalytic activity, since the G61D mutation creates a polar interaction with S-adenosyl methionine (**Figure 2.2C**). Disruption of the METTL5-TRMT112 interaction likely also decreases the stability of these METTL5 variants, negatively impacting their methyltransferase activity and contributing to

disease. To our knowledge, this is the first time that the METTL5-TRMT112 interaction has been investigated in the context of human disease. Intriguingly, our results further suggest the possibility that there are other pathways dysregulated by disease-causing METTL5 variants. For example, the availability of TRMT112 to its other binding partners, which include AlkBH8 and HemK2, could also be dysregulated in this context (Fu et al., 2010; Kusevic et al., 2016; Leetsi et al., 2019).

We also characterized METTL5 m^6A methyltransferase activity and investigated its substrates *in vitro* and *in vivo*, identifying 18S rRNA m^6A1832 as a major METTL5 substrate in mammalian cells, a result corroborated by recent reports (Ignatova et al., 2020a; van Tran et al., 2019; Xing et al., 2020). Interestingly, though the nucleolar localization of METTL5 pointed to its role in rRNA methylation, we also found significant cytoplasmic localization of METTL5 by fluorescence microscopy and biochemical fractionation (**Figure 2.3A, Figure 2.3B**). Significant cytoplasmic localization has also been reported in neurons (Richard et al., 2019) and *Drosophila melanogaster* (Leismann et al., 2020), raising questions about its cytoplasmic function. One plausible cytoplasmic role for METTL5 is late-stage methylation of rRNA, in line with analysis by van Tran *et al.* that found METTL5-TRMT112 complex near helix 44 only at late stages of processing (Ameismeier et al., 2018; van Tran et al., 2019). METTL5 may also remain associated with the ribosome in the cytoplasm, where it could regulate translation by recruiting factors to the ribosome. We note, however, that western blotting for METTL5 across polysome profiling fractions did not reveal an interaction with translating ribosomes in HeLa cells (**Figure 2.5C**). Additionally, cytoplasmic METTL5 localization could suggest the existence of other METTL5 methylation targets, and differences in abundance and localization of substrates in various cell lines could contribute to discrepancies in its localization. Our CLIP-seq and m^6A-seq identified a small subset of transcripts that were both associated with METTL5 and whose m^6A levels changed

in METTL5-KO cells, respectively, supporting this idea (**Figure 2.4A-D**). Although some of these may represent indirect interactions, it is interesting that motifs mimicking the 18S m^6A1832 sequence context are enriched in differentially methylated m^6A peaks on transcripts overrepresented in the METTL5 CLIP data (**Figure 2.4D**). Moreover, there is discordance among existing studies about non-18S rRNA substrates. Although Ignatova *et al.* demonstrated METTL5 activity on polyA-enriched RNA, van Tran *et al.* did not identify any non-rRNA sites in their miCLIP screen (Ignatova et al., 2020a; van Tran et al., 2019). This may be explained by the fact that in an *in vitro* experiment, in the absence of rRNA, activity is more likely detectable on other substrates such as polyA-enriched RNA. On the other hand, a miCLIP experiment could also miss more transient interactions on less abundant RNAs such as mRNAs. Altogether, existing data makes it difficult to draw conclusions as to other possible METTL5 substrates, and this question warrants future investigation.

METTL5 activity on 18S rRNA prompted us to characterize the effects of METTL5 depletion on translation. Though polysome profiling did not reveal global translational changes in HeLa or HepG2 METTL5-KO cells (**Figure 2.5B, Figure 2.5C, Figure 2.6B**) or in *Mettl5*$^{-/-}$ mice (**Figure 2.8G**), ribosome profiling revealed dysregulated transcripts in cell lines (**Figure 2.5E, Figure 2.8H**) and greater variation at the translational than the transcriptional level (**Figures 2.5D, Figure 2.7G**). These data are consistent with 18S rRNA being a major cellular substrate of METTL5. Interestingly, a more drastic effect on polysomes has been documented in mouse embryonic stem cells (Ignatova et al., 2020a; Rong et al., 2020; Xing et al., 2020). Clues to how METTL5 causes transcript- and context-specific effects may come from the location of the 18S m^6A1832 site at the tip of helix 44 near the decoding center (Natchiar et al., 2017; van Tran et al., 2019). Rong *et al.* proposed that the methyl group may fine-tune the conformation of the decoding

center and suggested that the methylated adenine and its base-pairing partner are in closer proximity to mRNA in the human ribosome than in structures lacking the methylation from other organisms (Rong et al., 2020). The location of m^6A1832 on the 3' end of 18S rRNA suggests that it could be surface exposed in some contexts and in close proximity to the mRNA being translated. The presence or absence of m^6A1832 could thereby alter interactions between the mRNA and the ribosome, possibly even between 18S rRNA and the mRNA directly, thereby modulating translation efficiency. Although we and others have observed through ribosome profiling that certain transcripts and codons are more affected than others (**Figure 2.5E, Figure 2.6D, Figure 2.8H**), the specific codons and transcripts affected are not consistent across datasets (Ignatova et al., 2020a; Xing et al., 2020). Loss of 18S m^6A1832 may also affect the position of helix 44, which is located near binding sites for key initiation and re-initiation factors, including eIF1, eIF1A, DENR, and eIF2D (Lomakin et al., 2017; Weisser et al., 2013, 2017). Indeed, Rong *et al.* reported altered binding of the initiation factors eIF3A and eIF4E to translating ribosomes and decreased phosphorylation of the translation initiation-related signaling protein RPS6 in METTL5-KO HEK293T and HeLa cells (Rong et al., 2020). It is thus possible that some of the effects of METTL5 may be mediated by changes in binding of initiation-related factors to the ribosome. Considering the larger context beyond the ribosome itself may also elucidate differences in the effects of METTL5 knockout on translation in different cell lines, as there are known differences in translation-related signaling pathways among cell lines, especially in stem cells (Tahmasebi et al., 2018).

Our findings, particularly from our *Mettl5* knockout mouse model, provide insight into hereditary METTL5-related diseases caused by human variants. To our knowledge, this is the third *Mettl5* knockout mouse model reported, but the first for which 18S m^6A loss was validated

(Ignatova et al., 2020a; Wang et al., 2022). Guided by the observations that *Mettl5*$^{-/-}$ mice weighed less (**Figure 2.7D**) and had decreased body fat (**Figure 2.8E**), we further investigated the metabolism of the mice through RNA-seq of their livers. We found dysregulated lipid, lipoprotein, and fatty acid metabolic pathways in knockout mice (**Figure 2.7F**) and changes in thyroid hormone signaling (**Figure 2.7E, Figure 2.8F**). Interestingly, metabolic dysregulation was also suggested in a report of *mettl-5* knockout in *C. elegans* (Liberman et al., 2020). The mechanism(s) leading to these changes remain unclear but could have clinical significance, as many patients with METTL5-associated microcephaly and intellectual disability reported by Richard *et al.* are reported to have reduced body weight (Richard et al., 2019). Unlike recent reports (Ignatova et al., 2020a; Wang et al., 2022), however, tests with our mouse model for neurological and behavioral deficits failed to show statistically significant differences between heterozygous and knockout mice in learning ability, motor activity, or exploratory activity (**Figure 2.7H, Figure 2.7I, Figure 2.8J-L**), potentially due to differences in experimental setup or mouse model design (**Figure 2.8A**). Another possibility is that the loss of activity in heterozygotes may be enough to affect neurological function, although the intellectual disability and microcephaly clinical phenotypes were reported to be autosomal recessive (Richard et al., 2019). The 18S rRNA m^6A levels were also similar in *Mettl5*$^{+/+}$ versus *Mettl5*$^{+/-}$ mice (**Figure 2.7A**). Thus, such an effect would likely be unrelated to the rRNA methyltransferase activity of METTL5-TRMT112. We also note that the patient-derived METTL5 mutants are expressed to some degree, meaning our complete knockout cell lines and mouse model do not entirely mimic the physiological conditions. It also remains possible that the neurological and behavioral defects caused by METTL5 mutations may be more subtle than is easily detectable by the simple tasks or sample sizes we tested but significant enough in disrupting complex tasks in humans to cause clinical intellectual disability.

Given that 18S m^6A1832 is currently the only validated substrate of METTL5 and the effects of METTL5 knockout seem to be most significant at the translational level, questions arise about how losing a single methyl group in the context of the whole ribosome may lead to organism-level effects. Notably, abnormalities in development and metabolism have been found in characterized clinical ribosomopathies (Yelick and Trainor, 2015). Furthermore, the differences seen in the effects of METTL5 in different cell types and tissues mirror other ribosomopathies that are very tissue-specific (Farley-Barnes et al., 2019). Since our findings and those reported by others have solidified the role of METTL5 in ribosome biogenesis and function, we suggest that diseases caused by METTL5 variants could represent a previously uncharacterized ribosomopathy. The clinical relevance of METTL5 underscores the importance of future investigation into the mechanisms by which mutations lead to these phenotypes. Such studies should include comprehensive identification and validation of METTL5 substrates alongside detailed investigations into protein translation and signaling pathway changes, which may in turn guide the development of novel clinical interventions.

SUPPLEMENTAL NOTE

We encountered METTL5 in the course of a biochemical fractionation experiment, whose purpose was actually to identify m^1A methyltransferases. In brief, nuclear extract was fractionated over a cation exchange column, and the fractions were collected and tested for m^1A methyltransferase activity. Active fractions were then pooled and fractionated over an anion exchange column, and fractions were again collected and tested for m^1A methyltransferase activity.

Thirty 15cm dishes of HeLa cells were grown to ~80-90% confluency. Cells were washed with ice cold 1x PBS prior to harvesting cells with cell lifters. The cells were pelleted by centrifugation at 500g for 10 minutes at 4ºC (cell pellet was ~4mLs). The cell pellet was gently washed in 12mLs Buffer A and pelleted by centrifugation at 500g for 5 minutes at 4ºC. The cell pellet was gently resuspended in 8mLs Buffer A (10mM HEPES pH 7.5, 1.5mM MgCl2, 10mM KCl, 1x SigmaFast Protease Inhibitor cocktail [Sigma]) and incubated on ice for 10 minutes. The cell suspension was then lysed with a dounce with a loose pestle (20 times, on ice). Nuclei were pelleted by centrifugation at 500g for 5 minutes at 4ºC, and the supernatant (cytoplasm) was removed. The nuclear pellet was then resuspended and lysed in 1.5mLs of Buffer C (20mM HEPES pH 7.5, 25% glycerol, 0.42M NaCl, 1.5mM MgCl2, 0.2mM EDTA, 0.5mM DTT, 1x SigmaFast Protease Inhibitor cocktail [Sigma]), rotating for 30 minutes at 4ºC. Nuclear lysate was spun at 20,000g for 30 minutes at 4ºC to pellet debris, and the supernatant (~2mLs) was dialyzed into 20mM HEPES pH 7.5, 50mM NaCl, 1.5mM $MgCl_2$, 0.5mM EDTA using a 3500 Da molecular weight cut off cassette for 6 hours at 4ºC.

The dialyzed lysate was then fractionated over a Fast Flow SP (GE Healthcare) column in 20mM HEPES pH 7.5,1.5mM $MgCl_2$, 0.5mM EDTA using a gradient from 50mM to 1M NaCl. Fractions were collected and concentrated and exchanged into methyltransferase buffer (MT buffer: 50mM Tris pH 8.0, 10% glycerol, 1.5mM $MgCl_2$, 0.5mM EDTA, and 100mM NH_4SO_4) through 10,000 Da molecular weight cut off spin filters. Each fraction was tested for methyltransferase activity by combining 20uL fraction with 1mM d_3SAM, 3% v/v SUPERasin RNase inhibitor, and 750ng of an RNA probe (5'-CUGACCCUCGCCUGCACCCGCCCCGAGAAGCCGAG-biotin-3'). Reactions were

incubated at 30ºC for 30 minutes and then snap frozen and stored at -80ºC until clean-up and processing for LC-MS/MS analysis.

For clean-up, 1mM EDTA and 1uL Proteinase K (Sigma, #P4850) was added to each 30uL methyltransferase reaction, and incubated for 1 hour at 37ºC. The samples were then diluted into 500uL IP buffer (5mM Tris pH 7.4, 1M NaCl, 1mM EDTA, 1% SUPERasin). 10uL MyOne C1 streptavidin beads per sample were washed into IP buffer and added to each sample. Samples with beads were rotated for 2 hours at 4ºC, after which beads were washed twice with IP buffer and eluted with Trizol (twice for 5 minutes at room temperature, shaking at 1000rpm). The eluates were then cleaned up with the DirectZol Micro clean up kit (Zymo). The samples were then digested into nucleosides and analyzed by LC-MS/MS as described in the 'LC-MS/MS' section (see **Materials and Methods**).

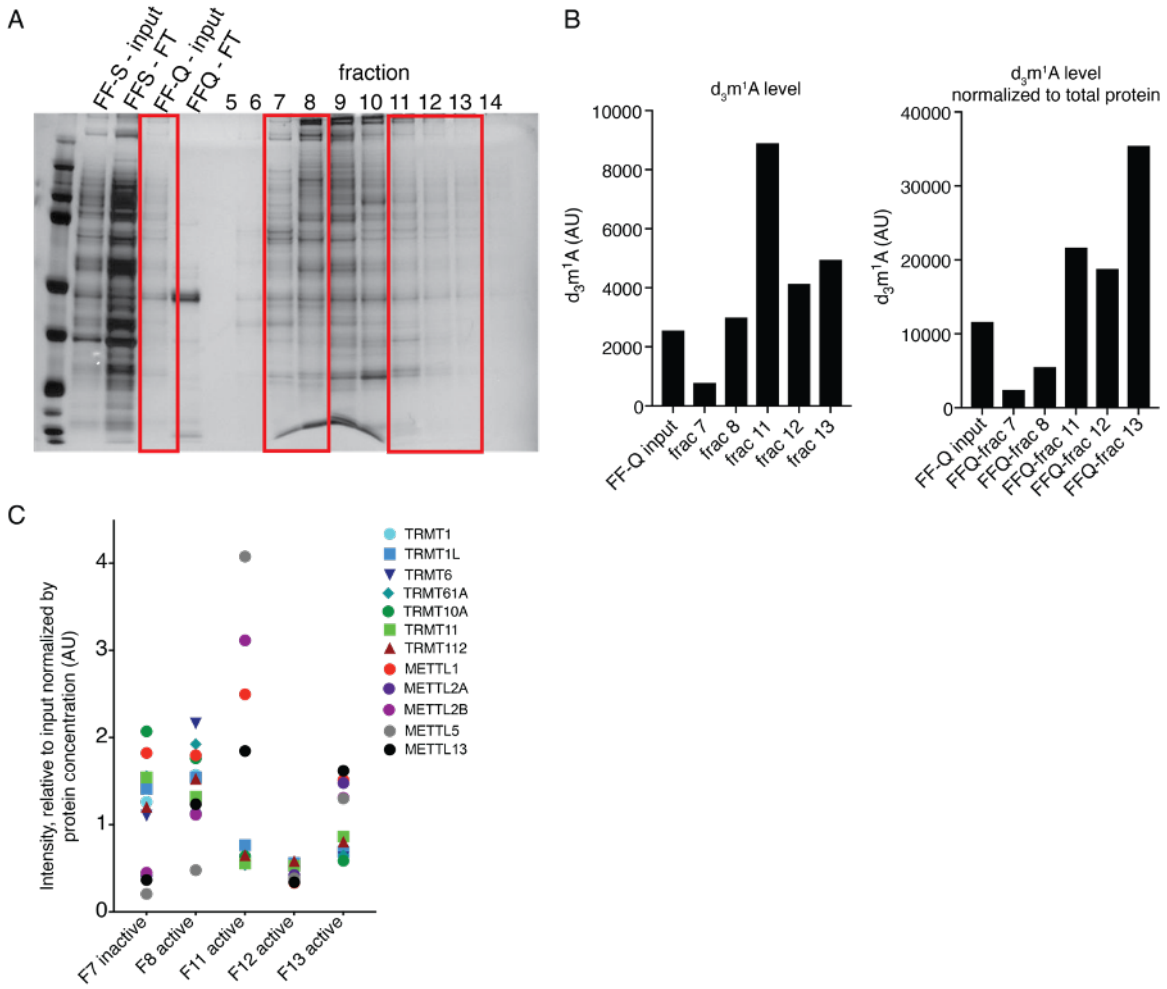


Figure 2.9. Biochemical fractionation to identify METTL5. (A) SDS-PAGE analysis of selected fractions, visualized by silver stain. The lanes indicated in red boxes correspond to those whose activity is shown in **(B)** and which were analyzed by TMT11-LC-MS/MS in **(C)**. FF-S, Fast Flow SP column; FF-Q, Fast Flow Q column.

Active fractions were then subjected to a second fractionation procedure, this time using a Fast Flow Q column (GE Healthcare) in 20mM Tris pH 7.5,1.5mM $MgCl_2$, 0.5mM EDTA using a gradient from 50mM to 1M NaCl (**Figure 2.9**A). Methyltransferase reactions were set up with each fraction as described above (**Figure 2.9**B) . After analysis of this second round of fractions for d_3m^1A activity by LC-MS/MS, fractions were selected for analysis by TMT11-LC-MS/MS proteomics analysis (see **Materials and Methods**: TMT11-LC-MS/MS), which yielded quantitation of numerous proteins in each active fraction. To narrow down possible candidates, the

abundance of all putative RNA modification enzymes was normalized by the total protein concentration of the original protein fractions, and plotted in **Figure 2.9**C. METTL5 appeared as an overrepresented protein in active fraction F11, and was pursued further as described in this manuscript.

MATERIALS AND METHODS

Cell culture

FreeStyle 293-F cells were obtained from Thermo Fisher (R79007), and HeLa (CCL-2) and HepG2 (HB-8065) cells were obtained from ATCC. HeLa and HepG2 cells were cultured in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum, sodium pyruvate, and penicillin/streptomycin (Gibco). FreeStyle 293-F cells were grown in suspension in FreeStyle 293 Expression Medium (Gibco).

Construction and validation of knockout cell lines

HeLa and HepG2 METTL5-KO cell lines were generated via CRISPR-Cas9-mediated gene disruption. Guides (see table below) were cloned into pSpCas9(BB)-SA-Puro (PX459) V2.0, a gift from Feng Zhang (Addgene plasmid # 62988 ; http://n2t.net/addgene:62988) (Ran et al., 2013). 100 picomoles of each guide and its reverse complement were phosphorylated with T4 polynucleotide kinase for 30 minutes at 37°C and then annealed by incubating at 95°C for 5 minutes and then ramping down to 25°C at 5°C per minute. Phosphorylated, annealed guides were ligated into BbsI-digested PX459 using T4 DNA ligase. Constructs were sequence verified prior to transfection into HeLa and HepG2 cells. Combinations of two constructs were used to generate deletions: PX459/METTL5-guide2 and PX459/METTL5-guide32 to generate a deletion across

exon 3 and 4, and PX459/METTL5-guide2 and PX459/METTL5-guide3 to generate a deletion across exons 2, 3, and 4 (**Figure 2.2A**). These combinations of plasmids were transfected into HeLa or HepG2 cells using Lipofectamine 2000 according to manufacturer instructions. After 48 hours, media was changed to media containing 1μg/mL of puromycin, and cells were allowed to grow for approximately a week, with media changes as needed as cell death progressed. Remaining cells were then trypsinized and diluted such that they could be plated at approximately 1 cell per well. Single cells were allowed to grow over the course of 2-4 weeks and collected and expanded as needed. METTL5 KO cell lines were identified via PCR to verify the appropriate deletion and verified by western blot.

Table 2.1. METTL5 guide sequences

M5 guide 2	CACCGAATTACTGTATCGAATGACT
M5 guide 2-RC	AAACAGTCATTCGATACAGTAATTC
M5 guide 3	CACCGAGAGAGTCGCCTGCAACAAG
M5 guide 3-RC	AAACCTTGTTGCAGGCGACTCTCTC
M5 guide 32	CACCGCAATGTCATCATAAGTGTTA
M5 guide 32-RC	AAACTAACACTTATGATGACATTGC

DNA constructs

METTL5 coding sequence was amplified from pCMV-Entry-METTL5-Myc-DDK (Origene cat# PS100001) by PCR. The PCR product and pCS2-FLAG (AddGene 16331) were digested with EcoRI and XhoI and ligated with T4 DNA Ligase. Point mutations were made using quickchange mutagenesis with primers overlapping mutated sequences and encoding the following mutations: amino acids 127-129 NPP>AAA, G61D, and K191Vfs*10delAA. pCS2-FLAG was a gift from Peter Klein (Addgene plasmid # 16331; http://n2t.net/addgene:16331; RRID:Addgene 16331).

FLAG-METTL5 co-immunoprecipitations

HeLa or Freestyle 293-F cells were transfected with constructs encoding FLAG-tagged METTL5 (WT or mutant) and were pelleted in cold 1x PBS by centrifugation at 1000g for 10 minutes at 4°C. Cells were lysed in 5 pellet volumes of lysis buffer containing 50mM Tris pH 7.4, 150mM NaCl, 10% glycerol, 1.5mM $MgCl_2$, 0.5mM EDTA, 0.2% NP-40, 1mM DTT, and 1x SigmaFast protease inhibitor cocktail (Sigma) and rotated for 1 hour at 4°C. The lysate was clarified by centrifugation at 20,000g for 20 minutes at 4°C. For immunoprecipitation, 30µL FLAG-M2 beads per 10 million cells transfected were washed into lysis buffer 3 times before adding to clarified lysate. The mixture of lysate and beads was rotated for 1 hour at 4°C, and then washed 3 times with 0.5mL of lysis buffer each time. After the final wash, 2x SDS with 100mM DTT was added to the beads and bound protein was eluted by heating at 65°C for 15 minutes, shaking at 1000rpm. Input and IP samples were separated on a 4-12% bis-tris gradient gel for western blotting (below).

Western blots

Cells were scraped into ice cold PBS and pelleted by centrifugation at 1000g for 10 minutes at 4°C. Cell pellets were resuspended in 3 pellet volumes of lysis buffer containing 50mM Tris pH 7.4, 300mM NaCl, 1% NP-40, 0.25% deoxycholate, 1mM DTT, and 1x SigmaFast protease inhibitor cocktail (Sigma). Lysates were kept on ice for one hour with periodic agitation, and then clarified by centrifugation at 20,000g for 20 minutes at 4°C. Protein concentrations were measured by bicinchoninic acid (BCA) assay (Pierce) for normalization. Protein samples were separated on 4-12% bis-tris gradient gels using MOPS (Thermo Fisher) or MES (Thermo Fisher) buffer, and then transferred to nitrocellulose membrane by semi-dry transfer at 17V for 50 minutes using 2x NuPage Transfer Buffer (Thermo Fisher) with 10% methanol. Membranes were blocked with 5% dry milk in 1x TBST (1x TBS with 0.1% Tween-20) for 30 minutes. Primary and secondary antibodies were diluted in 5% milk in TBST and incubated as indicated in the antibodies table (below). Membranes were imaged using a FluorChem R (ProteinSimple) or a ChemiDocMP (BioRad) and image analysis was performed with Fiji (Schindelin et al., 2012).

RNA purification

Total RNA was isolated from cells using Trizol (Invitrogen) according to the manufacturer's instructions. After collecting cells in Trizol, 0.2mL chloroform per 1mL Trizol was added and the samples were shaken vigorously for 15 seconds prior to incubating at room temperature for 2 minutes. Samples were then phase separated by centrifugation at 12,000g for 20 minutes at 4°C. The aqueous phase was collected, mixed with isopropanol (0.5mL per 1mL Trizol used), incubated at room temperature (RT) for 10 minutes and RNA was pelleted by centrifugation at 20,000g for 20 minutes at 4°C. The pellet was washed with 80% ethanol, centrifuged at 10,000g for 10 minutes

at 4°C, and then dried and resuspended in RNase free water. Polyadenylated RNA was isolated using the Dynabeads mRNA Direct kit (Life Technologies) as per manufacturer instructions. Ribosomal RNA was depleted using the RiboMinus Eukaryote Kit v2 (Life Technologies) as per manufacturer instructions.

siRNA transfection

One day prior to transfection, 150,000 HeLa cells were plated per 6cm cell culture dish in antibiotic-free cell culture media. Media was changed to fresh antibiotic-free media just prior to transfection with 20 picomoles siRNA and Lipofectamine RNAiMAX transfection reagent (Invitrogen). 48 hours after transfection, cells were harvested either in Trizol (for qPCR analysis) or in ice cold PBS (for western blot analysis).

Table 2.2. siRNAs used to investigate METTL5 function

Target/label	**Source**	**Cat number**	**Sequence (sense)**
Cont. #1	Sigma	SIC001-10NMOL	Proprietary sequence - siRNA Universal Negative Control #1
Cont. #2	Sigma	SIC002-10NMOL	Proprietary sequence - siRNA Universal Negative Control #2
METTL5 #1	Sigma	SASI_Hs01_00170949	CCAAGUCAUUCGAUACAGU[dT][dT]

Table 2.2, continued.

Target/label	Source	Cat number	Sequence (sense)
METTL5 #2	Sigma	SASI_Hs01_00170950	GAAGAGUUUGAGUUAACAA[dT][dT]
TRMT112 #1	Sigma	SASI_Hs01_00034520	CCGAUAACUUGCGUCUGAU[dT][dT]
TRMT112 #2	Sigma	SASI_Hs01_00034522	CGGCCGAUAACUUGCGUCU[dT][dT]
TRMT112 #3	Sigma	SASI_Hs01_00034521	GUCUGAUCCAGGUGCCGAA[dT][dT]

In vitro methyltransferase assays

Reactions were performed by combining the desired amounts of protein and total RNA or probe in 1x methyltransferase buffer (10mM HEPES pH 7.4, 4% glycerol, 100mM KCl, 1.5mM $MgCl_2$, 1mM DTT), with 2mM d_3SAM (CDN Isotopes) and 1mM (2% v/v) SUPERaseIN RNase inhibitor (Invitrogen), for a final volume of 50µL. Reactions were incubated at 37ºC for 1 hour before being snap frozen in liquid nitrogen and stored at -80ºC. Prior to LC-MS/MS analysis, samples were thawed, supplemented with 1mM EDTA, and treated with Proteinase K (Sigma) at 37ºC for 1 hour. Samples were purified by phenol/chloroform extraction and RNA precipitation (total RNA) or MyOne Streptavidin C1 beads (biotinylated probes; Invitrogen). For total RNA purification, an equal volume of acid phenol chloroform pH 4.5 (Invitrogen) was added to each sample and the mixture was mixed for two minutes and allowed to sit 2 minutes at RT. After centrifugation at

16,000g for 15 minutes at 4ºC, the aqueous layer was transferred to a new tube. 10% the volume of 3M sodium acetate pH 5.5 and 3 volumes of 100% ethanol were added, before mixing and incubating at -80ºC for 3 hours. RNA was pelleted by centrifugation at 20,000g for 20 minutes at 4ºC, and then the pellets were washed with 80% ethanol and resuspended in nuclease-free water. For probe purification, each sample was diluted into 350 uL of 1x IP buffer (50mM Tris pH 7.4, 300mM NaCl, 0.5mM EDTA, 1% v/v SUPERaseIN RNase inhibitor). Streptavidin C1 beads were pre-washed three times in 1x IP buffer, and 10µL bead resuspension was added to each sample. Samples were rotated for 2 hours at 4ºC, washed in 1x IP buffer once, and then eluted with Trizol by shaking at RT for 20 minutes. RNA was purified from Trizol using the DirectZol RNA Miniprep kit (Zymo Research).

LC-MS/MS

For each sample, 25-50ng of RNA was digested with 1µL Nuclease P1 (Sigma) in P1 buffer (25mM NaCl, 2.5mM $ZnCl_2$) in a final reaction volume of 20µL for 2 hours at 42ºC. Subsequently, 1µL of FastAP (Thermo Fisher) and 2.5µL of 10x FastAP buffer were added to each sample, and they were incubated at 37ºC for 4 hours. Samples were then diluted with an equal volume of water and filtered through a 0.2µm PVDF filter (0.2µm pore size, 0.4mm diameter, Millipore). 5µL of each filtered sample was separated by reverse phase ultra-performance liquid chromatography on a ZORBAX Eclipse XDB-C18 Rapid Resolution HT 2.1x50mm, 1.8µm column (Agilent) on an Agilent Technologies 1290 Infinity II liquid chromatography system, followed by mass spectrometry on a Sciex Triple Quad 6500 triple-quadrupole mass spectrometer in positive electrospray ionization mode. Nucleosides were quantified using nucleoside-to-base transitions of 282.101>150.100 (m^6A), 282.101>150.100 (m^1A), 284.982>153.100 (d_3m^6A), 284.983>153.200

(d_3m^1A), 296.101>164.100 ($m^{6,6}A$), 267.966>136.000 (A), and 284.004>152.100 (G). Standard curves were generated by injecting known concentrations of the corresponding pure nucleosides in the same run, and the percentage of modified to unmodified nucleosides were calculated based on the calibrated concentrations.

Quantitative PCR (qPCR)

cDNA synthesis was performed in 20µL reactions consisting of 1µL random hexamer primers at 50ng/µL (Invitrogen), 1µL 10mM dNTPs (NEB), 1µL 100mM DTT (Invitrogen), 1µL SUPERaseIN inhibitor (Invitrogen), 1µL Superscript IV reverse transcriptase (Invitrogen), 4µL 5x SSIV buffer, and nuclease-free water. Samples were incubated in a thermocycler at 23ºC for 10 minutes, 55ºC for 10 minutes, and 80ºC for 10 minutes, then held at 4ºC. 10µL qPCR reactions were set up with 1µL of cDNA, 5µL FastStart Essential DNA Green Master mix (Roche), and 0.1µL each of 10µM forward and reverse primers (see table below). qPCR was performed on a LightCycler 96 (Roche) with the following settings: 95ºC, 600 seconds (preincubation); 95ºC 20 seconds, 60ºC 20 seconds, 72ºC 20 seconds (three-step amplification). Analysis was performed with LightCycler 96 SW1.1 software (Roche).

Cell growth assays

To measure cell proliferation curves, four replicate 96-well plates were plated with 4 wells per cell line at 1000 cells per well (each plate represented one time point). For each time point, cell culture media was removed from each well and the plate was placed directly into a -80°C freezer for storage for the duration of the time course. The first time point was collected 3 hours after plating (just enough time for cells to adhere to the plate surface) and used as time 0, to normalize for small

differences in plated cell number. Subsequent time points were collected at 24, 48, and 96 hours post-plating, and frozen at -80°C for at least 24 hours. The number of cells per well was quantified using the CyQuant Cell Proliferation Assay Kit (Invitrogen) as per the manufacturer's instructions. Briefly, 200µL of lysis buffer was added to each well, pipetted up and down 4-6 times and left at room temperature for 5 minutes. 150µL of this lysate was transferred to a black, clear bottomed 96 well plate. Fluorescence was measured in a Synergy HTX plate reader (Biotek) using a 485/525 excitation/emission filter cube.

METTL5 binding partner ID

Freestyle 293-F cells were seeded at 10^6 cells/mL just prior to transfection with pCS2-FLAG-METTL5. 36 hours after transfection, cells were pelleted by centrifugation at 500g for 5 minutes at 4ºC. The cells were washed with ice cold PBS and pelleted by centrifugation at 500g for 5 minutes at 4ºC. The washed cell pellet was lysed in 5 pellet volumes of lysis buffer (50mM Tris pH 7.4, 50mM NaCl, 10% glycerol, 1.5mM $MgCl_2$, 0.5mM $CaCl_2$, 0.5% NP-40, 1x SigmaFast protease inhibitor [Sigma], 1x PhosSTOP [Sigma], and 10units/mL DNaseI [NEB]) by rotating for 1 hour at 4ºC. Lysate was first clarified by centrifugation at 10,000g for 10 minutes at 4ºC, and then the supernatant was transferred and clarified a second time by centrifugation at 20,000g for 20 minutes at 4ºC. 500µL FLAG-M2 magnetic beads (Sigma) were washed three times with 2mL lysis buffer per wash, and added to lysate and rotated for 30 minutes at 4ºC. After incubation with lysate, the beads were washed twice with 1.5mL lysis buffer, three times with 1.5mL high salt wash buffer (50mM Tris pH 7.4, 500mM KCl, 0.05% NP-40), and once with 1.5mL PBS with 0.1 Tween-20. Bound proteins were then eluted with 500uL of 2x Laemmli sample buffer (BioRad) supplemented with 100mM DTT by heating at 70ºC for 15 minutes, shaking at 1000rpm. Samples

were then separated on a 4-12% bis-tris gradient gel and stained by Silver Stain for Mass Spectrometry (Pierce). Bands of interest were excised using razor blades and stored in 1.5mL tubes at -80ºC prior to being shipped on dry ice and analyzed by MS BioWorks, LLC (Band ID service; Ann Arbor, Michigan, USA).

To prepare samples, trypsin digestion was performed using a robot (ProGest, DigiLab) and processed as follows: samples were washed with 25mM ammonium bicarbonate followed by acetonitrile, reduced with 10mM dithiothreitol at 60°C followed by alkylation with 50mM iodoacetamide at RT, digested with trypsin (Promega) at 37°C for 4 hours, quenched with formic acid and the supernatant was analyzed directly without further processing. The gel digest was analyzed by nano LC/MS/MS with a Waters NanoAcquity HPLC system interfaced to a ThermoFisher Q Exactive. Peptides were loaded on a trapping column and eluted over a 75µm analytical column at 350nL/min; both columns were packed with Luna C18 resin (Phenomenex). The mass spectrometer was operated in data-dependent mode, with MS and MS/MS performed in the Orbitrap at 70,000 FWHM resolution and 17,500 FWHM resolution, respectively. The fifteen most abundant ions were selected for MS/MS. Data were searched using a local copy of Mascot with the following parameters:

Enzyme: Trypsin

Database: Swissprot Human (forward and reverse appended with common contaminants)

Fixed modification: Carbamidomethyl

Variable modifications: Oxidation (M), Acetyl (Protein N-term), Pyro-Glu (N-term Q), Deamidation (NQ)

Mass values: Monoisotopic

Peptide Mass Tolerance: 10 ppm

Fragment Mass Tolerance: 0.02 Da

Max Missed Cleavages: 2

Mascot DAT files were parsed into the Scaffold software for validation, filtering and to create a nonredundant list per sample. Data were filtered using a minimum protein value of 90%, a minimum

peptide value of 50% (Prophet scores) and requiring at least two unique peptides per protein.

Fluorescence microscopy

20,000 cells were plated per acid-washed glass coverslip in a 24-well cell culture dish 24 hours prior to fixation with 4% paraformaldehyde in PBS for 10 minutes at RT. Coverslips were then washed 3 times with PBS and transferred to a humidity chamber. Coverslips were incubated in blocking solution (1X PBS, 1% goat serum, 10mg/mL BSA, 0.25% Triton X-100) for 30 minutes at RT, and then incubated with primary antibodies diluted in blocking buffer for 2 hours at RT. After three 5 minute washes with 1x PBS with 0.25% Triton X-100, coverslips were incubated with secondary antibodies diluted in blocking solution for 1 hour at RT. After three 5 minute washes with 1x PBS, the coverslips were mounted onto glass slides with ProLong Gold Antifade Mountant with DAPI (Life Technologies) and allowed to cure overnight at RT. Coverslips were imaged on an Olympus BX53 microscope with an Olympus U-HGLGPS light source using a 40X objective (NA 0.95). Antibodies used are listed in the table below.

Cellular fractionation

Cellular fractionation was performed essentially as previously described (Werner and Ruthenburg, 2015). HeLa and HepG2 cells were scraped into ice cold PBS and pelleted by centrifugation at 500g for 5 minutes at 4°C. The cell pellet was gently resuspended and washed in 5 pellet volumes of Buffer A (10mM HEPES pH 7.5, 4mM $MgCl_2$, 10mM KCl, 10% glycerol, 340mM sucrose, 1mM DTT, 1x SigmaFast Protease Inhibitor [Sigma]) three times, pelleting cells by centrifugation 500g for 5 minutes at 4°C between each wash. After the final wash, the pellet was resuspended in 2.5 pellet volumes of Buffer A. An equal volume of Buffer A + 0.2% Triton X-100 was added to this cell suspension with gentle mixing, and the samples then incubated on ice for 12 minutes with gentle inversion every 3 minutes. The lysate was then centrifuged 1200g for 5 minutes at 4°C, pelleting nuclei. The supernatant (cytoplasmic extract) was transferred to a new tube, and nuclei washed three times with Buffer A + 0.1% Triton X-100, pelleting nuclei by centrifugation at 500g for 5 minutes at 4°C between each wash and changing to fresh tubes once during these washes. The washed nuclear pellet was resuspended in one original cell pellet volume of NRB buffer (20mM HEPES pH 7.5, 75mM NaCl, 0.5mM EDTA, 50% glycerol, 1mM DTT, 1x SigmaFast Protease inhibitor [Sigma]). An equal volume of NUN buffer (20mM HEPES pH 7.5, 10mM $MgCl_2$, 300mM NaCl, 0.2mM EDTA, 1M urea, 1% NP-40, 1mM DTT, 1x SigmaFast Protease inhibitor [Sigma]) was gently added to the nuclear suspension and incubated on ice for 5 minutes, with periodic inversion. After centrifugation at 1200g for 5 minutes at 4°C, the supernatant (nuclear extract) was transferred to a new tube, and chromatin pellet washed twice with a 1:1 mixture of NUN/NRB buffer with centrifugation at 5000g for 3 minutes at 4°C between each wash. The pellet was then washed twice with Buffer A in the same manner and resuspended in Buffer A after the final wash. Cytoplasmic, nuclear, and chromatin extracts were then treated with

Benzonase (Sigma) for 30 minutes at 37°C before being separated by gel electrophoresis and analyzed by western blot.

Cross-linking assisted immunoprecipitation of METTL5 and sequencing

Immunoprecipitation: For each replicate experiment, fourteen 15cm plates of ~70% confluent HeLa cells were transfected with 5µg of pCS2-FLAG-METTL5 plasmid per plate, using Lipofectamine LTX Reagent (Invitrogen) as per manufacturer's instructions. 24 hours after transfection, the cells were treated with 100µM 4-thiouridine (Sigma) overnight. 36 hours after transfection, cells were washed with cold PBS and crosslinked twice with 150mJ per cm^2 254nM UV in a crosslinker (UVP CL-1000), with plates kept on ice for the duration. Cells were harvested by scraping in ice cold PBS and pelleting cells by centrifugation at 1000g for 10 minutes at 4ºC. Cell pellets were resuspended in 5 pellet volumes of lysis buffer (50mM Tris pH 7.4, 150mM NaCl, 10% glycerol, 1.5mM $MgCl_2$, 0.5mM EDTA, 0.5% NP-40, 1x SigmaFast protease inhibitor) and rotated for 1 hour at 4ºC. The lysate was clarified by centrifugation at 20,000g for 20 minutes at 4ºC, and an aliquot of lysate transferred to a new tube as an input sample. 200µL FLAG-M2 beads (Sigma) per 10mL lysate were washed into lysis buffer 3 times before being added to clarified lysate. The bead/lysate suspension was rotated for 2 hours at 4ºC, after which the beads were washed once with 1mL lysis buffer and transferred to new tubes, then 3 times with 1mL high salt buffer (50mM Tris pH 7.4, 500mM KCl, 0.05% NP-40), followed by two more 1mL washes with lysis buffer. The beads were then washed once with 1mL Proteinase K buffer (50mM Tris pH 7.4, 75 mM NaCl, 6.25mM EDTA, 1% SDS) before resuspending the beads in 450µL of this buffer (25µL of these beads were transferred to a new tube for analysis by western blot). Proteinase K (Sigma) was pre-incubated at 37ºC for 30 minutes prior to use. 50µL of Proteinase K was added

to beads resuspended in Proteinase K buffer. Inputs were similarly treated with Proteinase K by adding 225µL 2x Proteinase K buffer (100mM Tris pH 7.4, 150 mM NaCl, 12.5mM EDTA, 2% SDS) and 25µL Proteinase K to 250µL of input sample. Proteinase K treatment of both input and bead samples was done by incubating samples 50ºC for 1 hour, shaking at 1000rpm. Three sample volumes of Trizol were added to each, and RNA was purified using the DirectZol kit (Zymo) as per manufacturer's instructions. A technical note: our efforts to perform these experiments with the RNase digestion that is usually performed to reduce background and more precisely map protein binding sites were unsuccessful (Darnell, 2010), possibly due to the small footprint of METTL5 and/or the transient nature of interactions between RNAs and catalytic enzymes.

Preparation of sequencing libraries: Sequencing libraries were prepared using the TruSeq Stranded mRNA library preparation kit (Illumina), using 30-40 ng of RNA for each input and IP sample. Samples were sequenced in one lane of an SR50 flow cell on a HiSeq4000 (50bp single-end reads).

Data analysis: Analysis was performed very similarly to what would be done for an RNA immunoprecipitation (IP) experiment to assess which RNA transcripts are bound to a protein of interest. Three replicate experiments were performed as described above, with corresponding input and IP samples for each replicate. Quality of .fastq files was checked using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters were trimmed using Cutadapt (Martin, 2011) and files were then aligned to the human genome hg38 using hisat2 v2.1.0 (Kim et al., 2019) in splice-aware mode and resulting bam files were indexed and sorted with samtools version 1.7 (Li et al., 2009). DESeq2 (Love et al., 2014) and R version 4.0.3

(https://cran.r-project.org/) were used to analyze differentially expressed transcripts, thereby identifying transcripts that were enriched in the IP samples relative to input samples.

rRNA fragment purification

A 40-nt rRNA fragment was purified similarly to previously described (Ma et al., 2019c). Briefly, 3-4μg of a biotinylated DNA probe designed to bind to the target rRNA region was combined with 33μg of total RNA in 3.33x hybridization buffer (250mM HEPES pH 7, 500mM KCl) in a total volume of 150μL. The mixture was incubated at 90ºC for 7 min and then allowed to slowly cool to room temperature over 3.5 hours to allow hybridization. Single-stranded RNA and DNA were digested by adding 0.5μg RNase A (Thermo Scientific) to the mixture and incubating at 37ºC for 30 minutes, then 30 units of Mung Bean Nuclease were added along with 10x Mung Bean Nuclease buffer (NEB) to a final concentration of 1x, followed by a 30 minute incubation at 37ºC. Streptavidin T1 beads (20μL per sample; Invitrogen) were washed 3 times in 1mL 2.5x IP buffer (375mM NaCl, 125mM Tris pH 7.9, 0.25% NP-40), then resuspended in 100μL 2.5x IP buffer. The beads were combined with the 150μL RNA-DNA hybridization mix and rotated at 4C for 1 hour. After rotation, the beads were washed 3 times in 1x IP buffer (150mM NaCl, 50mM Tris pH 7.9, 0.1% NP-40), twice in nuclease-free water, and resuspended in 30μL nuclease-free water. The resuspended beads were heated at 70ºC for 5 minutes to denature the RNA from the probes, and quickly placed on ice before placing on a magnet to collect the RNA-containing supernatant.

m^6A-seq

Experiment: 1μg of 1x ribo-depleted RNA per sample (from HeLa METTL5-WT and METTL5-KO cells) was fragmented using the Bioruptor Pico sonicator (Diagenode) 30 sec on, 30 sec off,

30 cycles. About 1/20th of each sample was saved as input, and immunoprecipitation was performed using the EpiMark *N*^{6}-methyladenosine Enrichment Kit (NEB) as per manufacturer's instructions. Briefly, 1µL m^{6}A antibody per sample was coupled to pre-washed Protein G Magnetic Beads (NEB) in 1x reaction buffer (150mM NaCl, 10mM Tris-HCl pH 7.5, 0.1% NP-40 in nuclease-free water) with orbital rotation for 30 minutes at 4ºC. Beads were washed, and fragmented RNA was added in 1x reaction buffer with 2% v/v SUPERaseIN Rnase inhibitor (Invitrogen). Beads were washed twice each in 1x reaction buffer, low salt reaction buffer (50mM NaCl, 10mM Tris-HCl pH 7.5, 0.1% NP-40 in nuclease-free water), and high salt reaction buffer (500mM NaCl, 10mM Tris-HCl pH 7.5, 0.1% NP-40 in nuclease-free water), then resuspended at room temperature in 30µL Buffer RLT (Qiagen) per sample. Eluates were purified using the RNA Clean & Concentrator-5 kit (Zymo Research). Libraries were then constructed using the SMARTer Stranded Total RNA-Seq Kit v2 (Takara) per manufacturer's instructions and pooled and sequenced evenly across one lane of an Illumina NovaSeq6000 SP flow cell with 50bp paired-end reads.

Data analysis: Quality of .fastq files was checked using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). As instructed by the SMARTer Stranded Total RNA-Seq Kit v2 manual, the 3 bases at the start of all R2 reads were removed using Trimmomatic 0.39 (Bolger et al., 2014) with the option HEADCROP:3. Adapter trimming and quality filtering were also performed with Trimmomatic using paired-end mode with options LEADING:3, TRAILING:3, SLIDINGWINDOW: 4:15, and MINLEN: 21. To check for contamination, files were aligned to the mycoplasma genome using hisat2 v2.1.0 (Kim et al., 2019) and any matching reads were discarded. Files were then aligned to the human genome hg38 using

hisat2 v2.1.0 in splice-aware paired-end mode and resulting bam files were sorted with samtools version 1.7 (Li et al., 2009). Using the MeRIPTools package (Zhang et al., 2020b), reads were counted with the function CountReads, and peaks were called using Fisher's exact test with the function callPeakFisher. The R package QNB was used for inferential testing and differentially methylated peaks were called with an adjusted P-value < 0.1 (Liu et al., 2017a). Motif searches were performed using HOMER v4.11 (Heinz et al., 2010). For a background reference, sequences were extracted from random 200bp peaks that were sampled from an mRNA transcript (Zhang et al., 2020b).

Polysome and ribosome profiling

Polysome profiling was performed similarly to previous reports (Wang et al., 2015; Zhang et al., 2019). Four 15cm plates per sample were cultured to ~80% confluence. Each plate was treated with cycloheximide at a final concentration of 100µg/mL in antibiotic-free medium and incubated at 37ºC for 7 minutes. Plates were washed with 10mL of ice-cold PBS with 100µg/mL cycloheximide twice, then cells were collected in ice-cold PBS and centrifuged at 500g for 5 minutes. Each pellet was resuspended in 3 pellet volumes of lysis buffer (20mM HEPES pH 7.6, 100mM KCl, 5mM MgCl2, 1% Triton x-100, 100µg/mL cycloheximide, 1% v/v SUPERaseIN inhibitor in nuclease-free water) and cells were lysed with rotation at 4ºC for 20 minutes before being centrifuged at 16,000 g for 15 minutes. To the clarified lysate, 4µL of Turbo DNase (Thermo Fisher) was added and samples were incubated at room temperature 15 minutes then centrifuged again to clear the lysate. The absorbance at 260nm of each lysate was measured by Nanodrop (Thermo Fisher) and samples were adjusted to the same optical density with lysis buffer. One fifth of each sample was saved as input, whereas the remaining portion of each sample was loaded on

a 5-50% w/v sucrose gradient prepared in lysis buffer in a SETON Scientific open top polyclear tube (cat# 7042). Samples were then centrifuged at 28,000 rpm for 3 hours at 4ºC using an Optima L-100XP centrifuge with an SW28 rotor. After centrifugation, absorbance was read and fractions were collected by Gradient Station (BioComp) equipped with an ECONO UV monitor (BioRad) and fraction collector (Gilson). The max absorbance (AUFS) was set to 0.5, number of fractions to 30, and distance/fraction to 3.2mm.

Ribosome profiling was performed similarly, with a few key differences. Firstly, after DNase treatment, about 20% of the sample was saved as input with 3 volumes of Trizol added, while the other 80% was treated with 3µL MNase (NEB) in 1x MNase buffer for 15 minutes at room temperature to digest RNA not covered by ribosomes. Additional SUPERaseIN (Invitrogen) was added to a final concentration of 0.5mM to quench the reaction. Then, samples were loaded on 5-50% w/v sucrose gradients and fractionated as above. Secondly, after fractions were collected, samples were processed for sequencing. Fractions containing the digested monosomes were collected and pooled, and three volumes of Trizol (Invitrogen) was added to each sample. RNA purification was performed for both input and ribosome-protected fragment (RPF) samples by chloroform extraction followed by isopropanol precipitation, and one round of rRNA depletion was performed (see "RNA purification" section). Input RNA was fragmented by combining 20µL RNA with 7.5µL PNK buffer A (Thermo Fisher) and incubating at 94ºC for 25 minutes, then end repaired by diluting in 37.5µL nuclease-free water, adding 10µL T4 PNK (Thermo Fisher) and incubating at 37ºC for 30 minutes. Then, 12µL 10mM ATP (EMD Millipore), 2µL 10x PNK buffer A, and 6µL T4 PNK were added to incubate at 37ºC for 30 minutes. These samples were purified using the RNA clean and concentrator-5 kit (Zymo Research). Purified RPF RNA was run in 1x

Novex TBU sample buffer (Thermo Fisher) on a 10% TBE-urea gel (Thermo Fisher) in 1x RNase-free TBE buffer (Thermo Fisher) with an IDT 10/60 loading control (IDT) for 1 hour at 180V. The gel was stained 15 minutes in 1x SYBR gold in TBE with gentle shaking and visualized using a Gel Doc EZ Imager (Biorad). Fragments between ~26-34 nucleotides (corresponding to the region between ~20-30nt of the IDT loading control) were excised and purified by Small-RNA PAGE recovery kit (Zymo Research) per manufacturer's instructions. RPFs were end repaired similarly to inputs with T4 PNK. Input and RPF RNA samples were used to prepare libraries for sequencing using the NEBNext Small RNA Library Prep Kit (NEB) per manufacturer's instructions, and were sequenced by Illumina NovaSeq6000 with 100bp single-end reads on two SP flowcells.

Ribosome and polysome profiling of mouse tissues were performed similarly to the method in cell lines, with a few specific changes in sample acquisition, cycloheximide treatment, and lysis (Seimetz et al., 2019). Mice were sacrificed and livers were removed and snap-frozen very quickly following standard procedures (Seimetz et al., 2019), then stored at -80ºC . Each tissue sample was ground to powder under liquid nitrogen by mortar and pestle, and the powder was transferred to a sterile 10 cm cell culture plate on dry ice. An appropriate volume (~1 mL per ~200mg tissue) of lysis buffer containing cycloheximide (10mM Tris-HCl pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 0.5% w/v sodium deoxycholate, 1% v/v NP-40 alternative, 10 mM DTT, 150 µg/mL cycloheximide, 1 U/µL SUPERase In RNase inhibitor, and 1x EDTA-free Roche protease inhibitor cocktail in RNase-free water) was added to the powder sample immediately upon removal from dry ice to resuspend it, and the plate was placed on ice for 1-2 minutes. Each sample was resuspended by gently pipetting 10 times before transferring it to a clean microcentrifuge tube, and leaving on ice for 3-5 more minutes to ensure complete lysis. Samples were spun at 2000g for 5 minutes at 4ºC

to pellet the nuclei and tissue debris. The supernatant was transferred to a fresh tube and spun at 12,000g for 5 minutes at 4ºC to pellet the remaining cellular debris. Again the supernatant was transferred to a fresh tube to start DNase treatment. From there, the same protocols used for polysome and/or ribosome profiling of cell lines were followed, except that the max absorbance (AUFS) was set to 1.0 during fraction collection. Mouse ribosome profiling libraries were sequenced by Illumina NovaSeq6000 with 50bp paired-end reads on two SP flowcells.

Data analysis: Analysis of polysome profiles was performed using Microsoft Excel and GraphPad Prism. For ribosome profiling from cell lines, fastq files from the two flowcells were merged using the linux command "cat". For all ribosome profiling sequencing data, data quality was checked with FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters were trimmed and quality filtering was performed with BBDuk from BBTools version 38.84 using a reference fasta file of the NEBNext adapter sequences and settings ktrim = 4, qtrim = rl, trimq = 20, k = 31, mink = 11, hdist = 1, entropy = 0.5, entropywindow = 30, entropyk = 5, minlen = 25 and maxlen = 34 (for RPFs), and minlen = 15 (for inputs) (BBtools: (https://sourceforge.net/projects/bbmap/). For input files, alignment to human genome hg38 (Frankish et al., 2019) was performed using STAR 2.7.3a (Dobin et al., 2013) in quantMode TranscriptomeSAM. Bam files were sorted and uniquely mapped reads extracted with samtools v1.7 (Li et al., 2009), followed by featureCounts v1.6.0 (Liao et al., 2014) to create a counts table from mapped reads with options -s 1, -g gene_id, and -t exon. rRNA reads were removed from RPF Fasta files using BBDuk with a reference fasta file of human rRNA sequences. Analysis was then performed on the RiboToolkit server (http://rnabioinfor.tch.harvard.edu/RiboToolkit/) (Liu et al., 2020b). First, RPF files were uploaded for single-case submission in collapsed fasta format

with the following options: species = homo sapiens (hg38), RPF interval = 25-34, allowed mismatch = 2, max multiple-mapping = 1, no duplicate removal, offsets to infer P-sites - calculate by RiboToolkit, min coverage for pause sites = 10, fold change for pause sites = 20, ORF p-value = 0.05. Input counts table and single case Job IDs were then used as inputs for group case analysis with F value = 2 and P-value = 0.05.

Analysis was performed similarly for mouse liver ribosome profiling data except that alignments were performed to the mm10 genome (Frankish et al., 2019), and changes were made to accommodate 50bp paired-end data for inputs. For example, BBDuk and STAR 2.7.3a were run in paired-end mode with paired input files, and FeatureCounts was run using the flags -p and -countReadPairs. The STAR reference genome generated before mapping was made with sjdbOverhang of 49 (rather than 100) to accommodate the shorter reads. For RPF samples, only read 1 was used, and data was analyzed as above for cell line data. For RiboToolkit single-case submissions, mus musculus (mm10) was selected for the species.

RNA-seq, METTL5-KO HeLa cells

Experiment: Three biological replicates of wild type and 2 different clones METTL5-KO HeLa cells (clone 1 and 7) were grown to confluency prior to lysis in Trizol (Invitrogen). Total RNA was purified as described above ('RNA purification') and sequencing libraries were generated using an HIV reverse transcriptase evolved for m^1A detection, as previously described (Zhou et al., 2019a). Libraries were sequenced using an Illumina NovaSeq6000 on an S1 flowcell with 100bp paired end reads. For RNA-seq analysis, only read 1 (R1) was used.

Analysis: Quality of fastq files was checked using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters were trimmed using Cutadapt (Martin, 2011). To check for contamination, files were aligned to the mycoplasma genome using hisat2 v2.1.0 (Kim et al., 2019) and any matching reads were discarded. Files were then aligned to the human genome hg19 (Frankish et al., 2019) using hisat2 v2.1.0 (Kim et al., 2019) and resulting bam files were indexed and sorted with samtools version 1.7 (Li et al., 2009). Differential expression analysis was performed using DESeq2 (Love et al., 2014) and R version 4.0.3 (https://cran.r-project.org/). Gene ontology analysis was performed with MetaScape (Zhou et al., 2019b).

RNA-seq, mouse tissues

Experiment: Mouse organs were collected immediately upon sacrifice, washed in ice-cold PBS, placed in ~4 volumes of Trizol (Invitrogen), and then sonicated using a hand-held sonicator (OMNI International) with a few brief pulses until there were no more visible chunks. Samples were stored at -80ºC until use. Libraries were prepared using the SMARTer Stranded Total RNA-Seq Kit v2 (Takara) per manufacturer's instructions and sequenced on a NovaSeq6000 with 100 bp single-end reads.

Data analysis: Sequencing data quality was checked by FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapter trimming and quality filtering was performed with Trimmomatic (Bolger et al., 2014) in single-end mode with options LEADING:3, TRAILING:3, SLIDINGWINDOW:4:15, and MINLEN:36. Reads were then aligned to the mm10 genome (Frankish et al., 2019) using STAR 2.7.3a (Dobin et al., 2013). Read

counts mapping to each gene were obtained using featureCounts (Liao et al., 2014) with options -s 2, -g gene_id, and -t exon. Differential expression analysis was then performed in R version 4.0.3 (https://cran.r-project.org/) using DESeq2 1.28.1 (Love et al., 2014). Gene ontology analysis was performed with MetaScape (Zhou et al., 2019b).

Knockout mouse generation and mouse husbandry

Mettl5 knockout mice were generated at the Gene Targeting & Transgenic Facility at Janelia Research Campus. A *Mettl5* conditional knockout line was made in which exon 2 was floxed, such that presence of Cre would induce exon 2 deletion by creating a frame-shift mutation. The construct was electroporated with gRNA/Cas9 protein into ES cells that are F1 hybrid of C57bl/6 x 129S6. F1 mice were transferred from Janelia Research Campus to the University of Chicago where the second generation was backcrossed to the B6 strain. To generate a whole body knockout, the conditional knockout strain described above was then crossed with the B6.C-Tg(CMV-cre)1Cgn/J line (The Jackson Laboratory, 006054) (Schwenk et al., 1995). All mice were housed in a specific pathogen-free facility. Both male and female mice were used throughout the study, between 4 weeks and 6 months of age. All experiments were approved by the University of Chicago Institutional Animal Care and Use Committee.

ELISA

Blood was collected retro-orbitally from mice upon sacrifice, allowed to coagulate at room temperature for about one hour, and centrifuged at 10,000g for 15 minutes at 4ºC. The supernatant (serum) was transferred to a new tube and stored at -80ºC. The thyroid hormone (T3) concentration

was assayed using the Biomatik Mouse Tri-iodothyronine, T3 ELISA Kit (Cat# EKC40189), per manufacturer's instructions.

Mouse behavioral assays

Rotarod Experiment: A computer-controlled rotarod apparatus (Rotamex-5, Columbus Instruments) with a rat rod (7cm diameter) was set to accelerate from 4 to 40 revolutions per minute over 300 seconds, and time to fall was recorded. Mice received five consecutive trials per session, one session per day (about 30 seconds between trials).

Open field: Open field chambers were 40cm x 40cm (Med Associates) with lighting at 21 lux. Infrared beams recorded the animals' locomotor activity and rearing movements (vertical activity). Mice were put in the open filed chamber for 1 hour to record their activity.

Passive avoidance: The shuttle box used contained two chambers: one chamber illuminated and the other dark (Kinder Scientific). Mice were transported to the behavior room and were handled for three minutes for 3 days before the passive avoidance experiment. During tasks, the right chamber remained illuminated while the left chamber remained dark. Training began by placing the mouse into the illuminated chamber facing away from the shut guillotine door. The mouse was allowed to explore the illuminated chamber for 2 minutes. The door was then opened to let the mouse explore both the illuminated and dark chambers for 5 minutes. At the end of this exploration period, the door was shut after returning the mouse into the illuminated chamber. Two minutes later, the door was opened. Latency to step into the dark chamber was recorded by the computer as the baseline. Upon entering the dark chamber, the door was closed and one foot shock (0.2 mA,

2 seconds) was delivered. Ten seconds later, the mouse was removed from the dark chamber and put back to the home cage. After 24 hours, the mouse was put into the light chamber for 2 minutes and then the latency to step into the dark chamber was recorded as the 24 hour memory.

FR1: Mice were food deprived each day before the experiment. Mice had unlimited food access for 3 hours each day after the experiment. Each mouse was first trained in an illuminated operant box (Med Associates, St. Albans, VT, USA) on a FR1 schedule for 30 minutes each day for about three weeks, with additional random delivery of a food pellet for the first 15 minutes (one pellet per minute on average). The maximum number of food pellets by pressing the lever was set at 40. Mice were then trained on a FR1 schedule without free delivery of the food pellets. After the mice achieved more than 15 presses for two consecutive days, the results of the last of day FR1 were compared between genotypes.

Statistical analyses

Statistical analyses were done in GraphPad Prism. Throughout the figures, replicates are plotted as individual points, along with mean and standard error of the mean (s.e.m.). For experiments comparing two conditions, unpaired t-test was used. For experiments with multiple comparisons, one-way ANOVA was used to test specific comparisons as indicated (that is to say, not every single sample was compared to every other sample). A Dunnett's test was used to correct for multiple comparisons when all comparisons were made to a single control (**Figures 2.1F, Figure 2.3E, Figure 2.4G**). When multiple comparisons were made between experimentally determined pairs, the Sidak method was used to correct for multiple comparisons (**Figure 2.7A, Figure 2.8J**). In experiments with multiple variables and parameters tested, a two-way ANOVA with the

appropriate multiple comparison correction was used (Sidak test for **Figure 2.3F**, Tukey test for **Figure 2.3G**).

TMT11 proteomic analysis of biochemically active fractions

HeLa nuclear extract was fractionated into 11 fractions and a few active fractions were identified. Each fraction was loaded on 10% SDS-PAGE gel and run a short time with little separation. The proteins were then in-gel digested and labeled with 11 different tandem mass tags (TMT; Thermo Scientific). The TMT labeled peptides were pooled and separated into 6 fractions by an Ultra-micro spin C18 column by eluting with 10%, 20%, 25%, 30%, 35%, and 70% of buffer B respectively (buffer A: 10 mM NH_4COOH (pH=8.0); buffer B: buffer A plus 90% acetonitrile). Every fraction was further separated on a reverse phase column (30 cm X 75 μm, 1.9 μm C18 resin) during a 4 h gradient of 12-36% buffer B (buffer A: 0.2% formic acid, 3% DMSO; buffer B: buffer A plus 65% acetonitrile) and analyzed by Q-Exactive HF (Thermo Scientific) with one MS scan and up to 20 data-dependent high-resolution MS/MS scans. The MS/MS raw files were processed using the JUMP searching engine against the UniProt human database. Searches were performed using 15 ppm mass tolerance and allowing up to two missed trypsin cleavage sites. TMT tags on lysine residues and peptide N termini (+229.162932 Da) and carbamidomethylation of cysteine residues (+57.021 Da) were used for static modifications and the dynamic modifications included oxidation of methionine residues (+15.99492 Da). Proteins were quantified by summing reporter ion counts across all matched PSMs using our in-house software. The mass spectrometry proteomics data are available via the PRIDE database (http://www.proteomexchange.org) under accession number PXD028832.

Table 2.3. Table of primers and oligonucleotides for METTL5 investigation.

Description	**Application**	**Sequence**
Mouse Mettl5 FOR	qPCR	AGAACAGTATCCCACCAGGC
Mouse Mettl5 REV	qPCR	ATCCAACACACAACCCTGCT
Human METTL5 FOR	qPCR	GGGTTAGCCGGGAGATCCT
Human METTL5 REV	qPCR	GCAGGCGACTCTCTAGTTCC
Human TRMT112 FOR	qPCR	GGCCGATAACTTGCGTCTGA
Human TRMT112 REV	qPCR	GGGTGCCCTCTATCACTTCC
Human HPRT1 FOR	qPCR	TGACACTGGCAAAACAATGCA
Human HPRT1 REV	qPCR	GGTCCTTTTCACCAGCAAGCT
Mouse Hprt FOR	qPCR	CTGGTGAAAAGGACCTCTCGAAG
Mouse Hprt REV	qPCR	CCAGTTTCACTAATGACACAAACG
28S biotinyl. DNA probe (4200-4240)	rRNA fragment purif.	Biotin-CTCGCCTTAGGACACCTGCGTTACCGTTTGACAGGTGTAC

Table 2.3, continued.

Description	Application	Sequence
18S biotinyl. DNA probe (1821-1860)	rRNA fragment purif.	Biotin-TTCCGCAGGTTCACCTACGGAAACCTTGTTACGACTTTTA
28S rRNA 12mer	In vitro Mtase assay	CGGUAACGCAGG-biotin
18S rRNA 12mer	In vitro Mtase assay	UCGUAACAAGGU-biotin
18S rRNA 60mer	In vitro Mtase assay	UCUAGAGGAAGUAAAAGUCGUAACAAGGUUUCCG UAGGUGAACCUGCGGAAGGAUCAUUA-biotin

Table 2.4. Antibodies used for METTL5 investigation.

Target	Source/catalog number	Application/species/dilution
anti-METTL5	ProteinTech/16791-1-AP	Western blot/rabbit/1:500 dilution
anti-METTL5	Atlas/SAB2101471	IF/rabbit/1:250 dilution
anti-TRMT112	Santa Cruz Biotechnology/sc-398481	Western blot/mouse/1:100 dilution
anti-FLAG-M2	Sigma/A8592	Western blot/mouse/1:10,000 dilution
anti-RPL4	Proteintech/11302-1-AP	Western blot/rabbit/1:1000 dilution

Table 2.4, continued.

Target	Source/catalog number	Application/species/dilution
anti-RPS6	Abcam/70227	Western blot/rabbit/1:1000 dilution
anti-EIF3A	Cell Signaling Technology/3411	Western blot/rabbit/1:1000 dilution
anti-EIF4E	Cell Signaling Technology/2067	Western blot/rabbit/1:1000 dilution
anti-GAPDH-HRP	GenScript/A00192-100	Western blot/goat/1:10,000 dilution
anti-SNRP70	Abcam/ab83306	Western blot/rabbit/1:1000 dilution
anti-histone H3	Cell Signaling Technology/D1H2	Western blot/rabbit/1:2000 dilution
anti-fibrillarin	Abcam/ab4566	IF/mouse/1:250 dilution
goat-anti-rabbit-488	Thermo Fisher/A11008	IF secondary, 1:500 dilution
goat-anti-mouse-594	Thermo Fisher/A11005	IF secondary, 1:500 dilution
goat anti-rabbit HRP	Thermo Fisher/A16110	Western blot, secondary, 1:5000 dilution
goat anti-mouse HRP	Thermo Fisher/A16078	Western blot, secondary, 1:5000 dilution
donkey-anti-rabbit DyLight 800	Thermo Fisher/SA5-10044	Western blot, secondary, 1:5000 dilution

Table 2.4, continued.

Target	Source/catalog number	Application/species/dilution
donkey-anti-rabbit DyLight 680	Thermo Fisher/SA5-10170	Western blot, secondary, 1:5000 dilution

DATA AVAILABILITY

Raw and processed data files from all high throughput sequencing experiments have been deposited in the NCBI Gene Expression Omnibus (GEO) with the following accession numbers: GSE174435 (METTL5 CLIP-seq), GSE174503 (RNA-seq of HeLa cells), GSE174420 (m^6A-seq), GSE174418 (RNA-seq of mouse tissues), GSE174419 (HepG2 ribosome profiling), GSE188798 (mouse ribosome profiling). The mass spectrometry proteomics data are available via the PRIDE database (http://www.proteomexchange.org) under accession numbers PXD028832 (TMT11) and PXD029574 (METTL5 IP).

CHAPTER 3: LOCATION-SPECIFIC EFFECTS OF N^6-METHYLADENOSINE ON mRNA TRANSCRIPT STABILITY

Preface Note:

Figures 3.1 through 3.5 of the following section (Chapter 3) comprise analyses of public datasets solely performed by me. The two final figures (Figure 3.6, Figure 3.7) of the section include work done in collaboration with Zhongyu Zou (Z.Z.), published on BioRxiv (Zou et al., 2022), which was submitted for publication with me as second author and is currently under review. The full citation is: Zou, Zhongyu, Caraline Sepich-Poore, Xiaoming Zhou, Jiangbo Wei, and Chuan He. (2022). "The mechanism underlying redundant functions of the YTHDF proteins." bioRxiv. https://doi.org/10.1101/2022.05.05.490669. Specifically, Z.Z. performed the P body imaging, purification, and sequencing experiments shown (Fig. 3.7A-D), and I performed analyses involved in generating all parts of Figure 3.6. Overall, unless otherwise explicitly noted, all other work presented in this chapter was done by me, with guidance and feedback provided by Prof. Chuan He. Additional figures and information for the project not included in this dissertation can be found on BioRxiv (Zou et al., 2022).

INTRODUCTION

The nature of mRNA transcripts is fundamentally directional, as they are transcribed from DNA from 5' to 3' ends. This directionality involves regional divisions that determine key mechanisms of translation: the 5'UTR, CDS, and 3'UTR. For example, the process of canonical cap-dependent translation begins with the binding of translation initiation factors to the 5'cap, followed by scanning of the 5'UTR, assembly of the full ribosome at the start codon, and translation elongation through the CDS, ending at the stop codon (Leppek et al., 2018). Both the

5’UTR and 3’UTR comprise regulatory regions to which factors may bind and regulate translation (Leppek et al., 2018). Similarly, mRNA decay and stability are spatially regulated. Predominant cellular mechanisms of decay involve exonuclease activity initiating from either 5’ or 3’ ends (Lee et al., 2020). Conversely, the 5’ cap and polyA tail at the 3’ end contribute to protecting transcripts from these decay mechanisms (Lee et al., 2020; Nicholson and Pasquinelli, 2019; Ramanathan et al., 2016; Roundtree et al., 2017a; Torabi et al., 2021). RNA binding proteins involved in mRNA stabilization and decay also exhibit transcript regional specificity, including cap-interacting and polyA tail-interacting factors that promote stability or decay (Evdokimova et al., 2001; Torabi et al., 2021).

m^6A on mRNA commonly correlates with increased transcript decay, as demonstrated by how METTL3 (an m^6A writer) loss leads to increased half-life of methylated transcripts (Lasman et al., 2020; Wang et al., 2014a). YTHDF2 was the first m^6A reader shown to promote decay, and its targets similarly exhibited significantly increased half-lives upon YTHDF2 depletion (Wang et al., 2014a). This effect was mediated by interactions between YTHDF2 and the CCR4-NOT deadenylase complex, which accelerates mRNA degradation from the 3’ end (Du et al., 2016). Another mechanism was later uncovered involving YTHDF2-driven recruitment of the HRSP12/RNaseP/MRP endonuclease complex to methylated targets (Du et al., 2016; Park et al., 2019).

Critically, both of these transcript decay mechanisms are dependent on regional specificity. In the first mechanism, the CCR4-NOT deadenylase complex must initiate decay at the 3’ end, with the location of YTHDF2 binding facilitating this targeting. In the second, HRSP12 tends to bind approximately 800 nucleotides upstream of YTHDF2, and cleavage by RNaseP/MRP preferentially occurs about 400 nucleotides downstream of YTHDF2 (Lee et al., 2020; Park et al.,

2019). Interestingly, it has been proposed that a mechanism may exist through which YTHDF2 promotes mRNA decapping (Lee et al., 2020), but direct evidence of this does not yet exist.

Curiously, YTHDF2 depletion has been shown to cause upregulation and downregulation of different mRNA transcripts in about equal numbers (Dixit et al., 2021; Wang et al., 2014a). Recently, Dixit *et al*. reported that YTHDF2 can actually stabilize MYC and VEGFA transcripts in glioblastoma stem cells (GSCs) in an m^6A-dependent manner (Dixit et al., 2021). This stabilization effect linked YTHDF2 to tumor growth and potentially a clinically-relevant mechanism (Dixit et al., 2021).

Our group also recently uncovered a potential role for YTHDF2 in stabilization of methylated transcripts with regional specificity (Dixit et al., 2021; Hu et al., 2022). Analysis of single-base-resolution m^6A sequencing data, using a new method called m^6A-SAC-seq, revealed that transcripts with m^6A in the 5'UTR display longer lifetimes than those with m^6A in other regions or those that are unmethylated; moreover, this effect was reversed upon YTHDF2 knockdown (Hu et al., 2022). If validated, this 5'UTR-specific stabilization effect would contrast with canonical roles of YTHDF2, and it would further highlight the importance of location-specific mRNA methylation(s) as a regulator of RNA binding protein functions.

YTHDF2 is not the only m^6A reader that regulates mRNA stability and decay. For example, YTHDF3 was shown to facilitate decay in synergy with YTHDF2 (Shi et al., 2017a). Because triple knockdown of YTHDF1-3 promotes dramatic accumulation of methylated transcripts, some have proposed that all three YTHDF proteins actually have redundant functions in promoting mRNA decay (Lasman et al., 2020; Shi et al., 2017a; Zaccara and Jaffrey, 2020). However, conflicting studies have indicated translation promotion roles for YTHDF1, as well as YTHDF3 in certain cases (Rauch et al., 2018; Shi et al., 2017a; Wang et al., 2015). Furthermore,

in mESCs, triple YTHDF knockdown also led to the stabilization of unmethylated mRNA (Lasman et al., 2020; Rauch et al., 2018; Shi et al., 2017a; Wang et al., 2015). These data question whether functional decay redundancies among YTHDF proteins accurately explain YTHDF1-3 knockdown-mediated transcript stabilization, or whether other factors/mechanisms are at play.

One line of investigation that may resolve the apparent discrepancies is to take into account m^6A regional specificity along the transcript and in the cell. As described above, it is well known that the regional location of m^6A on mRNA affects its functional potential, including proximity to the cap and polyA tail, and its influence on RNA binding proteins and regulatory factors. Similarly, the cellular location of a m^6A-methylated transcript can influence mRNA stabilization or decay. For example, different decay pathways dictate degradation of mRNA in the nucleus versus the cytosol (Cheng et al., 2021; Lee et al., 2020). Cytosolic granules, including P bodies and stress granules, can also influence the concentration and segregation of RNA processing factors, with functional impacts on translation or decay of transcripts localized within them (Anderson and Kedersha, 2006).

In this Chapter, we describe efforts to elucidate mRNA stability and decay mechanisms mediated by YTHDF proteins when considering transcript-wise and cellular location as important contextual aspects. Through re-analysis of sequencing datasets, we first examine the impact of m^6A regional localization along a transcript on mRNA stability, finding that 5'UTR m^6A correlates with increased transcript stability across multiple cell lines whereas CDS and 3'UTR m^6A correlate with decreased stability. We then show that YTHDF2 depletion leads to significant destabilization of transcripts with 5'UTR m^6A and explore potential candidate partner proteins that could mediate this effect. Lastly, we reinvestigate the mechanism underlying mRNA stabilization of YTHDF1-3 triple knockdown with a focus on cytosolic contexture, finding that large-scale changes in P body

formation under these conditions may account for global stabilization of mRNA rather than only functional redundancies among YTHDF proteins. Overall, our studies highlight the importance of location on the transcript and in the cell when contemplating models of m^{6}A function in stability and decay.

RESULTS

Investigating the relationship between YTHDF proteins and regional m^{6}A location in regulating transcript stability

Inspired by the recent finding that m^{6}A in the 5'UTR correlates with increased transcript stability using m^{6}A-SAC-seq data (Hu et al., 2022), with differing regulation when in CDS and 3'UTR regions, we examined m^{6}A localization and transcript stability using two independent methylated RNA immunoprecipitation and sequencing (MeRIPseq) datasets. m^{6}A peaks from HeLa (GSE46705) and mESCs (GSE52681) were grouped by transcript region—5'UTR, CDS, and 3'UTR—and then data were filtered for transcripts with m^{6}A only in a single region. We paired these data with transcript lifetimes from identical cell-lines (HeLa: GSE49339; mESC: GSE148039) to compare half-lives of these transcript groupings, with reference to unmethylated transcripts (i.e., for which there were no m^{6}A peaks). This comparison revealed that, although transcripts with only CDS or 3'UTR m^{6}A (also referenced as "CDS-m^{6}A-only" and "3'UTR-m^{6}A-only") have significantly shorter half-lives than unmethylated transcripts, those with only 5'UTR m^{6}A (also referenced as "5'UTR-m^{6}A-only") have longer half-lives on average in HeLa (**Figure 3.1A**) and mESCs (**Figure 3.1B**). These data validated the previous finding by Hu *et al*. (Hu et al., 2022) with different experimental protocols, although with varying degrees of stabilization.

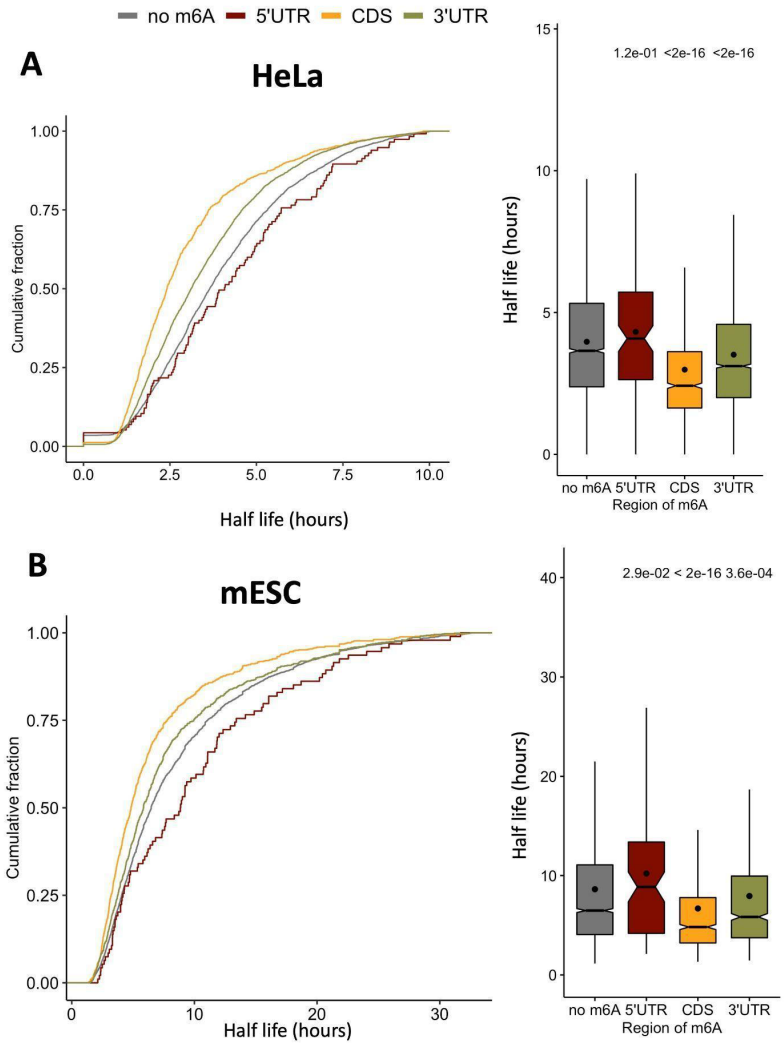


Figure 3.1. m^{6}A in the 5'UTR correlates with increased stability in WT cells. Cumulative fraction plots (left) and boxplots (right) displaying the half lives of transcripts with m^{6}A only in a single region, grouped by region, using public datasets from HeLa (**A**) and mESCs (**B**). The numbers of transcripts represented are: 3782 (no m^{6}A), 115 (5'UTR), 1107 (CDS), and 3494 (3'UTR) for HeLa and 3150 (no m^{6}A), 94 (5'UTR), 780 (CDS), and 827 (3'UTR) for mESCs. Lifetime data are from GSE49339 (HeLa) and GSE148039 (mESC). MeRIPseq data are from GSE46705 (HeLa) and GSE52681 (mESC). Negative lifetimes and outliers above 1.5x the interquartile range above the upper quartile were removed. For boxplots, the center line represents the median, the dot represents the mean, the box limits show the upper and lower quartiles, and the whiskers represent the 1.5x interquartile range. *P* values were determined by a Mann-Whitney-Wilcoxon test.

To gain insight into the mechanism underpinning this relationship, we investigated whether the effect may be mediated by one or more m^{6}A reader proteins. Thus, in HeLa cells, we knocked down each member of the YTHDF family of proteins—the first and most well-characterized cytosolic m^{6}A readers—and performed RNA sequencing (RNA-seq) using spike-in normalization.[1] In parallel, we re-analyzed public lifetime RNA-seq data from wild-type and single YTHDF knockout mESCs (GSE148039) (Lasman et al., 2020). We again grouped transcripts by those that had m^{6}A only in a single region, and compared the effects of depletion or loss of each YTHDF protein on transcript abundance or stability (**Figure 3.2**). In HeLa, transcripts with only CDS or 3'UTR m^{6}A were generally significantly increased in abundance compared to unmethylated transcripts by depletion of any of the three YTHDFs (**Figure 3.2A-C, left**). In contrast, transcripts with only 5'UTR m^{6}A were decreased in abundance by depletion of YTHDF2 or YTHDF3, and a much weaker effect was observed with YTHDF1 knockdown (**Figure 3.2A-C, left**). In mESCs, transcripts with only CDS m^{6}A were significantly stabilized by loss of YTHDF2 or YTHDF3, whereas 5'UTR-m^{6}A-only transcripts were destabilized under these conditions, albeit not significantly (**Figure 3.2B-C, right**). These results suggest that stabilization of transcripts with 5'UTR m^{6}A sites could be partly mediated by YTHDF2 and/or YTHDF3. Such a role would contrast with their destabilizing effects at other transcript regions, and would differ from their canonical functions.

[1] This specific experiment was performed by Zhongyu Zou, and its results are noted in (Zou et al., 2022).

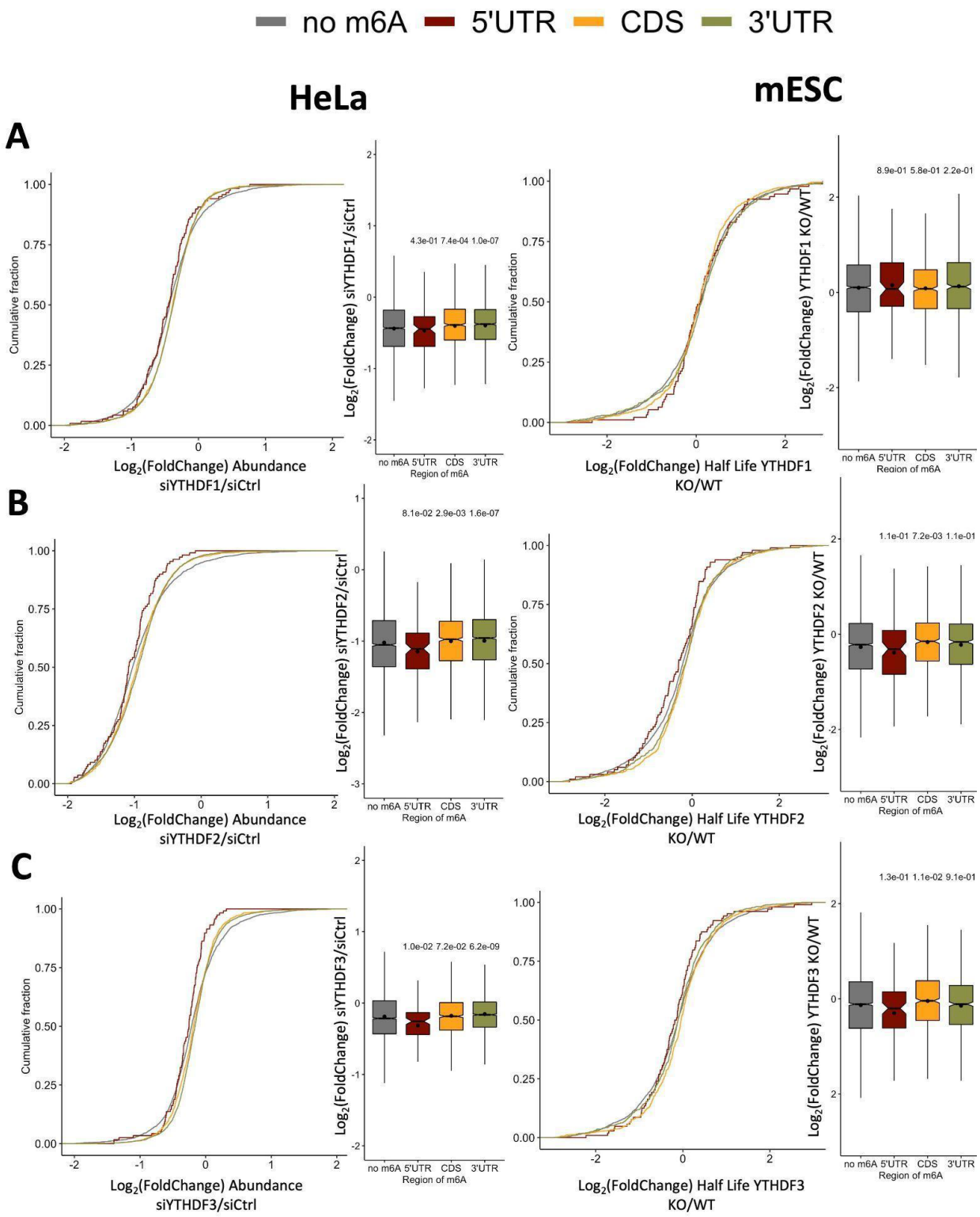


Figure 3.2. The effects of YTHDF loss on transcript stability by m^{6}A region. Cumulative fraction plots (left) and boxplots (right) displaying the $\log_2$(FoldChange) in half lives of transcripts with m^{6}A only in a single region, grouped by region, upon depletion of YTHDF1 (**A**), YTHDF2 (**B**), or YTHDF3 (**C**) in HeLa (left) and mESCs (right). The numbers of transcripts represented are: 6130 (no m^{6}A), 131 (5'UTR), 1619 (CDS), and 3592 (3'UTR) for HeLa and 7351 (no m^{6}A),

Figure 3.2, continued. 174 (5'UTR), 1214 (CDS), and 1540 (3'UTR) for mESCs. mESC lifetime data are from GSE148039, whereas abundance data for HeLa was newly generated from our lab. MeRIPseq data are from GSE46705 (HeLa) and GSE52681 (mESC). For boxplots, the center line represents the median, the dot represents the mean, the box limits show the upper and lower quartiles, and the whiskers represent the 1.5x interquartile range. *P* values were determined by a Mann-Whitney-Wilcoxon test.

Cellular regulation of transcript stability is intertwined with translation. It has been reported that transcripts with m^{6}A in the 5'UTR exhibit a significant reduction in translation efficiency (TE) upon METTL3 depletion, supporting a role for 5'UTR m^{6}A in translation promotion (Meyer et al., 2015). To gain insight into the potential relationship between YTHDF function in translation and in mRNA stability, as mediated by transcript region, we re-analyzed public ribosome profiling data, again grouping transcripts by those with m^{6}A only in a single region (GSE134380) (Zaccara and Jaffrey, 2020). We found that TE was significantly reduced for 3'UTR-m^{6}A-only transcripts by individual depletion of all three YTHDF proteins , but there was no significant difference in TE for transcripts with only 5'UTR m^{6}A (**Figure 3.3A-C**).

Combined with the results from **Figure 3.2**, these data suggest that YTHDF2 and/or YTHDF3 may stabilize transcripts with only 5'UTR m^{6}A via a mechanism independent of translation promotion. A limiting factor of these analyses is that they included all transcripts with m^{6}A only in a single region rather than confident targets of each depleted protein; however, intersection of existing CLIP-Seq datasets did not provide sufficient downstream 5'UTR-m^{6}A-only sample sizes for confident statistical comparisons (data not shown).

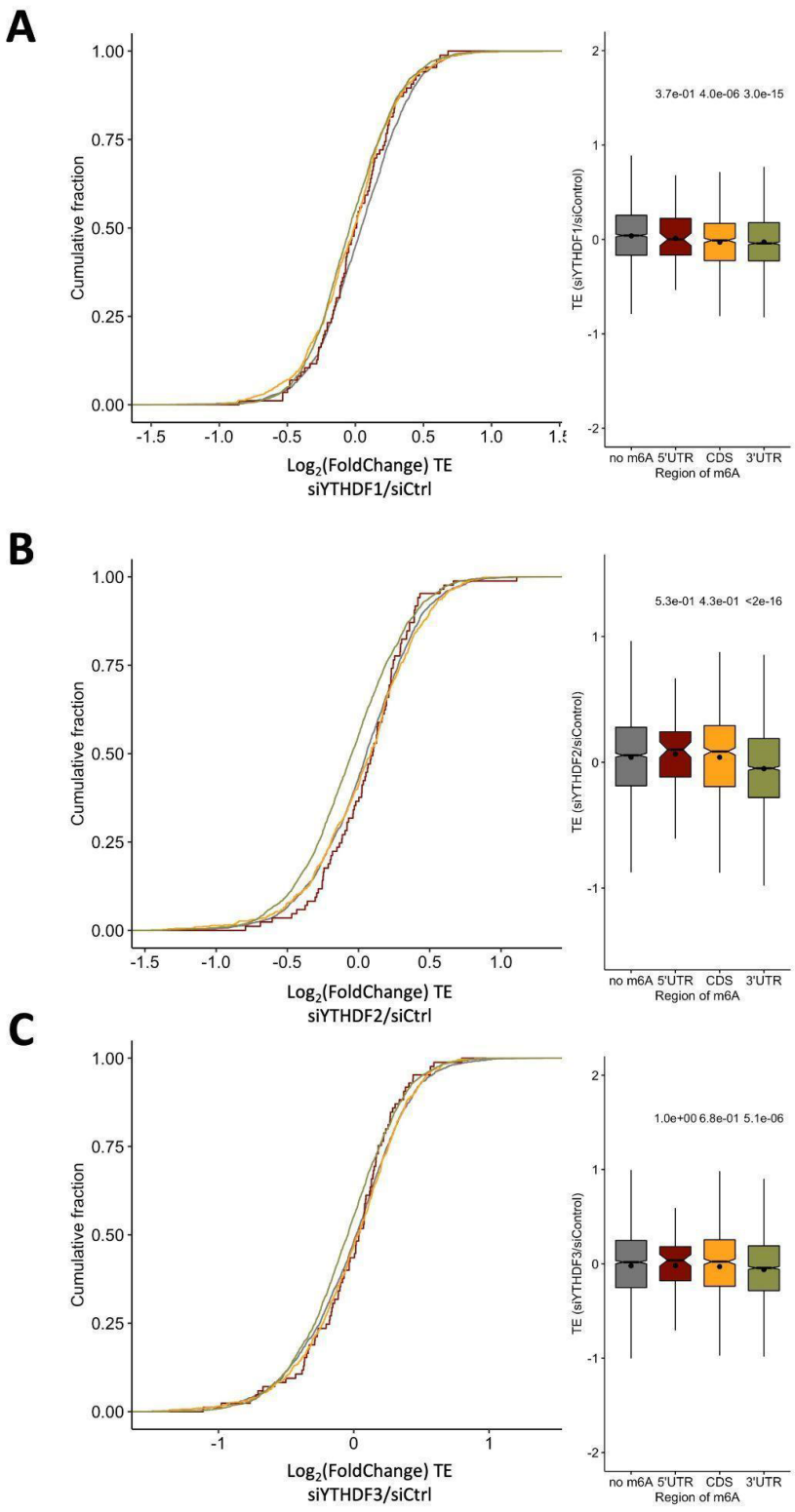


Figure 3.3. The effects of YTHDF loss on translation efficiency by m^{6}A region in HeLa cells. Cumulative fraction plots (left) and boxplots (right) displaying the $\log_2$(FoldChange) in translation efficiency (TE) of transcripts with m^{6}A only in a single region, grouped by region, upon depletion of YTHDF1 (**A**), YTHDF2 (**B**), or YTHDF3 (**C**) in HeLa. The numbers of transcripts represented

Figure 3.3, continued. are: 2210 (no m^{6}A), 87 (5'UTR), 721 (CDS), and 2298 (3'UTR). TE data are from GSE134380 and MeRIPseq data are from GSE46705. For boxplots, the center line represents the median, the dot represents the mean, the box limits show the upper and lower quartiles, and the whiskers represent the 1.5x interquartile range. *P* values were determined by a Mann-Whitney-Wilcoxon test.

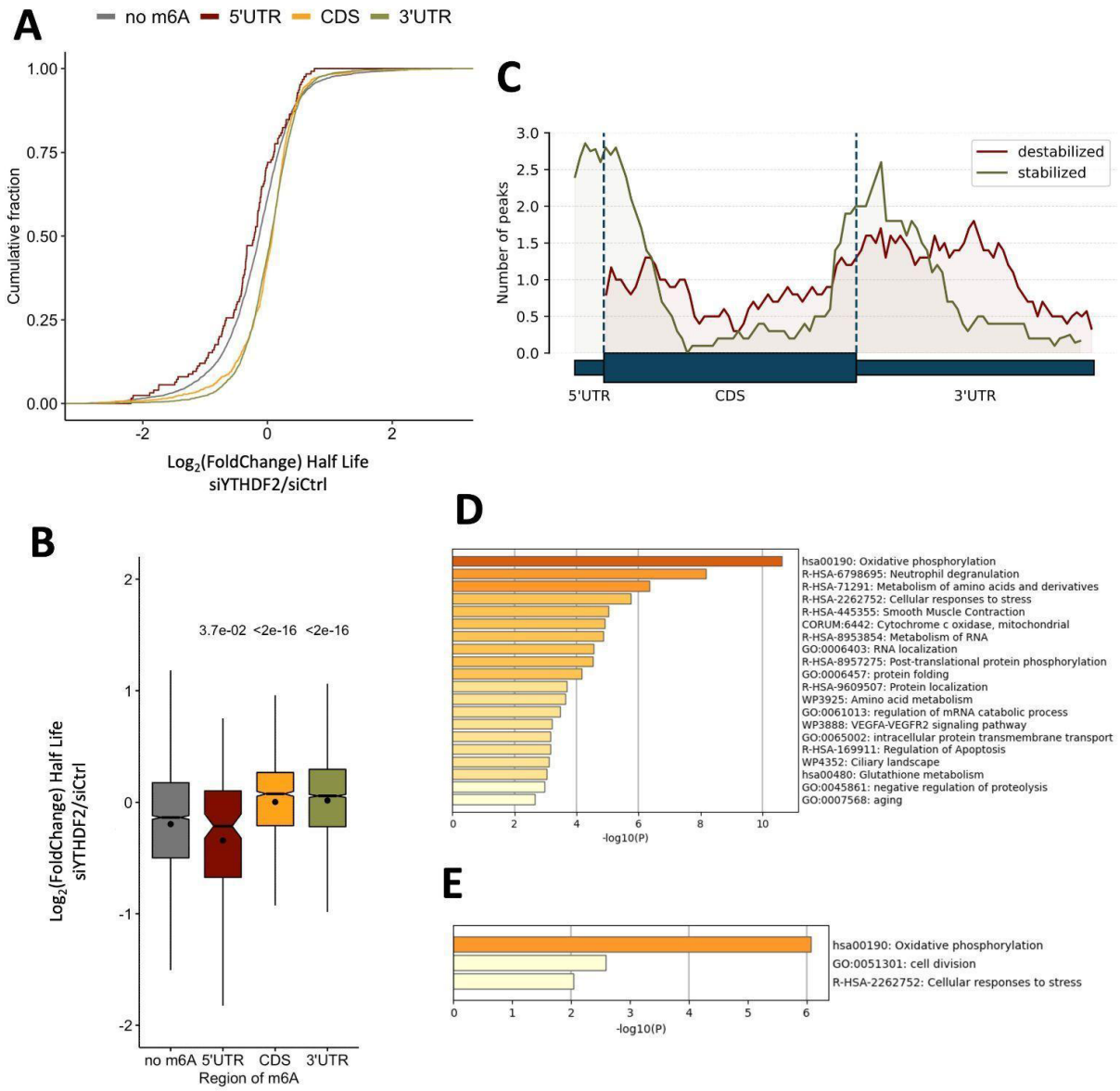


Figure 3.4. The effects of YTHDF2 depletion mRNA lifetime by m^{6}A region. Cumulative fraction plot (**A**) and boxplots (**B**) displaying the $\log_2$(FoldChange) in half life of transcripts with m^{6}A only in a single region, grouped by region, upon depletion of YTHDF2 in HeLa. The numbers of transcripts represented are 4136 (no m^{6}A), 130 (5'UTR), 1173 (CDS), and 3618 (3'UTR). Lifetime data are from GSE49339 and MeRIPseq data are from GSE46705. For boxplots, the

Figure 3.4, continued. center line represents the median, the dot represents the mean, the box limits show the upper and lower quartiles, and the whiskers represent the 1.5x interquartile range. *P* values were determined by a Mann-Whitney-Wilcoxon test. (**C**) A metagene plot showing the regional locations of the top 100 transcripts with m^{6}A only in a single region most stabilized and most destabilized by the presence of YTHDF2. Transcripts are assigned to be stabilized by YTHDF2 if they have a negative $\log_2$(FoldChange) in half life upon YTHDF2 depletion, and destabilized by YTHDF2 if they have a positive $\log_2$(FoldChange) in half life upon YTHDF2 depletion. (**D**) MetaScape gene annotations for the top 100 transcripts most stabilized by YTHDF2 in HeLa. (**E**) MetaScape gene annotations for transcripts with m^{6}A only in the 5'UTR in HeLa that are stabilized by YTHDF2.

To further investigate the stabilization mechanism in HeLa cells, we analyzed another dataset containing lifetime data from YTHDF2 knockdown and control cells (GSE49339) (Wang et al., 2014a). Of note, HeLa data presented in **Figures 3.1-3.3** involved either lifetimes from WT cells or abundances from cells with YTHDF knockdown, but not lifetimes in cells experiencing YTHDF knockdown[2]. As noted by the original authors, lifetime data analysis upon YTHDF2 knockdown of all YTHDF2 targets revealed its overall role in promoting transcript degradation (Wang et al., 2014a). In our analysis, grouping transcripts with m^{6}A only present in single regions revealed that transcripts with only CDS or 3'UTR m^{6}A have significantly increased lifetimes upon YTHDF2 depletion (**Figure 3.4A-B**). Conversely, the lifetimes of transcripts with only 5'UTR m^{6}A were significantly decreased upon YTHDF2 knockdown (**Figure 3.4A-B**). Although our analysis is not only limited to YTHDF2 targets, the results suggest that YTHDF2 likely promotes the stability of transcripts with 5'UTR m^{6}A directly or indirectly.

To further elucidate whether there exists a regional preference for YTHDF2 function in promoting stability or decay, we examined metagene profiles for the top 100 transcripts with m^{6}A

[2] Lifetime data acquisition requires treating cells with actinomycin D to inhibit transcription, followed by multi-time point sample collection and RNA-Seq with spike-in normalization. This permits analysis of transcript decay rates independent of transcription, as done in mESCs in **Figures 3.1-3.2.**

only present in a single region that were stabilized by YTHDF2 (i.e., a negative fold change upon YTHDF2 knockdown) or destabilized by YTHDF2 (i.e., a positive fold change upon YTHDF2 knockdown) (**Figure 3.4C**). Although zero of the transcripts destabilized by YTHDF2 had only 5'UTR m^6A, there was a peak among transcripts stabilized by YTHDF2 in the 5'UTR and early CDS (**Figure 3.4C**). There were also peaks for both sets of transcripts near the end of the CDS and beginning of the 3'UTR, which is a common distribution pattern for mRNA m^6A (Meyer et al., 2012). These results suggest that YTHDF2's function—whether in promoting stabilization or decay— is related to the position of methylation on its target transcript. Gene ontology (GO) analysis of the top 100 transcripts most stabilized by YTHDF2 in HeLa (**Figure 3.4D**; same as green line in **Figure 3.4C**), as well as those transcripts with only 5'UTR m^6A that are stabilized by YTHDF2 (**Figure 3.4E**), revealed enrichment of pathways important during stress: general cellular responses to stress, neutrophil degranulation (activated during infection), VEGF signaling, and oxidative phosphorylation pathways (modulated during hypoxia) (Fuhrmann and Brüne, 2017; Rosa et al., 2021).

To form a hypothesis for the mechanism by which YTHDF2 promotes stability, we profiled data from a large-scale, systematic, *in vivo* proximity-dependent biotinylation (BioID) study (Roux et al., 2018; Youn et al., 2018). Specifically, the authors included YTHDF1, YTHDF2, and YTHDF3 as "bait" proteins with a promiscuous biotin ligase (birA*) fused to the N-terminus or C-terminus (Roux et al., 2018; Youn et al., 2018)—however we note that for YTHDF1 only data with C-terminal birA* was made available. Upon addition of exogenous biotin, biotinylation of proteins within ~10 nm of the bait protein was induced, and biotinylated "prey" proteins were isolated with standard biotin affinity-capture and identified using mass spectrometry (Youn et al., 2018).

To visualize YTHDF2's top interactors, we plotted the log-fold change of counts of prey proteins in the YTHDF2-N-birA* bait purification versus the negative $\log_{10}$ of the Bayesian false discovery rate (BFDR) (**Figure 3.5A**). Although this visually depicted interaction enrichment, the data provided by the authors rounded the BFDR to two decimal points, causing any bait-prey relationship with a BFDR <0.005 to be assigned a BFDR of 0 (Teo et al., 2014; Youn et al., 2018). Thus, we also plotted the log-fold change versus the $\log_{10}$ of the sum of spectral counts of the prey protein in the YTHDF2-N-birA* bait purification (**Figure 3.5B**). In both plots, proteins with an average probability of interacting with YTHDF2 of greater than 0.95 are shown in red, notably including multiple members of the CCR4-NOT complex, which is known to interact with YTHDF2 to promote transcript decay (**Figure 3.5B**). We also identified several candidates that bind the 5' cap or interact with cap-binding factors: eIF4E, eIF4E2, EIF4ENIF1 (Dostie et al., 2000; Meyer et al., 2015; Teo et al., 2014; Uniacke et al., 2012), which are involved in translation regulation (EIF4G2, EIF4G3) (Sun et al., 2010; Yang et al., 2017), or which have known or presumed roles in binding RNA to regulate stability (YBX3, IGF2BP2, FMR1, FXR1, FXR2) (Cooke et al., 2019; Edupuganti et al., 2017; Huang et al., 2018; Kirkpatrick et al., 2001) (**Figure 3.5B**). These hits provide at least three possible pathways through which a YTHDF2-interacting protein could promote stability of YTHDF2 targets.

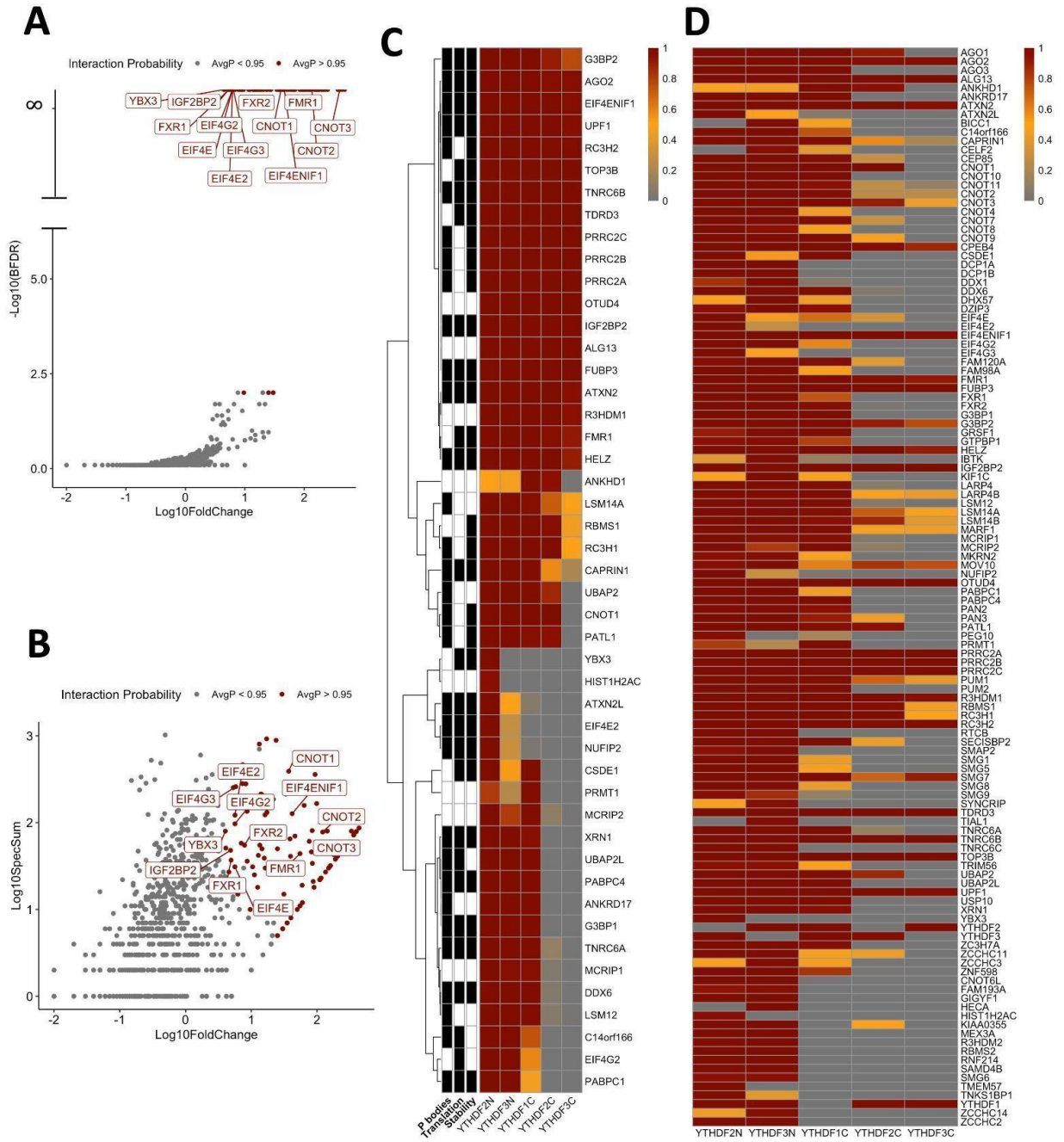


Figure 3.5. YTHDF interaction partners from public proximity-dependent biotin identification data. (**A**) Volcano plot of the Log_{10}(FoldChange) of counts in the YTHDF2-bait purification compared to the control versus the Log_{10} of the Bayesian false discovery rate. Those proteins with an interaction probability with YTHDF2 of > 0.95 are colored in red. (**B**) Scatter plot of the Log_{10}(FoldChange) of counts in the YTHDF2-bait purification compared to the control versus the Log_{10} of the sum of spectral counts for the prey protein. (**C**) Heatmap of the average interaction probabilities of the top YTHDF interactors as identified by Zaccara *et al*. (Zaccara and Jaffrey, 2020) from each of the samples with a YTHDF protein as bait in the proximity-dependent biotin identification (BioID) dataset. (**D**) Heatmap of the average interaction probabilities of all prey proteins that had a BFDR of 0 in any of the samples with a YTHDF protein as bait. The BioID

Figure 3.5, continued. dataset is from Youn *et al*., and notably does not include data for N-terminally-labeled YTHDF1 (Youn et al., 2018).

To pair our mechanistic exploration with effects of YTHDF1, YTHDF2, and YTHDF3 knockdown (cf. **Figure 3.2**, **Figure 3.3**), we created a heatmap of the top YTHDF1-3 interactors, as identified by Zaccara *et al*., and annotated these genes for known involvement in translation, stability, and P body association (**Figure 3.5C**) (Hubstenberger et al., 2017; Safran et al., 2002; Zaccara and Jaffrey, 2020). Since our results in **Figure 3.2** suggested greater effects of YTHDF2 and YTHDF3 than YTHDF1 to stabilize 5'UTR methylated transcripts, we searched for candidate partner proteins showing greater interaction with YTHDF2 and/or YTHDF3 than YTHDF1 (i.e., comparing columns 1-3 of the heatmap in **Figure 3.5C**). Candidates fitting this description included YBX3, EIF4E2, and PABPC1—all of which have previously reported roles in transcript stability (Behm-Ansmant et al., 2007; Cooke et al., 2019; Haimovich et al., 2013; Melanson et al., 2017; Weber et al., 2020). C14orf166 and EIF4E2 could also promote mRNA stability through their interactions with the 5' cap or cap-binding proteins (Melanson et al., 2017; Pazo et al., 2019), and EIF4G2 could affect stability through its translation functions (de la Parra et al., 2018). These interactions with YTHDF1-3 suggest several hypotheses through which 5'UTR-m^{6}A-containing transcripts may be stabilized.

This analysis also revealed that the YTHDF proteins share numerous top interaction partners, many of which are associated with P-bodies and RNA stability (**Figure 3.6C**). However, there are potentially several interaction partners that are not shared by all three YTHDFs, an observation that becomes particularly evident when comparing average interaction probabilities between each YTHDF protein and all other proteins that had a BFDR of ~0 in any one sample (**Figure 3.6D**). Notably, C-terminally tagged YTHDFs had fewer significant interactions,

potentially since the C-termini contain the YTH domains that bind m^{6}A; conversely, the more disordered N-termini are conventionally the primary effector domains. Although the lack of an N-terminally tagged YTHDF1 purification sample (by the original study) limits the scope of these conclusions, it is apparent that the three YTHDFs do not share the exact same set of interaction partners, which contrasts with a recently proposed model of their redundant functions in promoting mRNA decay due to a putatively shared interaction network (Zaccara and Jaffrey, 2020).

P body dysregulation mediates mRNA stabilization upon triple knockdown of YTHDFs

Our group and others have reported that triple knockdown of YTHDF1-3 proteins leads to greater mRNA stabilization than knockdown of any one or two YTHDF protein(s) (Lasman et al., 2020; Shi et al., 2017a; Zaccara and Jaffrey, 2020). This finding inspired a model wherein YTHDF1-3 proteins have redundant functions in promoting mRNA decay (Zaccara and Jaffrey, 2020). However, this model stands in contrast to previous findings that YTHDF1 and YTHDF3 play important roles in translation promotion (Shi et al., 2017a; Wang et al., 2015), as well as the results of multiple reporter assays, wherein tethering or targeting of YTHDF1 causes increased reporter translation with minimal impact on stability (Rauch et al., 2018; Wang et al., 2015).

Moreover, analyses by our group using data from Zaccara *et al*. revealed that the predominant role of YTHDF1 is translation promotion, but not decay, and this is evident when transcripts are grouped by YTHDF targets to assess the effects of YTHDF knockdown (Zaccara and Jaffrey, 2020; Zou et al., 2022). Furthermore, sequence alignments of the N-terminal low complexity domains (LCDs) displayed significant differences among YTHDF proteins, especially for YTHDF2, and the YTHDF2 LCD also was found to form distinct fibril-like structures in condensate formation, unique from YTHDF1 and YTHDF3 (Zou et al., 2022). Finally, analyses

of multiple datasets, including the BioID dataset presented in the previous subsection, revealed key differences in YTHDF interactomes (Go et al., 2021; Youn et al., 2018; Zou et al., 2022). Thus, YTHDF proteins likely have different functions, at least in certain cellular contexts, raising questions whether the significant stabilization effect of YTHDF triple knockdown is appropriately and solely modeled by putatively redundant YTHDF1-3 decay functions.

To continue exploring the potential mechanisms behind the stabilization effect of YTHDF triple knockdown, we also re-analyzed the ribosome profiling data published by Zaccara *et al*. (Zaccara and Jaffrey, 2020). Unexpectedly, we found that the authors used sequencing data from the inputs of ribosome profiling experiments and analyzed the effects of YTHDFs on the mRNA abundances of only actively translated genes (i.e., pre-filtered transcripts with >0 read counts in the ribosome-protected fragment samples). This limited their analysis to only 6,814 transcripts. In contrast, we re-analyzed the sequencing data to study all detected transcripts (16,595 transcripts with a sum of >10 read counts across all samples). In our re-analysis, we found that abundances increased significantly for both translated and expressed transcripts upon triple YTHDF knockdown, and for translated transcripts upon YTHDF2 or, to a lesser but still significant extent, YTHDF3 knockdown (**Figure 3.6A**). In these situations, more m^6A sites per transcript generally correlated with increased abundances. In contrast, we found that upon YTHDF1 or YTHDF3 knockdown, the abundance of expressed transcripts tended to decrease with an increase in m^6A sites, without a significant correlation between m^6A site numbers and RNA abundance changes for translated genes (**Figure 3.6A**). These differences highlight how cellular translation regulation impacts RNA abundance, thus, interrogating YTHDF protein function using only actively translated transcripts and leaving out untranslated transcripts is inappropriate. Moreover, there are additional changes in cellular processes introduced by treatment with a translation inhibitor which

impact the data and interpretation of results. Overall, these considerations highlight how choices in the experimental setup and analysis could have negatively impact the results on which the model for redundant YTHDF function in decay was built. At the same time, we validated the key result that triple YTHDF knockdown is associated with increased transcript stabilization, particularly with more highly methylated transcripts.

Interestingly, in our re-analysis of the RNA-seq data from single and knockdown conditions, we noticed a broad dysregulation of the cellular transcriptome upon triple YTHDF knockdown (**Figure 3.6B**). Although single knockdown of YTHDF1, YTHDF2, and YTHDF3 resulted in moderate transcript abundance changes, with the gene targeted by the siRNA in each case being one of the most significantly decreased, triple knockdown of all three YTHDFs resulted in much more significant alteration of the transcriptome (i.e., 5-19-fold more transcripts with adjusted P values ≤ 0.05 for differential expression in triple knockdown compared to single knockdown conditions). Although it is possible that redundant functions of YTHDF1-3 could lead to this result, caused by compensation upon singular gene loss that is abolished upon triple loss, the targets of the YTHDF proteins do not entirely overlap, especially in certain cell lines (Liu et al., 2018b). Additionally, we and others have found that YTHDF1-3 triple knockdown causes stabilization of unmethylated as well as methylated transcripts (Lasman et al., 2020; Zou et al., 2022). Combining these findings with the much larger scale of transcriptome alteration during triple knockdown compared to that during depletion of any single YTHDF (or their sum) suggests instead the perturbation of a fundamental cytosolic process regulating RNA stability.

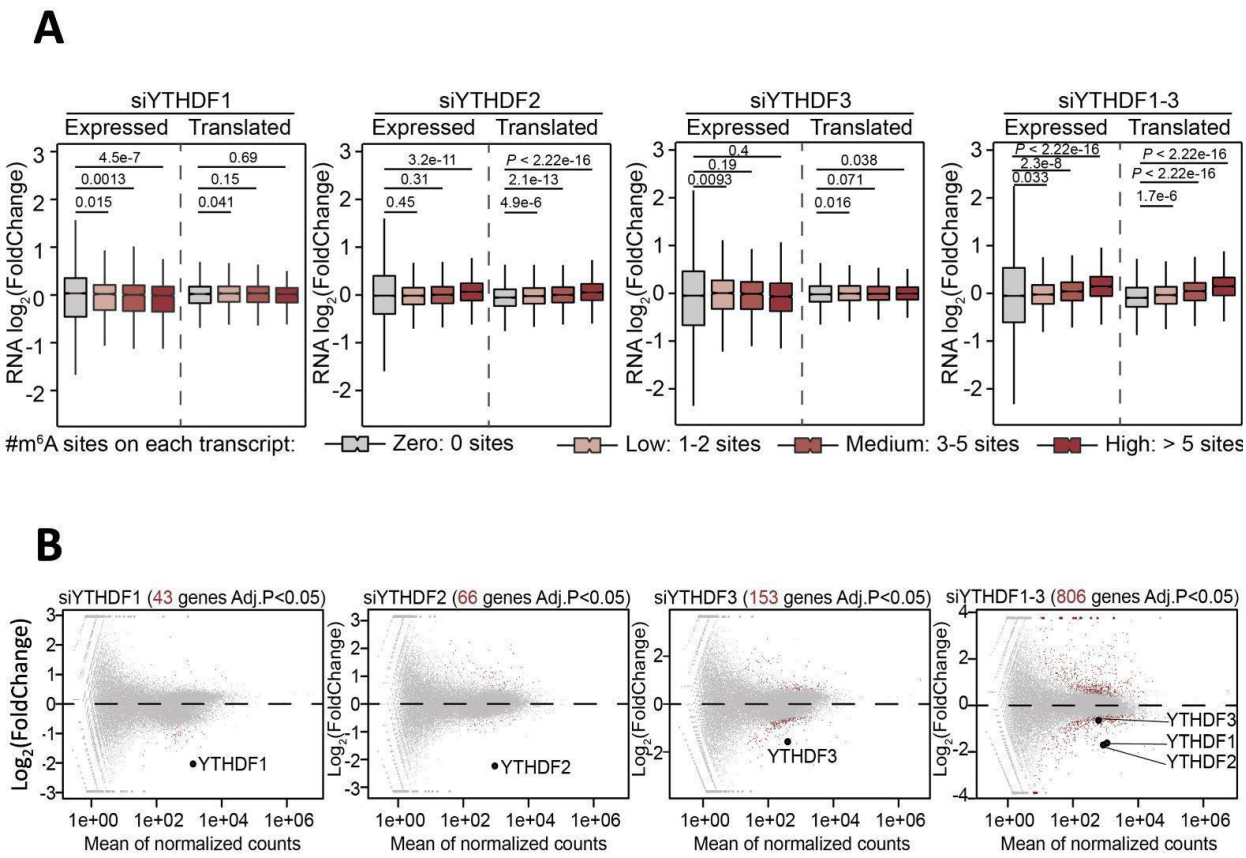


Figure 3.6. Triple knockdown of YTHDF proteins causes transcriptomic dysregulation. (**A**) Boxplots showing the $\log_2$(FoldChange) in abundance of transcripts with different numbers of m^{6}A sites when analyzing all expressed genes ("Expressed") or only actively translated genes ("Translated") after single or triple knockdown of YTHDF proteins in HeLa. Transcripts are categorized by the number of m^{6}A sites (Zero: 0 sites, Low: 1-2 sites, Medium: 3-5 sites, High: >5 sites). For boxplots, the center line represents the median, the dot represents the mean, the box limits show the upper and lower quartiles, and the whiskers represent the 1.5x interquartile range. *P* values were determined by a Mann-Whitney-Wilcoxon test. (**B**) MA plots showing the mean of normalized counts versus the $\log_2$(FoldChange) after single or triple knockdown of YTHDF proteins. Red dots denote significantly differentially expressed genes, and genes targeted by siRNAs are labeled. All data was retrieved from Zaccara *et al. (Zaccara and Jaffrey, 2020)* (GSE134380).

YTHDF2 is known to localize to P bodies (Wang et al., 2014a), and many top YTHDF interacting partners are associated with P bodies (**Figure 3.5C**). Since cytosolic P bodies are hubs for RNA processing, with reported roles in transcript stability (Hubstenberger et al., 2017; Luo et al., 2018), we investigated changes in P bodies upon YTHDF knockdown in cultured HeLa cells

by staining with the P body marker DCP1A (**Figure 3.7A**). We found that depletion of all three YTHDF proteins caused significantly increased numbers of P bodies (**Figure 3.7A-B**). Thus, to investigate how methylated and unmethylated transcripts are impacted by P body localization, we performed RNA immunoprecipitation and sequencing (RIP-seq) with antibodies to two individual P body markers, EDC3 and EDC4 (Zou et al., 2022). We then calculated the average $\log_2$-fold enrichment (versus the input) across the two datasets to define the P body transcriptome in HeLa cells and cross-referenced it with transcripts known to have m^{6}A (Zou et al., 2022).

Strong m^{6}A signals have been associated with P body imaging data (Fu and Zhuang, 2020). In line with this, we found that m^{6}A-modified transcripts are significantly enriched in P body-immunoprecipitate (both EDC3 and EDC4 datasets), and that more highly methylated transcripts are more likely to be enriched in P bodies (**Figure 3.7C**). Coming full circle, we analyzed abundance changes upon triple YTHDF knockdown[3] in HeLa from RNA-seq data with spike-in controls, this time categorizing transcripts by both methylation status and P body enrichment status from wild type (WT) cells (**Figure 3.7D**). This analysis revealed a global stabilization of transcripts upon triple knockdown. Those transcripts that are P body enriched in WT cells were most stabilized upon triple knockdown, with minor, albeit still significant, differences for methylated versus unmethylated transcripts (**Figure 3.7D**). An explanation for these findings could be that increased P bodies upon triple YTHDF knockdown leads to increased P body residence of most cellular transcripts, particularly favoring those traditionally enriched in P bodies. For transcripts that are P body depleted in WT cells, there is a significant difference in stabilization between methylated and unmethylated transcripts. This could be due to the loss of the decay-promoting function of one or more YTHDF proteins, particularly YTHDF2. It could also be related

[3] This specific experiment was performed by Zhongyu Zou, and its results are noted in (Zou et al., 2022).

to a preference for P body enrichment of methylated transcripts in the additional P bodies formed under these conditions, including methylated transcripts that were depleted under wild type conditions. The mechanistic relationship between P body changes and mRNA stabilization upon YTHDF triple knockdown is further elucidated in our report on BioRxiv (Zou et al., 2022).

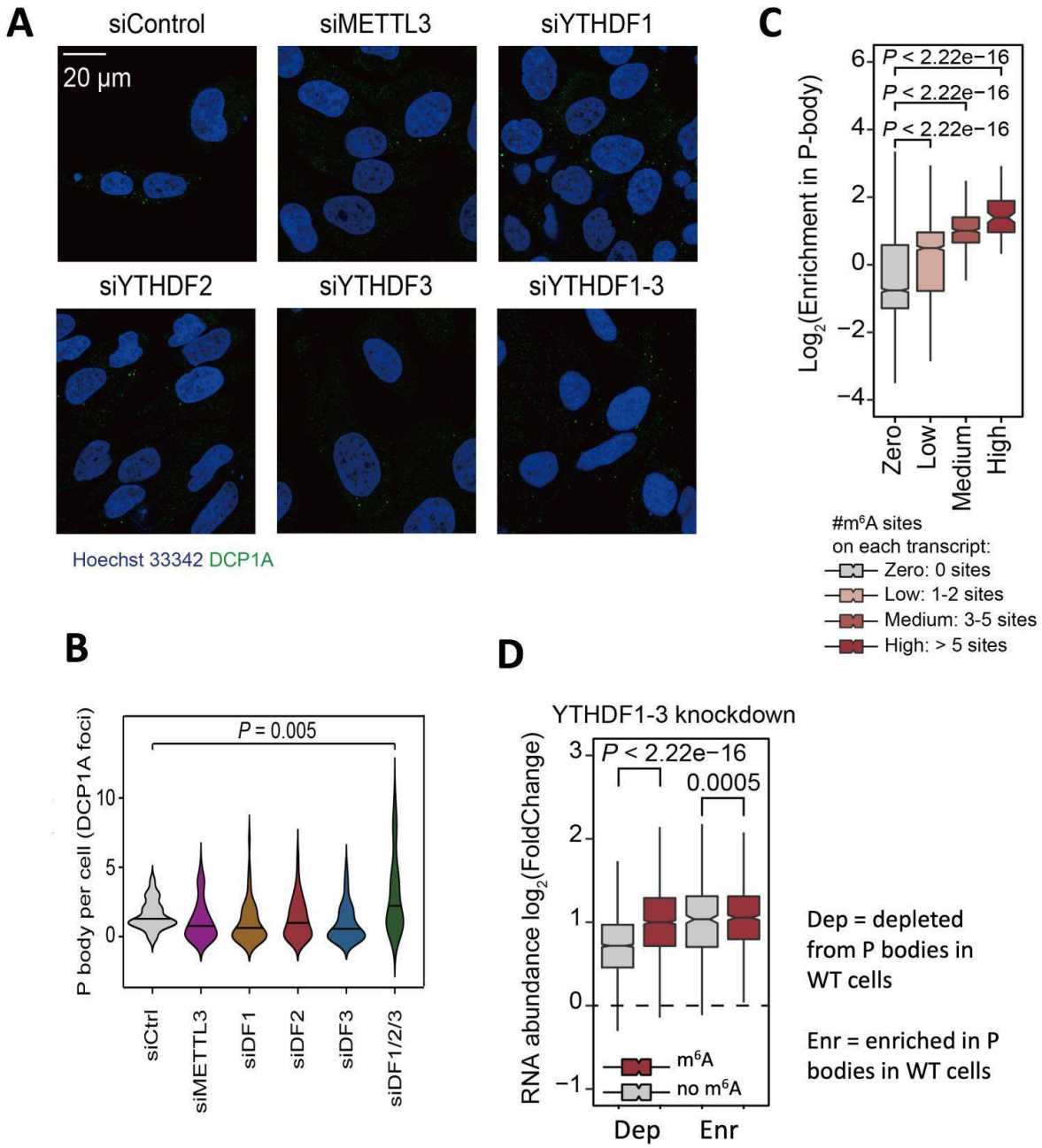


Figure 3.7. Increased P body formation after YTHDF triple knockdown. A, Representative images from P body imaging in HeLa cells. P-bodies were stained using a DCP1A antibody and cell nuclei were counterstained by Hoechst 33342. Eight images were captured for each condition

Figure 3.7, continued. in n = 3 individual experiments. **B,** Fluorescence microscopy analysis of P body numbers after single knockdown or triple knockdown of YTHDFs in HeLa cells. Numbers of DCP1A foci per cell were quantified with CellProfiler 3.0. **C,** Boxplots showing enrichments of different transcripts grouped by numbers of m^6A sites identified by MeRIP-seq (zero: 0 sites, low: 1-2 sites, medium: 3-5 sites, high: more than 5 sites). **D,** Boxplots showing the log_2(FoldChange) in abundance of transcripts upon triple knockdown of YTHDF proteins, categorized by P body enrichment and m^6A methylation status in wild-type HeLa (grey = unmethylated, red = methylated, "Dep" = P body depleted in WT cells, "Enr" = P body enriched in WT cells). For boxplots, the center line represents the median, the box limits show the upper and lower quartiles, and whiskers represent the 1.5 × interquartile range. *P* values were determined by a Mann-Whitney-Wilcoxon test. P body enrichment data and RNA-seq data were generated by our lab and are reported in Zou *et al*. (Zou et al., 2022).

It is likely that loss of stabilization by one or more YTHDF proteins contributes to stabilization of methylated transcripts upon triple knockdown; however, since unmethylated transcripts are also stabilized in this context, loss of YTHDF decay function paints an incomplete picture. Our results suggest that the dysregulation of P bodies upon triple YTHDF knockdown likely underlies the observed global transcriptome stabilization effect. The data further imply that the composition of P bodies with respect to transcript methylation status may also be altered under triple knockdown conditions, a question warranting future investigation.

DISCUSSION

In this study, we first called attention to the importance of the regional location of m^6A in determining its effect on transcript stability. We showed that transcripts with only 5'UTR m^6A have increased stability compared to those with only CDS or 3'UTR m^6A, as well as unmethylated transcripts, in multiple cell lines (**Figure 3.1**). We then found that depletion of YTHDF2 or YTHDF3 is associated with destabilization of these 5'UTR-m^6A-only transcripts (**Figure 3.2**, **Figure 3.4**), without an apparent significant effect on their translation efficiency (**Figure 3.3**). Importantly, since these analyses incorporated all expressed transcripts to ensure sufficient 5'UTR

m^{6}A transcript sample sizes for robust statistical comparisons, and not just targets of each YTHDF, indirect effects may also contribute to the observations. Further exploration of the effect of YTHDF2 knockdown on transcript lifetime in HeLa cells revealed an enrichment of 5'UTR m^{6}A among the top 100 transcripts with m^{6}A only in a single region that were stabilized by YTHDF2 (**Figure 3.4C**). These results suggest that YTHDF2 may function in stabilizing mRNA in certain contexts, and that this effect is more apparent when m^{6}A is in the 5'UTR.

There are two major mechanistic models that may explain a potential stabilizing effect of YTHDF2 on 5'UTR-m^{6}A-only transcripts. The first model involves direct stabilization of 5'UTR-m^{6}A-only transcripts by recruitment of a stabilizing protein via YTHDF2 whose mechanism of stabilization is 5'UTR-specific (e.g., by binding the 5' cap). The second model allows for YTHDF2 binding to a stabilization-promoting mediator whose mechanism is not necessarily 5'UTR-specific, but which may dominate in controlling transcript stability if the known mechanisms of YTHDF2-mediated decay are ineffective. For example, the mechanism by which YTHDF2 leads to endonucleolytic cleavage requires HRSP12 to bind approximately 800 nucleotides upstream of the YTHDF2 binding site, whereas the median 5'UTR length in humans is only 218 nucleotides, meaning that the necessary interactions may not occur when the m^{6}A is located within a shorter 5'UTR (Lee et al., 2020; Leppek et al., 2018; Park et al., 2019). The linear distance along the transcript between an m^{6}A site in the 5'UTR and the polyA tail may also render recruitment of the CCR4-NOT deadenylase complex to the latter by YTHDF2 impossible or ineffective in promoting deadenylation, unless substantial transcript looping occurs. As an alternative, the stabilizing mediator in the second model could be replaced with other stabilizing factors such as the relocalization of targets to P bodies, which have been documented as sites of mRNA storage (Hubstenberger et al., 2017; Zou et al., 2022). Still, further investigation of these

models of 5'UTR-m^{6}A-YTHDF2-mediated transcript stabilization likely necessitates investigating YTHDF2 binding partners.

Our analysis of public BioID data (**Figure 3.5**) revealed several candidate partner proteins for investigation. These include proteins with both 5'UTR-dependent and 5'UTR-independent stabilizing effects. For example, eIF4E and eIF4E2 both bind the 5'cap, and EIF4ENIF1 binds to them (Räsch et al., 2020). Interaction of any of these proteins with 5'UTR-bound YTHDF2 could stabilize their residence on the 5' cap, blocking access to decapping factors and preventing further 5'-exonucleolytic decay. YBX3 has also been shown to promote mRNA stability (Cooke et al., 2019), as has the closely related YBX1, which also has 5' cap-binding activity (Cooke et al., 2019; Evdokimova et al., 2001). On the other hand, IGF2BP2 and PABPC1 have been shown to mediate mRNA stabilization through mechanisms involving the 3'UTR (Huang et al., 2018; Su et al., 2020), so any stabilization mediated by them on 5'UTR-YTHDF2-bound transcripts would likely involve a mRNA looping mechanism. Interestingly, IGF2BP2, which was predicted to bind YTHDF2, is also known to bind m^{6}A but has few overlapping mRNA targets with it (Huang et al., 2018). In contrast, FMRP, which also shows a preference for binding methylated transcripts, interacts with YTHDF2 in an RNA-independent manner (Zhang et al., 2018). This study further showed that FMRP could maintain the stability of shared targets for which YTHDF2 could promote degradation (Zhang et al., 2018). However, investigations of the mechanisms by which FMRP promotes stability have not yet revealed any cap- or 5'UTR-dependencies. Further experimentation will be necessary to test each YTHDF2 putative binding partner more thoroughly to confirm/deny these interactions and their association with 5'UTR-m^{6}A-mediated stabilization.

Intriguingly, gene ontology analysis of YTHDF2-stabilized transcripts, especially those with 5'UTR m^{6}A, enriched pathways involved in cellular stress responses (**Figure 3.4D-E**). It has

been reported that m^{6}A increases in the 5'UTR under certain conditions of stress, including exposure to heat shock and UV radiation (Dominissini et al., 2012; Meyer et al., 2015; Zhou et al., 2015). Moreover, in circumstances of heat shock, 5'UTR m^{6}A aided translation of stress response-related transcripts and was associated with increased transcript abundance (Zhou et al., 2015). A similar stabilizing effect of 5'UTR m^{6}A was also observed in hypoxia (Wang et al., 2021b). Interestingly, pathways important for hypoxia response, including oxidative phosphorylation and VEGFA-VEGFR2 signaling, also appear as enriched pathways (**Figures 3.4D-E**).

This relationship between stress responses and 5'UTR-m^{6}A-dependent stabilization may inform mechanistic hypotheses, as some of the candidate YTHDF2 partners have special relationships to stress contexts. For example, eIF4E2 is a cap-binding protein with lower cap-binding affinity than eIF4E that generally stabilizes transcripts and represses translation under normal conditions, but has been shown to promote translation under hypoxic conditions through a mechanism involving HIF-2α (Melanson et al., 2017). The YBX3 paralog, YBX1, has also been described as a potent 5' cap-dependent mRNA stabilizer, particularly under stress conditions involving reduced cap binding by eIF4E (Evdokimova et al., 2001).

Future experiments to evaluate potential YTHDF2 candidate mediators could involve co-immunoprecipitations, as well as knockdown screens to monitor their effects on YTHDF2-stabilized targets. Selection of protein targets to screen could benefit from an assessment of their mRNA targets and binding sites for overlap with YTHDF2 targets and 5'UTR m^{6}A sites in stabilized transcripts. Overall elucidation of this mechanism would also highlight the importance of the regional localization of m^{6}A along a transcript in regulating mRNA stability.

The second section of our study investigated the role of cellular context on observed mRNA stabilization during YTHDF triple knockdown. We first re-examined the RNA sequencing

data from single and triple knockdown conditions that originally informed the model of redundant function among YTHDF1-3 proteins in promoting mRNA stability. We found that, although triple knockdown of all three YTHDFs indeed led to increased mRNA abundance correlated with m^6A number (**Figure 3.6A**), triple knockdown also resulted in significant transcriptomic dysregulation much greater than knockdown of any individual YTHDF or the sum thereof (**Figure 3.6B**), including unmethylated transcripts (Zou et al., 2022). Potentially explaining the latter, we then found significantly increased P body formation upon triple YTHDF knockdown (**Figure 3.7A-B**), which may also explain increased mRNA stabilization under those conditions. Additionally, we characterized P body enriched and P body depleted transcripts in HeLa under normal conditions, and assessed changes in abundance after triple knockdown when grouping transcripts by both P body enrichment status and methylation status in WT cells. Our analysis revealed that all four groups were stabilized upon triple knockdown, with P body enriched transcripts more stabilized than P body depleted transcripts, potentially owing to increased P body formation (**Figure 3.7D**). Although methylated transcripts were still more stabilized than unmethylated transcripts, supporting a role in protein-mediated transcript decay for one or more YTHDF proteins that is diminished upon triple knockdown, our results put forward an additional role for increased P body formation in general stabilization of transcripts under this condition.

The mechanism by which triple YTHDF knockdown causes increased P body number is still unclear. It has been reported that changes in P body numbers upon depletion are a general property of P body proteins (Luo et al., 2018). Although only YTHDF2 has been consistently reported to localize in P bodies (Fu and Zhuang, 2020; Luo et al., 2018; Wang et al., 2014a), all three YTHDFs display significant interaction with many P body-associated proteins, so YTHDF depletion may have a similar effect (**Figure 3.5C**). In particular, it has been shown in yeast that

depletion of Dcp1, Dcp2, or Xrn1 can lead to increased P body numbers due to accumulation of decay intermediates (Luo et al., 2018). Since YTHDF proteins—particularly YTHDF2—are involved in decay, their loss could also result in accumulation of decay intermediates, which may affect P body numbers. Notably, though, the structure of these intermediates would likely differ from those that accumulate upon Dcp1, Dcp2, or Xrn1 knockdown. YTHDF loss could also cause changes in RNA-protein and/or RNA-RNA interactions with cascading effects. For example, removal of YTHDFs from binding methylated RNA could permit binding of other RBPs potentially associated with P bodies. Loss of YTHDFs from their protein interactomes could also change the associations of P body related proteins with each other. Finally, since P body maintenance is dependent on the presence of translationally-repressed mRNA (Luo et al., 2018), it is possible that depletion of YTHDF1 and YTHDF3 may actually contribute to increased P body formation through decreased translation of their targets, resulting indirectly in transcript stabilization through redirecting targets from polysomes to P bodies.

Future investigations of the mechanistic relationships between YTHDF proteins and P bodies are warranted and would increase our understanding of YTHDF control of mRNA stability. In particular, imaging of m^6A and P bodies after triple knockdown could confirm whether m^6A is enriched in the increased P bodies, since so far the spatial association of m^6A in P bodies has been in WT cells. Characterization of the P body transcriptome after triple knockdown, rather than in WT cells, could also identify which transcripts are enriched in P bodies in the presence of YTHDFs versus without them. These experiments could clarify the mechanistic role of P body localization in general processing of methylated transcripts, which remains under-characterized to date.

Overall, the results presented in this chapter increase our understanding on the mechanisms by which YTHDF proteins regulate mRNA stability and decay, and how the molecular and cellular

location of m^{6}A modulates these processes. Our findings also add nuance to our understanding of YTHDF reader function. For example, we find that YTHDF2, which is the YTHDF protein most heavily associated with mRNA decay, can promote mRNA stabilization in certain cellular contexts, and that this mechanism could be related to the position of m^{6}A along the transcript. Furthermore, we find that YTHDF proteins play a role in the regulation of P bodies with profound effects on the global cellular transcriptome. We suggest that this process at least partially underlies the significant mRNA stabilization observed upon YTHDF triple knockdown, since we find that the localization of methylated transcripts to P bodies influences their stability. Overall, these studies display the multifaceted role of location, both molecular and cellular, as a critical contextual aspect in determining m^{6}A and m^{6}A-reader function.

MATERIALS AND METHODS

Table 3.1. Datasets analyzed to study the effects of mRNA m^{6}A location

Data Type	Cell Line	GEO Accession	Reference
Lifetime RNA-seq	HeLa	GSE49339	(Wang et al., 2014a)
Lifetime RNA-seq	mESC	GSE148039	(Lasman et al., 2020)
MeRIP-seq	HeLa	GSE46705	(Lasman et al., 2020; Liu et al., 2014)
MeRIP-seq	mESC	GSE52681	(Batista et al., 2014)
WT, YTHDF1 KD, YTHDF2 KD, YTHDF 3KD, and YTHDF1-3KD RNA-seq	HeLa	N/A	(Batista et al., 2014; Zou et al., 2022)

Table 3.1, continued.

Data Type	Cell Line	GEO Accession	Reference
WT, YTHDF1 KD, YTHDF2 KD, YTHDF 3KD, and YTHDF1-3KD RNA-seq	HeLa	GSE134380	(Zaccara and Jaffrey, 2020)
WT, YTHDF1 KD, YTHDF2 KD, YTHDF 3KD, and YTHDF1-3KD ribosome profiling	HeLa	GSE134380	(Zaccara and Jaffrey, 2020)
P body immunoprecipitation and sequencing	HeLa	N/A	(Zou et al., 2022)

Analysis of relationships between m^{6}A location and RNA stability or translation efficiency

For all data except RNA-seq data from Zaccara *et al*., the analyses were performed starting with processed data files, accessed either from supplemental files in the publication or on the Gene Expression Omnibus website (see **Table 3.1** for a list of GEO accession numbers). Downstream analyses were performed using R v. 4.0.3 (https://cran.r-project.org/). Site and lifetime data joining as well as site filtering were performed using dplyr v. 1.0.7 (https://CRAN.R-project.org/package=dplyr). Specifically, transcripts were filtered to only protein coding transcripts, tables were joined, and sites were selected using the count() function by transcript region, followed by filtering for transcripts with peaks in only one region. Transcripts with peaks only in the 5'UTR, CDS, or 3'UTR, or unmethylated transcripts, were selected for further analysis. Cumulative fraction and boxplots were created using ggplot2 v. 3.3.5 (https://ggplot2.tidyverse.org). Metagene plots were created using an in-house python package

developed by Dr. Chang Ye. Gene ontology analysis was performed using MetaScape (Zhou et al., 2019b). For boxplots, outliers were not shown but were included in statistical analyses.

Analysis of *in vivo* proximity-dependent biotinylation (BioID) data

Processed BioID was downloaded from Table S1 published by Youn *et al*. 2018 (Youn et al., 2018; Zhou et al., 2019b). Analysis was performed using R v. 4.0.3 (https://cran.r-project.org/). Relevant data was selected using dplyr v. 1.0.7 (https://CRAN.R-project.org/package=dplyr), and plots were made using ggplot2 v. 3.3.5 (https://ggplot2.tidyverse.org) and pheatmap v. 1.0.12 (https://CRAN.R-project.org/package=pheatmap).

Analysis of publicly available RNA-seq data

Raw fastq files were downloaded from Gene Expression Omnibus (accession number GSE134380) (Zaccara and Jaffrey, 2020), and quality checked using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters were trimmed using Cutadapt (Martin, 2011). Reads were then mapped to the hg38 human genome (Frankish et al., 2019) using HISAT2 v. 2.1.0 (Kim et al., 2019) with the option --rna-strandedness R. Output sam files were converted to bam files, sorted, and indexed using samtools v. 1.7 (Li et al., 2009). Read counts mapping to each gene were obtained using htseq-count (Anders et al., 2015) with reference to the hg38 annotation gtf file with options -s reverse, -t exon, -f bam, -i gene_id, -m intersection-nonempty and -r pos. Before further analysis, the sum of counts across all samples was computed for each gene in R v. 4.0.3 using the rowSums function, and genes with 10 or fewer mapped reads were removed. The remaining genes passing this filter represented "expressed" genes. Differential expression analysis of expressed genes was then performed with DESeq2 (Love et al., 2014) in R v. 4.0.3. Data on N^6-methyladenosine sites and differential expression of "translated" genes in

HeLa was obtained directly from processed data files published on Gene Expression Omnibus from the study of interest. Further analyses were performed and plots were created in R v. 4.0.3 using the following packages: biomaRt v. 2.44.4 (Durinck et al., 2009), ggplot2 v. 3.3.5 (https://ggplot2.tidyverse.org), dplyr v. 1.0.7 (https://CRAN.R-project.org/package=dplyr), ggpubr v. 0.4.0 (https://CRAN.R-project.org/package=ggpubr), and forcats v. 0.5.1(https://CRAN.R-project.org/package=forcats). When performing analyses involving methylation status, unnamed genes were removed since the published file describing m^6A sites contained only known gene names and m^6A site numbers. For boxplots, outliers were not shown but were included in statistical analyses.

CHAPTER 4: SUMMARY AND PERSPECTIVES

Advancements in understanding the role of m^{6}A in gene expression regulation

About a decade ago, two key breakthroughs led to intense interest in mRNA m^{6}A research. Firstly, immunoprecipitation-based transcriptome-wide mapping technologies provided a tool to study the location and prominence of m^{6}A, revealing its presence among thousands of human and mouse transcripts and its enrichment around stop codons and in 3'UTRs (Dominissini et al., 2012; Meyer et al., 2012). Secondly, the METTL3/METTL14 writer complex was shown to mediate mRNA m^{6}A methylation (Liu et al., 2014), and identification of its eraser, FTO (Jia et al., 2011), uncovered the essential elements of a dynamic regulatory system. These breakthroughs enabled the study of mRNA m^{6}A in space and time.

Further research has improved upon these initial steps. New mapping technologies, including miCLIP and later m^{6}A-SAC-seq, have permitted single-base-resolution sequencing of mRNA m^{6}A (Hu et al., 2022; Linder et al., 2015). Additional m^{6}A effector proteins, including several m^{6}A reader proteins, such as YTHDF1-3, YTHDC1-2, and IGF2BP1-3, have also since been characterized, with functional effects throughout the mRNA lifecycle (Huang et al., 2018; Roundtree et al., 2017a). In particular, YTHDF1-3 proteins have been studied extensively as key mediators of the effects of m^{6}A on mRNA fate in the cytoplasm. Initial reports suggested that YTHDF2 promotes decay of its target transcripts (Wang et al., 2014a), whereas YTHDF1 enhances translation (Wang et al., 2015), and YTHDF3 can work in synergy with YTHDF1 or YTHDF2 to promote translation or decay (Shi et al., 2017a). More recent studies have refined this model, with reports that YTHDF2 can promote target stability under certain circumstances, whereas YTHDF1 can promote decay (Dixit et al., 2021; Xia et al., 2021); additionally, triple

knockdown of YTHDF1-3 leads to significantly increased stability of methylated transcripts (Lasman et al., 2020; Shi et al., 2017a; Zaccara and Jaffrey, 2020). Beyond these cellular consequences, several studies involving m^{6}A effector proteins have also revealed the critical roles of m^{6}A in animal development (Frye et al., 2018) and human disease (Hsu et al., 2017). In aggregate, these studies portray a dynamic system of mRNA m^{6}A regulation in time and space, with far-reaching consequences on health and pathology.

Nonetheless, the picture is incomplete. Effector proteins for m^{6}A on noncoding RNAs are only starting to be characterized by us (Chapter 2) and others (Ma et al., 2019c; Sepich-Poore et al., 2022; van Tran et al., 2019; Warda et al., 2017), and many questions remain unanswered about m^{6}A function(s) on multiple types of noncoding RNA. Furthermore, there exist discrepancies and nuances in putative mRNA m^{6}A function between cell lines and cellular contexts (e.g., stress) that obfuscate a clear set of overarching rules of cellular m^{6}A coordination (Shi et al., 2019). Explicit accounting for cellular context, including the location of m^{6}A on the molecule and in the cell, as well as how those locations influence its functional impact and regulation, may prove useful for elucidating a clearer, more nuanced understanding of epitranscriptomic regulation.

Contributions of this thesis to m^{6}A biology and future directions

Characterization of the METTL5-TRMT112 complex, its role in translation, and future directions

In this work, I presented the characterization of the 18S rRNA m^{6}A writer complex, METTL5-TRMT112 (Chapter 2) (Sepich-Poore et al., 2022). I showed that, although METTL5 includes the m^{6}A catalytic site, TRMT112 is essential for METTL5's stability, and human disease-associated METTL5 mutations disrupt their interaction (**Figure 2.1-2.2)**. The identification of the writer of 18S m^{6}A1832 permitted studying the function of the modification through knockdown

and knockout studies in cell lines, which revealed its roles in cell proliferation and translation regulation (**Figure 2.5-2.6**). Although I found that loss of METTL5 in cell lines had only a minor effect on global translation, the translation of certain transcripts was more affected than others. Findings of 18S m^6A loss and dysregulated translation were recapitulated using a *METTL5* knockout mouse model (**Figure 2.7-2.8**), and phenotypic characterization of these mice revealed significant growth and metabolic defects, potentially related to the underlying translation dysregulation.

Although these findings greatly advanced the functional understanding of the 18S m^6A1832 site, its deposition, and its biological significance, there remain unanswered questions. Regarding cellular localization, imaging and biochemical fractionation showed that, although some METTL5 localizes to the nucleolus, the majority of METTL5 in HeLa and HEK293T cells is located in the cytosol (**Figure 2.3A-B**). The reason for this localization is unknown, but there are several possible explanations worth investigating. Firstly, cytosolic localization of the methyltransferase suggests the possibility that METTL5 and/or the METTL5/TRMT112 complex may still be bound to ribosomes during translation initiation and/or elongation. To this point, western blotting of polysome profiling fractions from wild-type HeLa cells did not display substantial visible signals in monosome or polysome fractions (**Figure 2.5C**). A second explanation for significant cytosolic localization of METTL5 could be that it methylates additional cytosolic substrates. Relatedly, I found 25 METTL5 targets from the CLIP-seq data that had significantly dysregulated m^6A peaks upon METTL5 knockout in the m^6A-seq data (**Figure 2.4D**). Ten of these peaks contain a UAAC motif, similar to the 18S m^6A site, and it would be important to investigate whether any of these transcripts are also METTL5 methylation targets. This could be done using m^6A-immunoprecipitation followed by qPCR (m^6A-IP-qPCR), SELECT (Xiao et

al., 2018), or LEAD-m^{6}A-seq (Wang et al., 2021a) using RNA from wild type and METTL5 knockout cells. A third potential explanation for the cytosolic localization of METTL5 could be late-stage deposition of 18S m^{6}A during ribosomal maturation. Evidence for this has been reported by van Tran *et al*., whose analysis of cryo-EM structure density maps of the 40S subunit revealed an unassigned density potentially attributable to METTL5/TRMT112 only in a the density map for one of the latest phases of maturation (Ameismeier et al., 2018; van Tran et al., 2019). Since eukaryotic rRNA processing extends from the nucleolus to the cytoplasm, methylation of 18S m^{6}A by METTL5/TRMT112 at a very late stage would indeed require cytosolic localization of METTL5. Nonetheless, this late-stage methylation has not been experimentally validated. Such validation could be performed by biochemical fractionation of wild type and METTL5 knockout cells into chromatin-associated, nuclear, and cytosolic fractions, followed by purification of the 18S segment surrounding the m^{6}A site and LC-MS/MS, as done in **Figure 2.3E**. Overall, these experiments, individually or in combination, would shed light on the potential biological mechanism(s) underlying the observed cellular localization pattern of METTL5.

Similar to the discovery of mRNA m^{6}A effector proteins, the characterization of the METTL5/TRMT112 writer complex enables its use as a tool for studying (i) potential dynamic effects of 18S m^{6}A and/or (ii) whether the presence or level of 18S m^{6}A is more important in certain disease, development, or stress contexts. Such studies could also provide further insight into the emerging concept of ribosome heterogeneity (Sloan et al., 2017). Recent reports have demonstrated sites of partial 2'-*O*-methylation and pseudouridylation on rRNA, some of which are regulated in response to environmental changes (Birkedal et al., 2015; Buchhaupt et al., 2014; Schwartz et al., 2014; Sloan et al., 2017; Taoka et al., 2016). An additional study revealed that the universally conserved dimethyladenosine (m^{6,6}A) sites near the 3' end of 18S rRNA can sometimes

be singly methylated as m^6A in yeast, and that there are striking differences in translation between those ribosomes with m^6A versus $m^{6,6}A$ at these sites, particularly for transcripts related to sulfur metabolism (Liu et al., 2021). Importantly, these sites are in close proximity to 18S m^6A1832 that is methylated by METTL5/TRMT112. These reports indicate that rRNA modification status could represent an important aspect of ribosome diversity, which could be involved in selectively adapting ribosomes for translation of specific mRNA subsets.

A few lines of evidence from this study suggest that 18S m^6A 1832 could potentially be differentially methylated in distinct contexts to selectively regulate translation of certain transcripts. Firstly, slightly different levels of 18S m^6A were found by mass spectrometry between mouse brain and liver tissue (**Figure 2.7A**). Secondly, loss of METTL5, and thus loss of 18S m^6A, was found to affect translation of certain transcripts more dramatically than others in human cell lines and mouse liver tissue (**Figures 2.5-2.8**). Future systematic studies of 18S m^6A levels across cell lines and tissues, and in response to various stresses, could reveal the extent to which heterogeneity in methylation exists at this site. In light of these possibilities, it may also prove interesting to study whether cells lacking METTL5 respond uniquely to these stresses.

The 18S m^6A1832 site is also located near the binding sites of multiple important translation initiation and re-initiation factors, such as eIF1, eIF1A, and DENR-MCT-1 (Lomakin et al., 2017; Weisser et al., 2013). Its presence or absence could thus alter the local ribosomal structure and thereby affect binding of these factors. Indeed, a report by Rong *et al*. presented attenuated translation initiation activity in METTL5 knockout cells, evidenced by reduced ribosome binding of initiation factors eIF4E and eIF3A; moreover, there was reduced phosphorylation of RPS6 kinase (S6K), a kinase that is part of the AMPK and mTOR pathways and is involved in regulating translation initiation in response to environmental changes (Rong et

al., 2020). Notably, the polysome profiling experiments presented in Chapter 2 did not reproduce the diminished level of polysomes or the differences in EIF binding to ribosomes presented in their report (**Figure 2.5C**). This discrepancy could be due to differences in cell line and/or cell state. Nonetheless, given the central role of translation initiation in gene expression regulation, the role of METTL5 and 18S m^6A1832 in this process warrants further investigation. For example, screens could be performed involving polysome profiling and western blotting from multiple cell lines and/or conditions to monitor if or how METTL5 loss affects translation initiation and S6K phosphorylation. If positive hits arise, follow-up experiments could trace the signaling pathway by which S6K phosphorylation is linked to METTL5 loss. Elucidating such a pathway would be important for understanding how METTL5 and 18S m^6A may be involved in regulating translation in response to environmental cues.

The creation of a *METTL5* knockout mouse model (Chapter 2) enables future investigation of such potential signaling pathways and mechanisms *in vivo*. Although I performed some basic characterization of the mouse model that validated the loss of 18S m^6A and revealed developmental (**Figure 2.7B, 2.8D**), growth (**Figure 2.7C-D**), and metabolic defects (**Figure 2.7F, 2.8 E-F**), additional phenotypic characterizations could be useful to couple with future mechanistic studies. Firstly, since *METTL5* knockout mice are smaller than their heterozygous littermates even by 4 weeks of age (i.e., the starting point for weight measurements) and are born at frequencies lower than expected, further investigation should be performed to examine the exact developmental stage at which a defect is observable by sacrificing mice at key stages of gestation, followed by observational and genetic analysis of the embryos. Relatedly, a study by Ignatova *et al*. reported that loss of *METTL5* in mESCs led to loss of pluripotency, compromised differentiation potential, and decreased global translation. These data suggest that it is likely that

a developmental defect arises at a very early stage with cascading consequences. Additionally, given the lesser abdominal fat in *METTL5* knockout mice (**Figure 2.7C-D, 2.8E**) and the dysregulation of metabolically-related pathways observed in sequencing experiments (**Figure 2.7F, 2.8H-I**), it would also be valuable to further characterize their metabolism. Key future experiments would include food intake measurement, glucose and insulin tolerance tests (Alquier and Poitout, 2018), and energy expenditure measurements (Speakman, 2013) comparing wild type and knockout littermates. Collectively, these studies would provide insight into the relative contributions of developmental and metabolic differences on resultant mouse weights.

Unexpectedly, basic phenotypic characterizations did not reveal significant deficits in motor activity, exploratory activity, or learning ability in *METTL5* knockout mice compared to heterozygous littermates (**Figure 2.7H-I, 2.8J-L**). This was surprising since human METTL5 mutations have been linked to microcephaly and intellectual disability (Richard et al., 2019), and since other groups have reported neurological and behavioral defects in *METTL5* knockout mice (Ignatova et al., 2020a; Wang et al., 2022). The discrepancy could be due to differences in experimental setup (i.e., sample sizes, ages, and genotypes used) or the types of tests applied. Exactly replicating the tests performed by the other groups and comparing the results would confirm whether loss of METTL5 does indeed cause neurological and behavioral defects, and would also reveal whether and how mouse model design contributes to these effects. It is also possible that the defects caused by METTL5 mutations may be more subtle than is easily detectable by the tests presented in Chapter 2, and that subjecting the mice to more complicated tasks may lead to more significantly observable differences. Additionally, certain key phenotypes observed by Richard *et al*. in human patients have not yet been studied in mouse models, including

aggression and autistic behavior, both of which could be monitored via social interaction tests (Barabas et al., 2021; Silverman et al., 2010).

Comprehensive phenotypic characterization of the mouse model, paired with molecular data from knockout cell lines, would guide future *in vivo* mechanistic studies. For example, it may be valuable to investigate whether translation initiation and/or initiation-related signaling pathways are dysregulated in mouse tissues and correlated to observed metabolic, developmental, or neuronal phenotypes. Extending this idea, if dysregulated S6K activity were to be validated upon METTL5 loss in cell lines and mouse tissues, subsequent studies could examine the impact of this dysregulation on mouse metabolism or learning and memory through pharmacological inhibition of mTOR or AMPK pathways (Ballou and Lin, 2008; Bavley et al., 2018). Additionally, potential heterogeneity in 18S m^6A1832 occupancy could be examined in different mouse tissues, throughout mouse development, or in response to stimuli. In summary, the creation of *METTL5* knockout mouse models in combination with the critical initial biochemical characterization of METTL5/TRMT112 function have laid the groundwork for future mechanistic studies into the roles of METTL5 *in vivo*.

Investigation of the impact of location on stabilization and decay of methylated transcripts

Modifications on mRNA differ from those on rRNA in their functional potential due to fundamental differences in the roles and processing of these molecules. In Chapter 3, I examined the impact of m^6A location on mRNA transcript stability. First, I identified evidence from two cell lines that the regional location of m^6A on a transcript may influence mRNA stability, and that transcripts with m^6A in the 5'UTR exhibit greater stability via a mechanism potentially mediated by YTHDF2 and/or YTHDF3. Then, I introduced data suggesting that the cellular location of

methylated transcripts has a major influence on mRNA stability. Specifically, I presented evidence that (i) m^6A-modified transcripts are enriched in P bodies and (ii) P bodies increase in number upon knockdown of the three m^6A readers of the YTHDF family, YTHDF1-3. These results support a model wherein increased mRNA stability upon triple YTHDF depletion is at least partially mediated by P body dysregulation, and they promote a major role for P bodies in regulating the stability of methylated transcripts. Additionally, these findings highlight the role of location in the mechanisms by which m^6A impacts transcript stability, and bring up several potential avenues of further research.

Although I demonstrated through analysis of published datasets that m^6A in the 5'UTR was correlated with increased stability of methylated transcripts, and that this could be mediated by YTHDF2 and/or YTHDF3, these findings should be experimentally validated in cell lines. For simplicity, future experiments described below are proposed with a focus on YTHDF2, but they can be conducted similarly using YTHDF3. First, top targets with m^6A in the 5'UTR that are stabilized by YTHDF2 could be validated by siRNA knockdown of METTL3 and YTHDF2 followed by qPCR. If these transcripts are indeed stabilized by YTHDF2 via its binding of m^6A in the 5'UTR, depletion of both METTL3 and YTHDF2 should result in their decreased abundance. To specifically assess the lifetimes of the targets, qPCR could be performed on mRNA collected from various time points after actinomycin D treatment of knockdown and control samples. To more directly demonstrate regional differences in the influence of YTHDF2 on target stability, a fusion construct of dCas13b-YTHDF2 could be employed with guide RNAs targeting the 5'UTR, CDS, or 3'UTR of a target transcript that is normally unmethylated. If YTHDF2 can indeed stabilize 5'UTR-methylated transcripts, qPCR analysis would reveal an increased abundance of

the target transcript when the 5'UTR-targeting guide is used, with an expected decrease in abundance from samples involving the CDS- or 3'UTR-targeting guides.

If such experiments validate that YTHDF2 can stabilize targets with 5'UTR methylation, the underlying mechanism should be further elucidated. As described in Chapter 3, there are two major mechanistic models which could explain this effect, both involving a critical partner protein (**Figure 4.1**). The first model ("Model A") involves recruitment by YTHDF2 of a partner protein with a 5'UTR-specific stabilizing mechanism, such as binding to the 5' cap to prevent degradation from the 5' end. The second model ("Model B") involves stabilization by a mechanism that is not necessarily 5'UTR-specific but which becomes dominant in situations where YTHDF2-mediated decay may not be as effective. This stabilization could also be mediated by a partner protein(s). Thus, the next key research step would be to screen potential protein partner(s) that could bind to YTHDF2 to increase transcript stability.

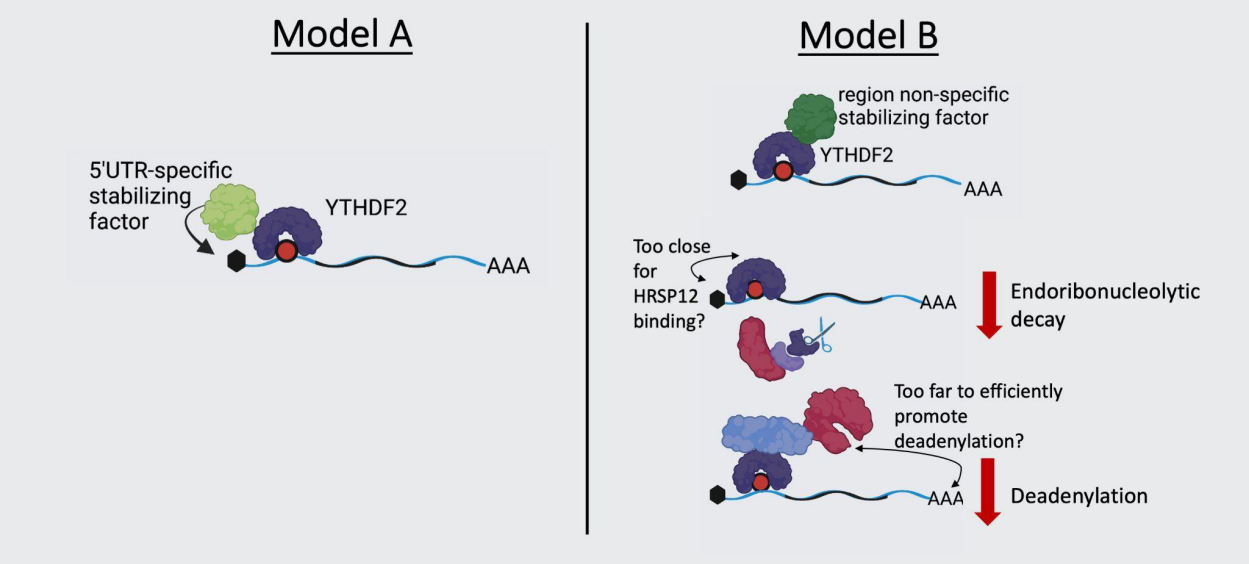


Figure 4.1. Potential models of stabilization of 5'UTR-methylated transcripts by YTHDF2. Model A (left) involves binding of m^6A in the 5'UTR by YTHDF2 and subsequent recruitment of a binding partner that stabilizes the transcript through a mechanism specific to the 5'UTR, such as binding to the 5' cap and protecting from 5'-mediated decay. Model B (right) states that the stabilization of 5'UTR-methylated transcripts could be the consequence of multiple factors, including the recruitment by 5'UTR-bound YTHDF2 of a region nonspecific stabilizing factor along with the decreased efficiency of the canonical pathways by which YTHDF2 promotes

Figure 4.1, continued. transcript decay. Since the mechanism of endoribonucleolytic cleavage related to YTHDF2 requires binding of HRSP12 about 800 nucleotides upstream of YTHDF2, it is unlikely to occur when YTHDF2 is bound in most 5'UTRs, since the median 5'UTR length for human transcripts is approximately 200 nucleotides. Additionally, promotion of deadenylation by YTHDF2 could be less efficient when YTHDF2 is bound to the 5'UTR due to increased distance from the transcript's polyA tail. For either model, identification of the key YTHDF2 interacting partner protein(s) is a critical next step.

The analysis of published bioID data in Chapter 3 identified several candidates for further investigation. Additional computational analyses of publically available data focused on the effect of knockdown of these candidates on RNA abundance changes based on m^6A region could reveal whether any of them preferentially stabilize 5'UTR-methylated transcripts. Experimentally, candidates could be investigated via knockdown screens in cell lines, revealing changes in the ability of YTHDF2 to stabilize targets in the absence of its partner protein—if one is indeed necessary for stabilization. Although single knockdown of YTHDF2 or its critical partner protein should decrease stability of those targets it stabilizes, double knockdown of both would not be expected to decrease target stability further than knockdown of the candidate partner alone. The dCas13b-YTHDF2 construct and region-specific guides could also be employed in the context of knockdown of potential partner proteins, as YTHDF2 targeted to the 5'UTR would likely not be able to stabilize a transcript without its stability-mediating partner. Finally, the interaction between YTHDF2 and its partner protein would need to be validated through co-immunoprecipitation. Identification of the critical partner protein(s) and validation of interaction with YTHDF2 would enable in-depth mechanistic studies in biological systems and disease. For instance, YTHDF2 has been reported to stabilize key oncogenic transcripts in glioblastoma stem cells (Dixit et al., 2021), and it would be valuable to investigate whether the identified partner protein mediates the stabilization of those specific transcripts. Other future directions may depend on the identity of the

partner protein, as it would be important to investigate potential contributions of YTHDF2 to other mechanisms in which it is involved.

In Chapter 3, I additionally presented evidence of major cytosolic changes upon depletion of YTHDF1-3 that globally affect mRNA stability, including a significant increase in the number of P bodies. Experiments, as presented in Chapter 3 and in our submitted manuscript under review (performed with Zhongyu Zou), demonstrated that these P body changes are important for the RNA stabilization observed upon triple YTHDF knockdown (Zou et al., 2022). Those transcripts enriched in P bodies in wild type HeLa, whether methylated or not, were the most stabilized upon triple knockdown (**Figure 3.7D**). Furthermore, knockdown of DDX6, a protein essential for P body assembly, led to global destabilization of RNA; moreover, knockdown of DDX6 in combination with YTHDF1-3 abolished the stabilization effect of triple YTHDF1-3 knockdown (Zou et al., 2022).

Although these observations establish the essential role of P bodies in the RNA stabilization effect observed upon YTHDF1-3 knockdown, they also point to directions of further research. Firstly, it would be helpful to perform overexpression of DDX6 in the context of YTHDF1-3 knockdown to validate that RNA is globally stabilized under those conditions. Secondly, it would be valuable to weigh the relative contributions of P body incorporation versus loss of YTHDF-mediated decay on RNA stabilization during YTHDF1-3 triple knockdown. In **Figure 3.7D**, for transcripts depleted from P bodies in wild type HeLa, those that are methylated are more stabilized after YTHDF1-3 triple knockdown than unmethylated transcripts, but it is unclear whether this is due to increased P body incorporation of methylated transcripts or a decrease in their degradation due to YTHDF depletion. One experiment that could shed light on this question would be purification of the P body transcriptome after YTHDF1-3 triple knockdown

to determine which transcripts are stabilized through P body incorporation, as well as whether there is a preference for incorporation of methylated transcripts and/or YTHDF targets. The potential redundancy in YTHDF function could also be investigated in the context of single knockdown of each YTHDF protein by performing lifetime analyses along with CLIP-seq of the other two YTHDF proteins. These experiments would indicate the extent to which each YTHDF protein contributes to decay, and whether they compensate for each other in certain contexts.

Broader implications of this thesis to the field of m^{6}A biology

Overall, these studies uncovered important aspects of m^{6}A mRNA biology and suggest multiple avenues for future investigation. Clearly, the functional effects of m^{6}A on different RNA types, as well as in distinct molecular and cellular contexts, highlights the significance of location when investigating the roles of m^{6}A in gene expression regulation (**Figure 4.2**).

In Chapter 2, I proposed that the location of the 18S m^{6}A1832 site at the tip of helix 44 near the decoding center is important for its function, since this region of the ribosome interacts with mRNA, tRNA, and several important initiation factors. In line with this, another study found that 18S m^{6}A1832 may fine-tune the interaction between the decoding center and mRNA, and that METTL5 promotes proper translation initiation (Rong et al., 2020; Shi et al., 2019). In contrast, the 28S m^{6}A4220 site is further from the decoding center and closer to the surface of the 60S subunit, near the interface with the 40S subunit (Ma et al., 2019c; Pinto et al., 2020). Although loss of the methyltransferase for this modification (ZCCHC4) also affects translation and cell proliferation (Ma et al., 2019c; Pinto et al., 2020), it was additionally found to impact ribosome subunit levels and assembly (Ma et al., 2019c). Comparing the impacts of loss of the same modification at different sites on the ribosome highlights the importance of local context in

determining modification function. As both methylation sites are on rRNA, they impact translation, but the exact mechanisms by which they do so have key differences based on their location.

The findings presented in Chapter 3 add nuance to existing models of YTHDF function by revealing differences in functional effects based on location at two different levels: on the transcript and in the cell. Specifically, in contrast to transcripts with m^6A in the 3'UTR or CDS which have reduced stability, transcripts with m^6A in the 5'UTR display increased stability compared to unmethylated transcripts (**Figure 3.1**). Transcripts that are enriched in the P bodies of wild-type cells also display a different stabilization pattern upon triple YTHDF1-3 knockdown than those depleted from P bodies, indicating that the cytosolic location of methylated transcripts plays an influential part in the role(s) of YTHDF binding in mRNA stability (**Figure 3.7D**). Future investigation may elucidate the precise mechanisms underlying the observations presented in Chapter 3, but the observations themselves substantiate the importance of location in determining the effects of m^6A on transcript stability.

For RNA modifications, understanding their regulatory capacity requires not only a knowledge of their location, but also an assessment of their dynamics–it matters not only where the modification is, but when it is there. For example, the addition or removal of m^6A to or from a transcript in response to an environmental stimulus may determine how proteins interact with that transcript, thus promoting either stability or decay. Over a decade ago, breakthroughs in the ability to study mRNA m^6A in time and space, using transcriptome-wide mapping and modulation of effector proteins, led to a burst of interest in the field. As advancements have been made, the specific questions have changed, but an awareness of context, including where a modification is

and when it is there, is still essential for elucidating the complex impacts of RNA modifications on cellular functions.

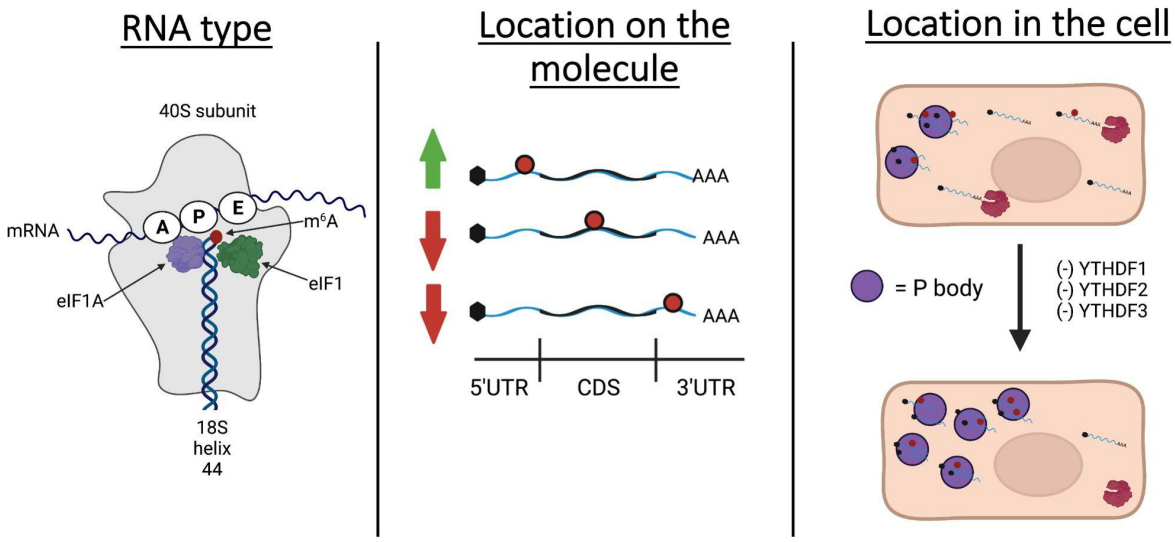


Figure 4.2. Aspects of location that impact m^{6}A function. The functional importance of m^{6}A location is highlighted in this thesis in three major ways. Firstly (left), the type of RNA on which the modification exists can determine its functional effects, as 18S m^{6}A was shown to have an important effect on translation (Chapter 2). Furthermore, the location of the m^{6}A on the molecule at the tip of helix 44 near the decoding center and binding sites of initiation factors likely plays an important role in modulating translation of specific transcripts. Secondly (middle), the location of m^{6}A on mRNA, whether in the 5'UTR, CDS, or 3'UTR, can be involved in determining the impact of the modification on transcript stability, with 5'UTR m^{6}A associated with increased transcript stability, whereas m^{6}A in the CDS or 3'UTR correlates with increased decay (Chapter 3). Thirdly (right), the location of methylated transcripts within the cell can impact the functional consequences of m^{6}A, as localization of methylated transcripts to P bodies can play an important role in transcript stability; moreover, increased P bodies may be a significant mediator of increased mRNA stability upon YTHDF1-3 knockdown (Chapter 3).

Perspectives for future investigations of RNA modifications

Moving forward, investigations into temporal and spatial contexts comprise the forefront of research on RNA modification biology. Regarding rRNA modifications, I (Chapter 2) and others have revealed that changes in modification levels can alter ribosome structure and translation (Erales et al., 2017; Rintala-Dempsey and Kothe, 2017; Sepich-Poore et al., 2022). Recently, it has been proposed that ribosome composition, including rRNA modification(s), may

be heterogeneous, allowing the formation of "specialized ribosomes" to regulate translation in different contexts (Genuth and Barna, 2018; Shi et al., 2017b). Such a model requires further elucidation and, in particular, it is unknown whether or how rRNA m^6A methylation may change under different conditions to physiologically regulate translation. However, the recent identification of rRNA m^6A writer proteins helps to enable investigations addressing this idea.

Regarding mRNA m^6A, my data argue that it is critical to carefully consider location in RNA modification study designs. My results suggest that YTHDF2 differs in its function at the 5'UTR, although the exact mechanism remains unclear. It is possible that other RNA modification readers may similarly display differences in function based on transcript region. Findings presented in Chapter 3 about the critical role of P bodies in regulating stability of methylated transcripts further emphasize the importance of considering physical cellular location when investigating mechanisms of mRNA m^6A function. Future studies should be performed to elucidate the mechanistic role(s) of P bodies in affecting the stability of methylated transcripts. More broadly, future research of mRNA m^6A must standardize analyses by mRNA region and cellular location to create more complete and accurate models. Promisingly, as an example of this, recent studies have placed a larger focus on cellular fractionation and granule purification to identify differences among m^6A effectors in distinct cellular regions (Liu et al., 2020a; Roundtree et al., 2017b; Zou et al., 2022).

Moving beyond m^6A, there remain many other modifications that are less well-characterized. Critical steps for the further study of these modifications will likely reflect the past path of m^6A research: the development of high-resolution mapping methods, and the characterization of key effector (i.e., reader, writer, eraser) proteins. Although this has already been achieved in some cases (Legrand et al., 2017; Xue et al., 2020; Zhang et al., 2019; Zhou et

al., 2019a), accomplishing these goals for all RNA modifications is essential for developing a more complete understanding of their dynamic regulation of gene expression in time and space.

APPENDIX: PROGRESS IN SINGLE-BASE RESOLUTION MAPPING OF N^3-METHYLCYTIDINE ON mRNA AND caRNA

Preface Note:

This work was done in collaboration with Dr. Qing Dai, Dr. Hui-Lung Sun, Dr. Xiaoyang Dou, and Kinga Pajdzik. Overall ideation and library preparation method development were done by Dr. Qing Dai. Initial screening of reverse transcriptases with m^3C-containing probes (**Figure A.2A**), the first set of libraries using spike-ins, and tRNA analyses (**Figure A.3**) were performed by Kinga Pajdzik. Data analysis for these two experiments was performed by Dr. Xiaoyang Dou. RNA samples from HeLa shControl and shMETTL8 knockdown stable cell lines were provided by Dr. Hui-Lung Sun. All other experiments were performed by me, including mass spectrometry, qPCR, library preparation and bioinformatic analysis for the paired control and METTL8 knockdown mRNA and caRNA samples from HeLa cells, as well as AlkB expression, purification, and activity testing. All work was guided and supervised by Prof. Chuan He.

INTRODUCTION

N^3-methylcytidine (m^3C) is an RNA modification involving methylation of cytidine at the Watson-Crick face (**Figure A.1A)**, and has been found in human tRNA (Xu et al., 2017), mitochondrial tRNA (mt-tRNA) (Kleiber et al., 2022; Lentini et al., 2022; Schöller et al., 2021; Xu et al., 2017), and mRNA (Xu et al., 2017). Outside of tRNA, however, m^3C remains understudied compared to many other modifications because of its low abundance (2-4 x 10^{-5} m^3C/C in human mRNA), the lack of suitable antibodies for immunoprecipitation-based mapping, and few known m^3C effector proteins (Bohnsack et al., 2022; Cui et al., 2021; Xu et al., 2017).

Still, the chemical properties of m^3C, including disrupted base pairing and charge effects, as well as its important tRNA functions, suggest its potential potency in other RNA types.

Specifically, the location of the methylation on m^3C causes significant disruption of the Watson-Crick C:G pair stability, and hinders base pairing discrimination between C:G and mismatched base pairs (Mao et al., 2021). Increased base pair mismatching can cause destabilization of secondary structure, as well as stopping of or mutations during reverse transcription (Mao et al., 2021). In addition, the presence of the m^3C methyl group may regulate binding by reader proteins, although none have yet been characterized (Bohnsack et al., 2022).

In contrast, several m^3C writers and erasers have been characterized. In humans, ALKBH1 and ALKBH3 have been shown to demethylate m^3C-modified RNAs, including mRNA and tRNA (Bohnsack et al., 2022; Ma et al., 2019a; Ueda et al., 2017). There exist four m^3C methyltransferase-like proteins in humans: METTL2A, METTL2B, METTL6, and METTL8 (Bohnsack et al., 2022; Cui et al., 2021). In human tRNA, METTL2A and METTL2B install m^3C_{32} in threonine isoacceptors (*Hs*tRNA$^{Thr(AGU/CGU/UGU}$), and METTL6 introduces m^3C_{32} in serine isoacceptors (*Hs*tRNA$^{Ser(AGA/CGA/UGA/GCU)}$) (Ignatova et al., 2020b; Lentini et al., 2020; Xu et al., 2017). The writer(s) for m^3C in leucine and methionine tRNAs remain unknown (Bohnsack et al., 2022). Although METTL8 was originally reported to be a mRNA m^3C methyltransferase, recent reports have characterize it as the writer for m^3C_{32} in mt-tRNA^{Thr} and mt-tRNA$^{Ser(UCN)}$ (Kleiber et al., 2022; Lentini et al., 2022; Schöller et al., 2021). Interestingly, RNA m^3C can also be introduced by direct, non-enzymatic methylation by endogenous and environmental agents, such as methyl methanesulfonate (MMS) (Wollen et al., 2021).

Of the characterized m^3C writers, the function(s) of METTL8 are perhaps the most ambiguous. Although its mt-tRNA m^3C writer function is now well-established (Kleiber et al.,

2022; Lentini et al., 2022; Schöller et al., 2021), mRNA m^3C levels on mRNA have been found to be significantly lower in liver samples from METTL8 knockout mice compared to wild-type controls (Xu et al., 2017). Related to potential activity on mRNA, METTL8 has also been found in mouse embryonic stem cells (mESCs) to interact with the transcript encoding MAPKBP1, a component of the c-Jun-N-terminal kinase signaling pathway, causing potential consequences for mESC differentiation (Gu et al., 2018). Additionally, METTL8 has been implicated in regulating the formation of R-loops (DNA-DNA:RNA hybrid structures), particularly in the nucleolus, with functional impacts on genome stability (Zhang et al., 2020a).

One possible explanation for the differing potential functions observed for METTL8 may derive from its diverse isoforms (Badri et al., 2008; Schöller et al., 2021). The most highly expressed isoforms in human cells contain a mitochondrial targeting signal (MTS), but isoforms lacking the MTS are also detectable (Schöller et al., 2021); moreover, when the MTS is removed, METTL8 accumulates in the nucleus (Lentini et al., 2022; Zhang et al., 2020a). Therefore, METTL8 isoforms lacking the MTS could mediate m^3C deposition on mRNA, caRNA (including in R-loop RNA), and/or other RNA types located in the nucleus. Intriguingly, over a decade ago, METTL8 isoforms were first identified as a set of transcription regulators (then referred to as tension-induced/inhibited proteins, "TIPs"), which could modulate transcription in response to stretch, and were found to interact with the histone acetyltransferase p300 (Badri et al., 2008). A single mechanistic model connecting all of these disparate and isoform-driven functions for METTL8 is yet to be elucidated.

To investigate m^3C sites on less abundant RNAs, and the concomitant role(s) of METTL8 or other effector proteins regulating them, it is critical to map their locations. Although many techniques have been developed to study m^3C on tRNA and mt-tRNA, none have yet been reported

to reproducibly map m^3C on mRNA. Mass spectrometry techniques provide quantitative information about m^3C levels of in bulk RNA, but they do not provide information about specific location(s) (Bohnsack et al., 2022). Primer extension assays can investigate modification(s) on specific target sites, but they cannot perform unbiased site discovery (Bohnsack et al., 2022). No suitable m^3C antibodies have yet been found for application in immunoprecipitation-based sequencing methods (Cui et al., 2021). Moreover, despite the development of numerous next-generation sequencing-based methods, such as DM-tRNA-seq (Zheng et al., 2015), DAMM-seq (Zhang et al., 2021), ARM-seq (Cozen et al., 2015), mim-tRNA-seq (Behrens et al., 2021), AlkAniline-seq (Marchand et al., 2021), and HAC-seq (Cui et al., 2021), none of these methods have reported m^3C sites on mRNA, even though mRNA m^3C has been detected by mass spectrometry (Xu et al., 2017). Notably, some of these methods (e.g. HAC-seq, AlkAniline-seq) cause extensive RNA degradation, potentially making it difficult to prepare libraries for less abundant RNA types.

Without the ability to identify specific m^3C sites on mRNAs and/or caRNAs, their specific molecular and cellular roles, as well as their possible relevance to human disease, will remain poorly understood. Therefore, building upon previous progress, we endeavored to map m^3C in human mRNA and caRNA, and, by doing so, gain insight into its potential functions.

RESULTS

We first confirmed the presence and abundance of m^3C in multiple HeLa cell RNA types. This comprised purifying mRNA, caRNA, and small RNA (smRNA) from HeLa cells, followed by LC-MS/MS, which revealed that the levels of mRNA and caRNA m^3C were about three orders of magnitude lower than in smRNA, which includes tRNA (**Figure A.1B**). Moreover, there was a

higher level of m^3C in caRNA than in mRNA (**Figure A.1B**). Although the mRNA and caRNA m^3C levels were low, they remained similar across experiments (biological replicate experiment for caRNA not shown), and the mRNA m^3C level observed on mRNA was similar to that reported in HCT116 cell mRNA (Xu et al., 2017).

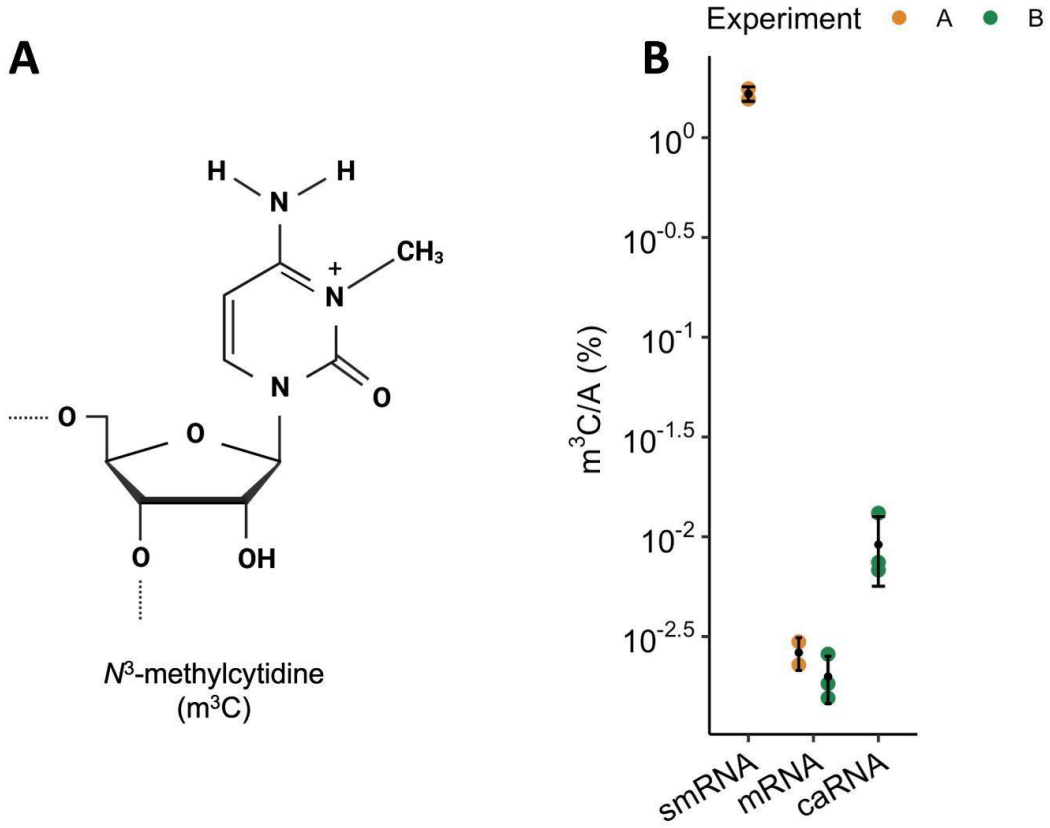


Figure A.1. N^3-methylcytidine in human RNA. (**A**) The chemical structure of N^3-methylcytidine (m^3C). (**B**) LC-MS/MS analysis of m^3C levels normalized to adenosine (A) from different RNA types isolated from HeLa cells, shown using a $\log_{10}$ scale. smRNA = small RNA, mRNA = messenger RNA, caRNA = chromatin-associated RNA. Data is shown from two separate experiments, both including mRNA samples (yellow = experiment A, green = experiment B). n=2 replicates were included in experiment A, n = 3 replicates were included in experiment B. Mean and s.e.m. are plotted.

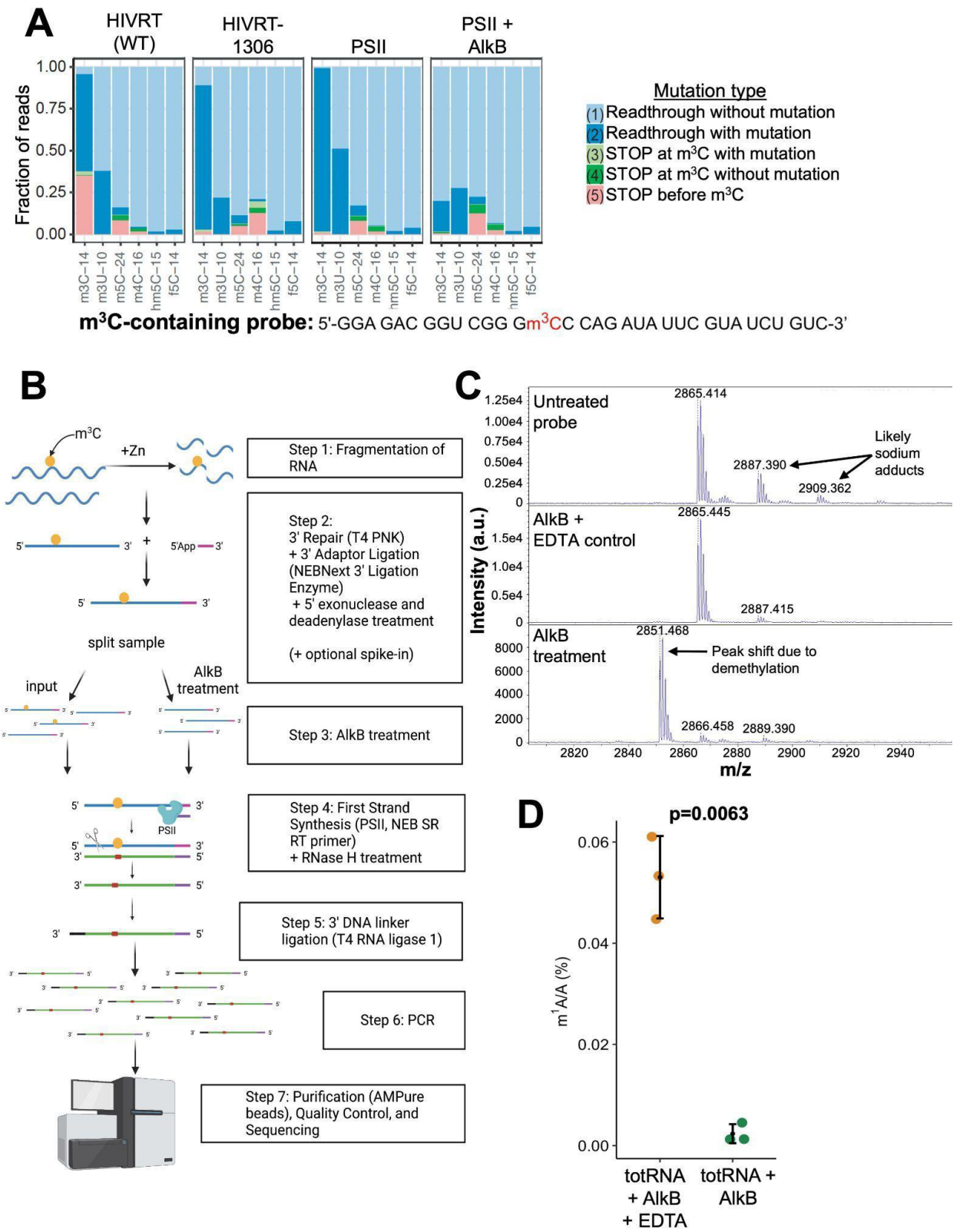


Figure A.2. Method development for sequencing m^{3}C in human mRNA and caRNA. (**A**) Fractions of reads with different mutation types at modification sites on probes from reverse transcription reactions using three different reverse transcriptases, HIVRT, HIVRT-1306 (evolved

Figure A.2, continued. for m^1A detection), and PSII, as well as from a reaction with PSII after AlkB treatment. The sequence for the m^3C-containing probe is displayed below. Mutation types are indicated by color (light blue = readthrough without mutation, dark blue = readthrough with mutation, light green = stop after m^3C site with a mutation, dark green = stop after m^3C site without a mutation, pink = stop before the m^3C site). (**B**) A schematic of the library preparation method employed for m^3C-sequencing in this study. (**C**) Intensity versus mass-to-charge ratio from matrix-assisted laser desorption ionization (MALDI) characterization of m^3C-containing probes from three conditions: untreated, AlkB treated with EDTA to chelate iron, and AlkB treated. A shift in the major peak of -15 is shown in the AlkB treated sample, indicating loss of a methyl group. (**D**) LC-MS/MS analysis of m^1A levels normalized to adenosine (A) from *in vitro* demethylation reactions of AlkB on HeLa total RNA are shown with (left) or without (right) EDTA. n=3 replicates. Mean and s.e.m. were plotted. *P* value was calculated using a two-sided T-test.

Given mass spectrometry evidence that m^3C consistently exists in these RNA types, we explored reverse-transcriptase-mediated strategies for mapping it. We included an m^3C-containing probe in a screen of reverse transcriptases reported to be mutagenic at sites of RNA modification (Mao et al., 2021; Potapov et al., 2018) (**Figure A.2A**), finding that Protoscript II (PSII, third panel) was highly mutagenic at m^3C sites, with low levels of mutation at other C modification sites (i.e., m^4C, m^5C, hm^5C, and f^5C). Use of PSII also induced few stops before or after the m^3C site compared to HIVRT. Although the HIVRT-1306 mutant developed by Zhou *et al*. displayed similar results, PSII is more widely commercially available (Zhou et al., 2019a). Therefore, we concluded that PSII may be a suitable reverse transcriptase for m^3C sequencing.

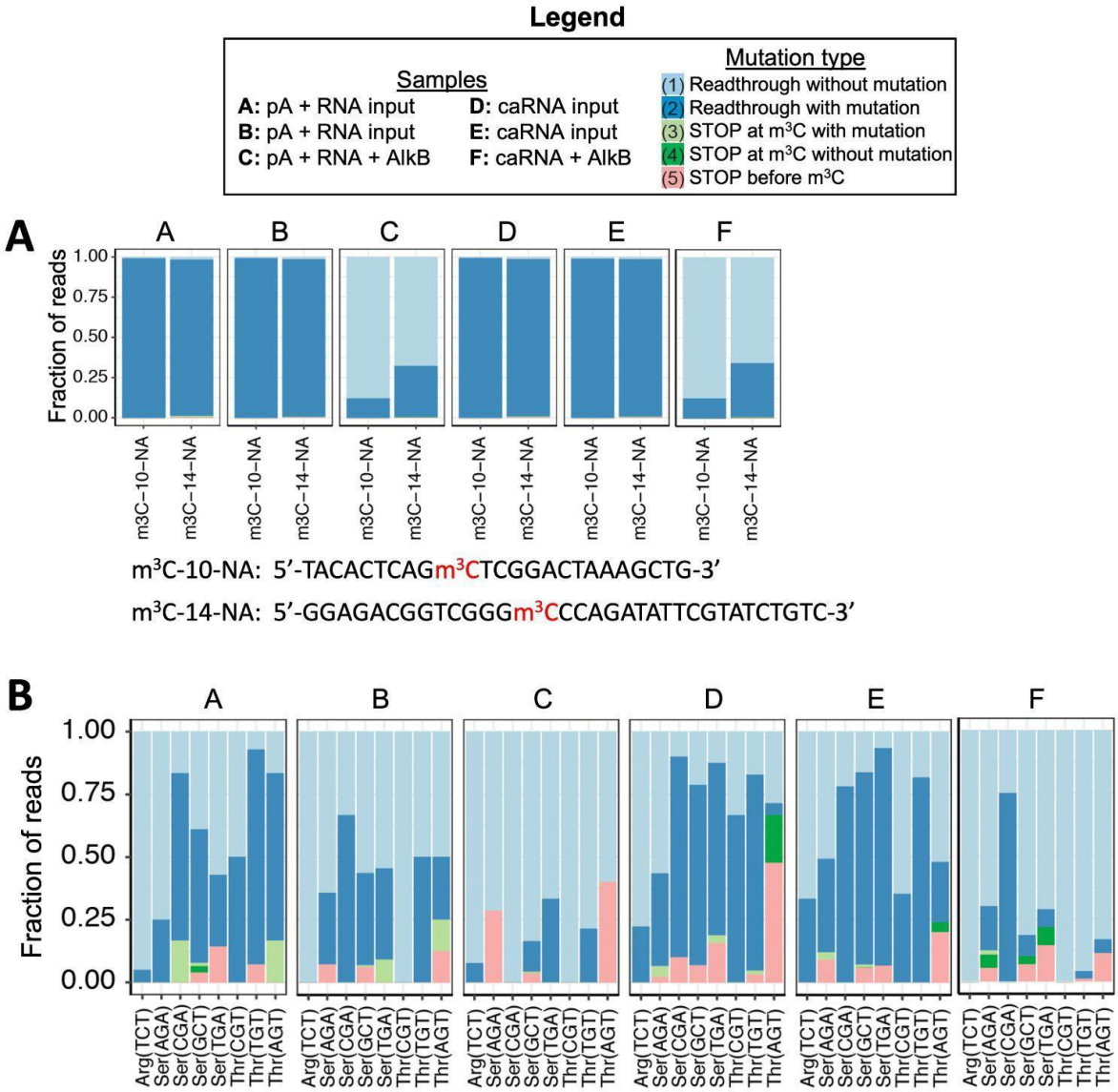


Figure A.3. Detection of mutations at m^{3}C sites in probes and tRNA from libraries. (**A**) Fractions of reads with different mutation types at m^{3}C sites on probes spiked into mRNA and caRNA sequencing libraries. Probe sequences are shown below the plot. (**B**) Fractions of reads with different mutation types at a set of known m^{3}C sites on human tRNA from mRNA and caRNA libraries. Library samples are annotated as follows: A = polyA-enriched RNA input 1, B = polyA-enriched RNA input 2, C = polyA-enriched RNA AlkB control, D = caRNA input 1, E = caRNA input 2, F = caRNA AlkB control. Mutation types are indicated by color (light blue = readthrough without mutation, dark blue = readthrough with mutation, light green = stop after m^{3}C site with a mutation, dark green = stop after m^{3}C site without a mutation, pink = stop before the m^{3}C site).

By modifying existing library preparation methods, we then devised a strategy for implementing PSII-mediated reverse transcription to selectively map m^{3}C on mRNA and caRNA

(**Figure A.2B**) (Cozen et al., 2015; Zhang et al., 2021; Zheng et al., 2015; Zhou et al., 2019a). Similar to several existing methods for sequencing m^3C on tRNA (Cozen et al., 2015; Zhang et al., 2021; Zheng et al., 2015) and m^1A on mRNA (Zhou et al., 2019a), we split each sample and employed AlkB demethylation; this created matched samples in which m^3C is either present at cellular stoichiometries or demethylated (**Figure A.2B**). To use in these libraries, we expressed and purified AlkB, verifying its m^3C demethylating activity on a synthetic probe (**Figure A.2C**), as well as its activity in demethylating m^1A from HeLa total RNA (totRNA) (**Figure A.2D**).

We next tested this method in two pilot experiments. In the first, we sequenced one sample each of HeLa mRNA and caRNA, wherein each sample was split into three fractions (i.e., two inputs and one that underwent AlkB demethylation), for a total of six libraries. The second experiment was similar but added paired METTL8 knockdown samples for mRNA and caRNA (i.e., samples from HeLa stable cell lines with either non-targeting control or METTL8-targeting shRNAs), for a total of twelve libraries. The first experiment included spike-in probes, as well as more extensive fragmentation and shorter read lengths; however, these conditions led to suboptimal mapping rates (<80%). Thus, for the second experiment, we fragmented for a shorter duration, although this precluded spike-in probes because of their much smaller length compared to the larger library size.

Examining spike-in probes from the first experiment, we found a 100% mutation rate at the m^3C site in all input samples for both probes, as well as significantly reduced mutation rates in the AlkB-treated counterparts (**Figure A.3A**). Furthermore, although these libraries comprised polyA-purified RNA and caRNA to map m^3C on mRNA, we also checked mutation rates on a subset of six known tRNA m^3C sites (**Figure A.3B**). These data, especially from the polyA-enriched samples, were limited by lower read depth but indicated that, at least for $tRNA^{Thr}$ and

$tRNA^{Ser}$, the presence of m^3C induced a high rate of mutation in the input samples that was diminished in the AlkB-treated samples. A low mutation rate was observed in mRNA input library 2 (panel B) for $tRNA^{Thr(CGT)}$, but only two reads overlapped that site (after deduplication), making the conclusiveness of the result difficult. Moreover, although the m^3C site on the arginine isoacceptor displayed much lower mutation rates, especially in mRNA samples, it has been shown to have a lower modification fraction than other tRNA sites we examined (Cui et al., 2021). Repeating this experiment with a focus on isolating smRNA may indeed confirm the ability of the method to detect these known tRNA m^3C sites.

The fragmentation-optimized second experiment identified hundreds of potential m^3C sites on mRNA and caRNA from control HeLa cells when the analytical pipeline involved filtering for sites with at least 20 reads, an average mutation rate in the input samples of 0.2, and at least 10% demethylation by AlkB (**Figure A.4A**). Requiring a more stringent demethylation cutoff of 50% still left 12 mRNA sites and 8 caRNA sites that overlapped between both input samples from control HeLa cells (**Figure A.4B**). In polyA-purified RNA samples, sites were most commonly found on mRNA (**Figure A.4C**), whereas in caRNA samples, sites were predominantly identified on promoter-associated RNA (paRNA) (**Figure A.4D**). An examination of tRNA sites was not conducted, because there were very few reads aligning to tRNAs due in part to the larger library sizes used.

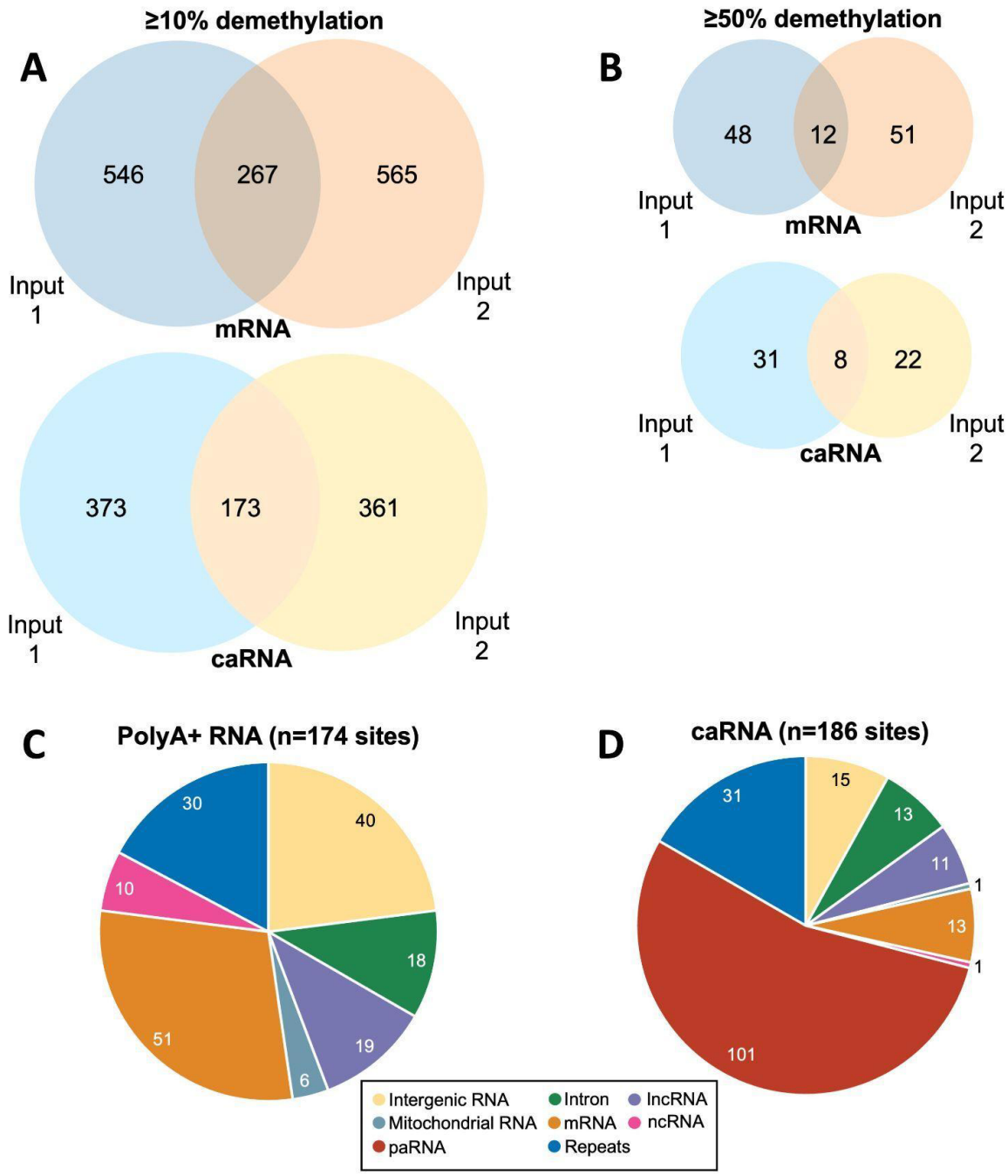


Figure A.4. Putative m^{3}C sites detected in HeLa cells. Venn diagrams of the numbers of putative m^{3}C sites detected in polyA-enriched RNA (i.e. "mRNA") and caRNA m^{3}C-seq libraries from control HeLa cells using a cutoff for demethylation by AlkB of ≥ 10% (**A**) or ≥ 50% (**B**). Pie charts of the types of RNA containing putative m^{3}C sites in libraries from polyA-enriched RNA (**C**) and caRNA (**D**). PolyA+ denotes polyA-enriched.

The inclusion of METTL8 knockdown samples in the second experiment enabled exploration of its potential m^3C-related roleson these RNA types. qPCR confirmed a high knockdown efficiency(**Figure A.5A**). Analyzing the sequencing data for differences in METTL8 knockdown versus control samples revealed 108 sites in polyA-purified samples and 117 sites in caRNA samples that decreased with METTL8 depletion, after discarding known SNP sites (**Figure A.5B**). Of these, 27 sites in polyA-purified RNA and 36 sites in caRNA decreased more than 20% upon METTL8 knockdown (**Figure A.5B**). One prominent example is SLC25A6 (also known as ADT3), which displayed high mutation rates in both control HeLa input samples (i.e., m^3C presence) that were significantly decreased in the AlkB control and all METTL8 knockdown samples (NB: SLC25A6 is on the minus (-) strand, so the RNA $m^3C \rightarrow A$ mutation is displayed as to G→T in IGV) (**Figure A.5C**). We also found a striking distribution of m^3C in caRNA, with the majority of sites mapping to paRNA, followed by caRNA transcribed from repeat elements ("repeats") (**Figure A.5D**). Because these RNA types have been associated with regulation of transcription (Han et al., 2007; Liu et al., 2020a; Mapelli et al., 2019), we performed gene ontology analysis of the genes nearest the caRNA m^3C sites, finding metabolic pathway enrichment, specifically RNA metabolism and TP53-regulated metabolic genes (**Figure A.5E**). We also noticed pathway enrichment of mechanical stimulus responses, potentially relating to METTL8's original characterization as a tension-induced transcription regulator (Badri et al., 2008; Jakkaraju et al., 2005).

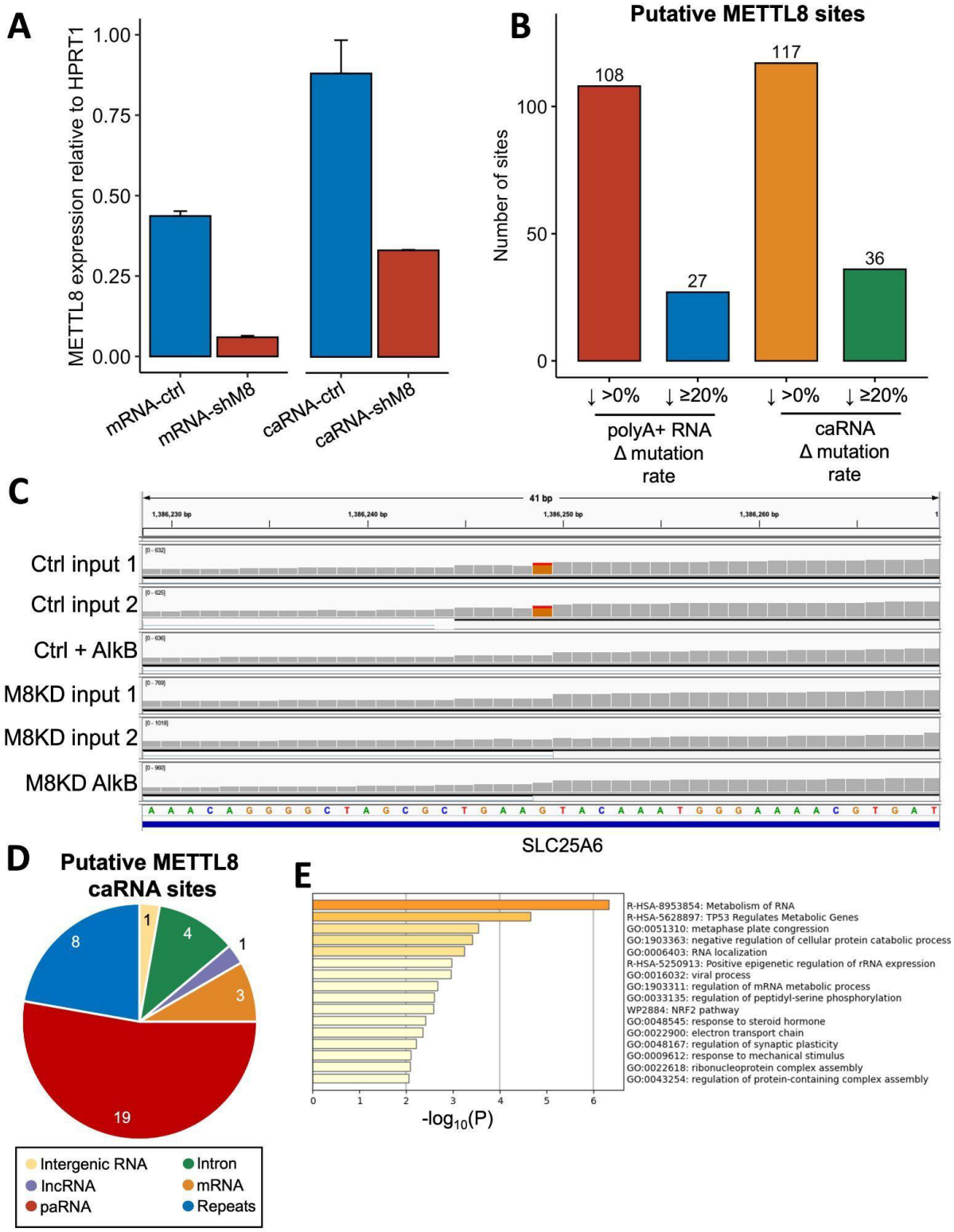


Figure A.5. Effects of METTL8 depletion on m^{3}C sites. (**A**) METTL8 expression relative to HPRT1 (housekeeping gene) in polyA-enriched RNA (left) and caRNA (right) purified from HeLa stable cell lines expressing non-targeting control shRNA (blue) or shRNA targeting METTL8 (red). (**B**) Numbers of putative m^{3}C sites detected in polyA-enriched RNA (left) or caRNA (right)

Figure A.5, continued. libraries that decreased in mutation rate in METTL8 knockdown samples compared to control samples by >0% or ≥20%. (**C**) Visualization of the reads aligned to SLC25A6 in HeLa control input, HeLa control AlkB-treated, HeLa METTL8 knockdown, and HeLa METTL8 knockdown AlkB-treated libraries, adapted from Integrated Genomics Viewer. Nucleotides at bottom represent reference bases from hg38. (**D**) Pie chart of the RNA type of putative m^3C sites detected in caRNA libraries that decrease by ≥20% in METTL8 knockdown compared to control libraries (total n = 36 sites). (**E**) MetaScape gene annotation terms for nearby genes to putative m^3C sites detected in caRNA libraries that decrease by ≥20% in METTL8 knockdown compared to control libraries (n = 36).

To further investigate the relationship between caRNA m^3C and transcription regulation, we compared our control and METTL8 knockdown polyA-enriched input samples using differential gene expression (NB: to be clear, differential expression determination did not consider methylation status). Notably, the samples strongly separated based on METTL8 knockdown status, signifying a large transcriptional perturbation (**Figure A.6A**). Specifically, we found >5000 differentially expressed transcripts ($p_{adj} \leq 0.05$) in the METTL8 knockdown samples compared to controls (**Figure A.6B**). Subsequent gene ontology analyses further revealed dysregulation of mechanosensitive pathways (cf. **Figure A.6C**), including bone and vasculature development, again highlighting METTL8's role as a tension-responsive transcription regulator. Interestingly, the top two transcription factors that regulated these differentially expressed genes (mRNA) upon METTL8 knockdown were STAT3 and TP53, both of which represent major cancer pathways (**Figure A.6D**) (Olivier et al., 2010; Siveen et al., 2014). One study also reported that STAT3 regulates METTL8 expression, with consequences for mESC differentiation (Gu et al., 2018). Moreover, in our prior analysis that examined caRNA m^3C site-proximal genes (**Figure A.5E**), TP53-regulated metabolic genes were also enriched, complementing the METTL8 knockdown differential expression finding. Overall, the enrichment of TP53-regulated metabolic genes in these two analyses may suggest a mechanistic role for caRNA m^3C in regulating TP53 target transcription levels.

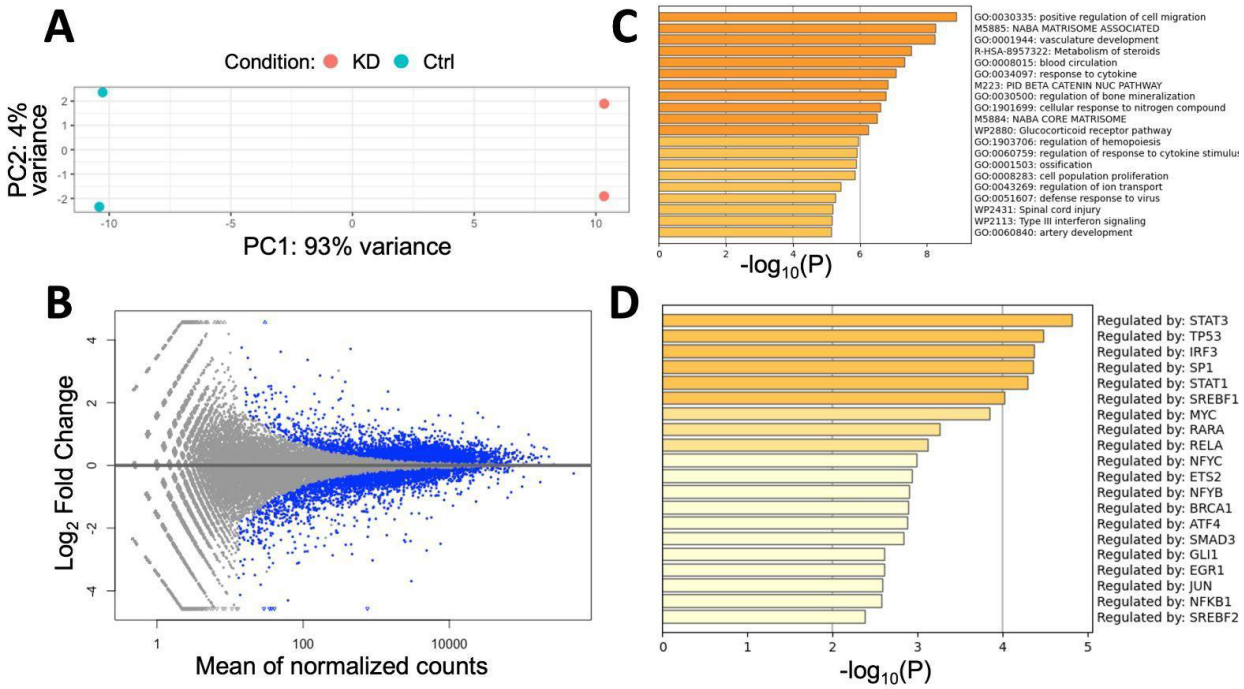


Figure A.6. Changes in the transcriptome with METTL8 depletion. (**A**) Principal component analysis (PCA) of normalized RNA-seq data from control and METTL8 knockdown HeLa. (**B**) MA plot showing the mean of normalized counts versus the $\log_2$(FoldChange) comparing control and METTL8 knockdown samples. Blue dots denote significantly differentially expressed genes ($p_{adj} \leq 0.05$). (**C**) MetaScape gene annotation terms for transcripts differentially expressed in METTL8 knockdown samples compared to control. (**D**) MetaScape identification of transcription factors that regulate transcripts differentially expressed in METTL8 knockdown samples compared to control.

DISCUSSION

Our pilot studies reveal that PSII has a high readthrough and mutation rate for m^3C sites (**Figure A.2A**). Utilizing it in combination with AlkB demethylation for sequencing of m^3C led to the detection of several known m^3C sites in tRNA, as well as many potential m^3C sites in mRNA and caRNA (**Figures A.3B, A.4, A.5)**. However, it is important to emphasize that these results are from pilot studies that did not include biological replicates, with only two technical replicates for input samples and one AlkB demethylation control replicate for each condition.

Several shortcomings of the method were detected that suggest needed optimization. Although several tRNA m^3C sites were picked up with high PSII-mediated mutation rates, the mutation rates for $tRNA^{Arg(UCU)}$ were low in libraries from polyA-enriched RNA (**Figure A.3B**). This may indicate method variability depending on surrounding sequences, structures, and chemical modifications, or it may be due to simple lack of smRNA enrichment. Additionally, low overlap existed between replicates (**Figure A.4A**), and we often observed high variation in mutation rates and AlkB demethylation rates between replicates, even for the same sites (data not shown). These results emphasize that further optimization should be explored, potentially including enzyme evolution, condition optimization (e.g., temperature, reagent concentrations), and adjustments to the bioinformatic analysis pipeline. Future endeavors beyond these pilot experiments should also include additional biological and technical replicates to measure robustness and reproducibility.

Despite these limitations, several interesting insights emerged from the data produced here that hint at avenues of future investigation. The METTL8 knockdown experiments revealed many sites with decreased mutation rates upon METTL8 knockdown (**Figure A.5B**), supporting previous reports that METTL8 may have functions beyond mere mt-tRNA methylation (Tsao et al., 2021; Xu et al., 2017; Zhang et al., 2020a). The existence of multiple METTL8 isoforms, including those without an MTS that localize to the nucleus, may enable these mRNA and caRNA sites to indeed be methylated by METTL8 (Schöller et al., 2021; Zhang et al., 2020a).

Most interestingly, we found numerous potential caRNA m^3C sites, the majority of which were on paRNA (**Figure A.4D**), including many with significantly decreased PSII-mediated mutation rates upon METTL8 knockdown (**Figure A.5D**). Notably, METTL8 was previously shown to bind to and regulate R-loops (Zhang et al., 2020a), which are enriched at promoters, and

paRNAs and R-loops at promoters modulate transcription (Dumelie and Jaffrey, 2017; Mapelli et al., 2019). Therefore, to assess whether and/or how METTL8 depletion may impact transcription, we examined the transcriptome between control and METTL8 knockdown polyA-purified input samples (**Figure A.6**). This analysis showed large-scale perturbations, with >5000 differentially expressed genes (**Figure A.6B**). It is important to note, though, that this analysis was done with only two replicates, which may limit statistical power and/or robustness of conclusions. In parallel, we examined genes closest to the caRNA m^3C sites, including downstream genes regulated by paRNA-inhabited promoters. Gene ontology assessment suggested that METTL8-deposited caRNA m^3C sites may influence transcription of nearby genes; for example, TP53-regulated genes were commonly found near caRNA m^3C sites with decreased PSII-mediated mutation rates upon METTL8 knockdown (**Figure A.5E**) and exhibited significant differential mRNA expression upon METTL8 depletion (**Figure A.6D**).

The potential relationship between METTL8-regulated caRNA m^3C sites, particularly on paRNA, and modulation of transcription, is reminiscent of some of the original studies into the function of METTL8 (Badri et al., 2008; Jakkaraju et al., 2005). In these studies, Schuger and colleagues found that METTL8 isoforms (referred to as "TIPs") comprised transcription regulators that were responsive to mechanical tension and influenced mesenchymal cell lineage determination (Jakkaraju et al., 2005). They further found that METTL8 interacted with histone acetyltransferase p300 to facilitate histone acetylation at the promoter of PPARγ2, a key adipogenic transcription factor, thereby regulating adipogenic differentiation of 3T3-L1 cells (Badri et al., 2008). At the time, it was known that METTL8 contained a SAM motif characteristic of *S*-adenosyl-L-methionine binding proteins, such as many RNA methyltransferases, but the contribution of potential methyltransfer activity to its transcription-regulation mechanism was not

investigated. Interestingly, our analyses in HeLa cells revealed an enrichment of caRNA m^3C sites near genes related to mechanical stimulus responses (**Figure A.5E**), suggesting that METTL8-mediated m^3C deposition may be related to mechanically-stimulated transcription regulation at promoters.

A potential connection between METTL8's transcription-regulation mechanism and TP53 arises in that p300, the histone acetyltransferase that binds METTL8, stimulates p53-mediated transcriptional activity on p53-regulated promoters (Avantaggiati et al., 1997). Thus, one interesting hypothesis is the possibility that p300 and METTL8 could be recruited together to p53-regulated promoters to influence transcriptional activity in a mechanism potentially involving caRNA m^3C deposition. The existence and details of such a mechanism remain unclear. Nonetheless, our findings suggest that caRNA m^3C may influence transcription control of this pathway, warranting future investigations.

Elucidating METTL8 function(s) may have important implications for cancer research beyond connections with TP53. Analyses of data in The Cancer Genome Atlas recently revealed that METTL8 is upregulated in several cancers and negatively correlated with survival in pancreatic cancer (Schöller et al., 2021). However, in low-grade gliomas, higher expression of METTL8 is associated with increased survival (Begik et al., 2020) Zhang *et al*. also found that METTL8 loss significantly reduced tumor growth in a mouse xenograft model (Zhang et al., 2020a),. These sometimes conflicting data suggest that METTL8 can influence carcinogenesis and/or progress, but since many of its functional roles and mechanisms remain unclear, it is difficult to assess relative contributions or specific mechanisms. Future studies will require dissecting its roles in mitochondrial tRNA methylation alongside its putative roles in the nucleus. More fundamentally, because the molecular and cellular locations of a modification are essential

to understanding its function, mechanistic studies of mRNA and caRNA m^3C will require the development and refinement of high-resolution mapping methods, for which this study lays important groundwork.

MATERIALS AND METHODS

RNA purification

Total RNA was isolated from HeLa cells using Trizol (Invitrogen) according to the manufacturer's instructions, and as described in Chapter 2. Polyadenylated RNA was isolated from total RNA using the Dynabeads mRNA Direct kit (Life Technologies) as per manufacturer instructions. Chromatin-associated RNA was purified as previously described by Liu *et al*. (Liu et al., 2020a). Small RNA was isolated from total RNA using the RNA Clean and Concentrator kit (Zymo) according to the manufacturer protocol for purification of small and large RNAs into separate fractions, saving the small RNA fraction.

Quantitative PCR

Quantitative PCR was performed as described in Chapter 2. The primers used to target METTL8 had the following sequences: TGCGAGTCCTTCTGGAAGAGCA (forward) and CTCAACAGCCAATTACGATCCTTG (reverse).

Mass spectrometry

LC-MS/MS was performed similarly to the description in Chapter 2, with a few key differences. For liquid chromatography, samples were separated on a HypersilGold aQ50 C18 50x3mm

column. C and m^3C nucleosides were quantified using the mass transitions 258.2>126.1 (m^3C) and 244>112 (C).

For MALDI-time-of-flight (MALDI-TOF) and HPLC, probes from AlkB demethylation reactions were purified by ethanol precipitation. Samples were then mixed with MALDI matrix, spotted and dried on a MALDI sample plate, and analyzed by a Bruker Ultraflextreme MALDI-TOF-TOF mass spectrometer in a reflector, positive mode similarly to a previous report (Hu et al., 2022).

AlkB expression, purification, and activity testing

AlkB was overexpressed using a vector from Addgene (#79050) in *E. coli* BL21 (DE3) following previously reported protocols (Zhou et al., 2019a). Cells were induced to overexpress AlkB by adding IPTG to a final concentration of 1 mM and iron (ferrous) sulfate to a final concentration of 20 μM. After shaking at 16°C overnight, the culture was pelleted and resuspended in lysis buffer (20 mM Tris-HCl pH 7.5, 300 mM NaCl, 5 mM imidazole, 1x EDTA-free protease inhibitor (Roche)). The recombinant AlkB protein was purified using a Ni-NTA super-flow column (QIAGEN). Fractions were checked for the presence and purity of AlkB by running them on a 4-12% Bis-Tris PAGE gel in 1x MOPS buffer and staining with Coomassie blue. Fractions with strong signals for AlkB at ~25 kD and low background contaminants were pooled and stored at -80°C for future use. The AlkB demethylation activity assay was performed as described by Zhang *et al*. (Zhang et al., 2021).

m3C-seq

Preparation of sequencing libraries:

For sequencing of HeLa sh-control and sh-METTL8 libraries (experiment 2), 200 ng of each RNA sample was brought to 18 µL with RNase-free water. Then 2 µL 10x fragmentation reagent (100 mM $ZnCl_2$, 10 mM Tris pH 7.5) was added to each sample, followed by incubation at 70°C for 5 minutes and placement on ice for 1 minute. Then, 0.5 M EDTA pH 8 was added to each sample, followed by purification using the Oligo Clean and Concentrator kit (Zymo) with a 1:2:8 ratio of sample to binding buffer to ethanol. 3' repair was performed by adding 2 µL of 10x PNK buffer and 1 µL T4 PNK (NEB) to 17 µL RNA sample, incubating at 37°C for 1 hour. Then, samples were purified using the RNA Clean and Concentrator kit (Zymo) using 40 µL binding buffer, 32 µL ethanol, 160 µL prep buffer, and 700 µL wash buffer, followed by elution with 7.5 µL of RNase-free water. 3' ligation was initiated by combining 6 µL of 3'-repaired RNA with 1 µL of 5.625 µM 3'-adaptor (5-'rApp-NN NNN ATC ACG AG ATC GGA AGA GCA CAC-3') and heating the mixture at 70°C for 2 minutes, then placing on ice 1 minute. 10 µL 3' ligation buffer and 3 µL enzyme mix (from the NEBNext Multiplex Small RNA Library Prep Set for Illumina) were added, and the mixture was incubated at 16°C for 16 hours. 1 µL 5' deadenylase (NEB #M0331) was added followed by incubation at 30°C for 1 hour, then 1 µL RecJF (NEB) was added followed by incubation at 30°C for 1 hour. Samples were purified with the RNA Clean and Concentrator kit (Zymo) using 44 µL binding buffer, 30.8 µL ethanol, 176 µL prep buffer, 700 µL wash buffer, and 9.5 µL RNase-free water for elution. About ⅛ of the sample was aliquoted into each of two new tubes to create two input samples. The rest was subjected to AlkB treatment, and filled up to 20 µL with RNase-free water. AlkB treatment was performed as previously reported (Zhang et al., 2021). Samples of 50 µL were then purified using the Oligo Clean and Concentrator

kit (Zymo) using 100 µL binding buffer, 400 µL ethanol, 750 µL wash buffer, and 11 µL RNase-free water for elution. Reverse transcription was then initiated by combining 9.5 µL RNA sample in water with 1 µL 1.25 µM SR RT primer (NEBNext Small RNA kit) and incubating at 65°C for 5 minutes, then placing on ice for 2 minutes. To each sample, 4 µL of 5x ProtoScript II buffer, 2 µL 0.1M DTT, 2 µL 10 µM dNTPs, 0.5 µL RNase inhibitor, and 1 µL ProtoScript II (NEB) were added and mixed, and the reaction was incubated at 42°C for 1 hour. After the reverse transcription reaction, 1 µL of RNase H buffer and 1 µL of RNase H (NEB) were added followed by incubation at 37°C for 30 minutes. Purification of samples was performed using the RNA Clean and Concentrator kit with 44 µL binding buffer, 44 µL ethanol, 176 µL prep buffer, and 700 µL wash buffer. Then, the 3' cDNA linker (5'-/Phos/NNN NNG ATC GTC GGA CTG TAG AAC TCT GAAC /3SpC3) was ligated by mixing 1 µL 50 µM linker and 1 µL DMSO with 5.5 µL cDNA sample, and incubating at 70°C for 2 minutes followed by placement on ice for a few minutes during preparation of the ligation mixture. The ligation mixture for each sample included 2 µL 10x RNA ligation buffer (NEB), 100 mM ATP, 50% PEG8000, and 1.5 µL T4 RNA ligase 1 (NEB M0437M, high concentration). Importantly, all reagents except T4 RNA ligase 1 were prepared as a master mix before aliquoting to samples, and T4 RNA ligase 1 was added directly to each sample with immediate mixing. Ligation was performed at 25 °C for 12 hours, followed by heat inactivation at 65°C for 5 minutes. Samples were purified using the RNA Clean and Concentrator kit (Zymo) by using for each 20 µL sample 40 µL binding buffer, 24 µL ethanol, 160 µL prep buffer, 700 µL wash buffer, and 21 µL RNase-free water for elution. qPCR reactions to test cycle numbers were set up with 10 µL FastStart Essential DNA Green Master Mix (Roche), 1 µL index 1 primer (NEB small RNA kit), 1 µL Universal primer (NEB), 1 µL cDNA template, and 7 µL RNase-free water. qPCR was performed as described above. PCR was performed for each sample

using the appropriate number of cycles per sample by combining 12.5 µL 2x LongAmp Taq Mix (NEB small RNA kit), 0.625 µL SR primer (NEB small RNA kit), 1.25 µL RNase-free water, 10 µL cDNA, and 0.625 µL of an index primer unique to each sample (NEB). Samples were pre-incubated at 94°C for 30s, followed by cycling for a number of cycles determined by qPCR at 94°C for 15 seconds, 62°C for 30 seconds, and 70°C for 15 seconds, and finishing with a 70°C, 5 minute final extension. Libraries were purified by AMPure bead purification (Beckman Coulter) per manufacturer instructions. Sample concentrations were measured by Qubit and quality control analysis was performed by Bioanalyzer. Libraries were sequenced on a NovaSeq6000 SP flow cell twice, once with 100bp single-end reads and again with 50bp paired end reads (due to a technical processing error at the sequencing facility).

Data analysis:

Read1 files from both sequencing experiments were merged using the Linux command cat and then quality checked using FastQC v0.11.5 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Then, analysis was performed using a pipeline constructed by modifying the framework used by Zhou *et al.* (Zhou et al., 2019a). Unique molecular identifiers (UMIs) were extracted using UMI-tools (Smith et al., 2017). Adapters were trimmed using BBDuk from BBTools version 38.84 and mapped to human genome hg38 (Frankish et al., 2019) using hisat2 v2.1.0 (Kim et al., 2019). Sorting of strands was performed using samtools view with options -F 0x10 for positive strand reads and -f 0x10 for negative strand reads, followed by sorting and indexing with the samtools sort and index commands (Li et al., 2009). Deduplication was then performed with UMI-tools (Smith et al., 2017). To count the number of reads with each base at each position, we employed bam-readcount

(https://github.com/genome/bam-readcount). To count mutation rates at each base that is C in the reference genome, apply total read cutoff and mutation cutoff filters, and clean up site lists, we used versions of the python scripts "parse.py", "control_parse.py", and "clean_untreat.py" from Zhou *et al*. modified to use C as a reference base for positive files and G for negative files (Zhou et al., 2019a). We then used R v. 4.0.3 (https://cran.r-project.org/) for downstream analyses, beginning with versions of the scripts "Prepare.R" and "Callm1A-input.R" modified for analysis of m^3C (Zhou et al., 2019a). Cutoffs for read counts, mutation rate, and demethylation rate are indicated in the text and figure captions in relevant sections. Putative m^3C sites were checked and visualized using Integrated Genomics Viewer (https://software.broadinstitute.org/software/igv/). Gene ontology analyses were performed using MetaScape (Zhou et al., 2019b). Plots were created using R v. 4.0.3 (https://cran.r-project.org/). Differential expression analysis of RNA-seq data was performed as described in Chapter 2.

REFERENCES

Agris, P.F., Eruysal, E.R., Narendran, A., Väre, V.Y.P., Vangaveti, S., and Ranganathan, S.V. (2018). Celebrating wobble decoding: Half a century and still much is new. RNA Biol. *15*, 537.

Alexandrov, A., Martzen, M.R., and Phizicky, E.M. (2002). Two proteins that form a complex are required for 7-methylguanosine modification of yeast tRNA. RNA *8*, 1253–1266.

Alquier, T., and Poitout, V. (2018). Considerations and guidelines for mouse metabolic phenotyping in diabetes research. Diabetologia *61*, 526–538.

Alseth, I., Dalhus, B., and Bjørås, M. (2014). Inosine in DNA and RNA. Curr. Opin. Genet. Dev. *26*, 116–123.

Ameismeier, M., Cheng, J., Berninghausen, O., and Beckmann, R. (2018). Visualizing late states of human 40S ribosomal subunit maturation. Nature *558*, 249–253.

Anders, M., Chelysheva, I., Goebel, I., Trenkner, T., Zhou, J., Mao, Y., Verzini, S., Qian, S.-B., and Ignatova, Z. (2018). Dynamic m6A methylation facilitates mRNA triaging to stress granules. Life Sci Alliance *1*, e201800113.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Anderson, P., and Kedersha, N. (2006). RNA granules. J. Cell Biol. *172*, 803–808.

Avantaggiati, M.L., Ogryzko, V., Gardner, K., Giordano, A., Levine, A.S., and Kelly, K. (1997). Recruitment of p300/CBP in p53-dependent signal pathways. Cell *89*, 1175–1184.

Badri, K.R., Zhou, Y., Dhru, U., Aramgam, S., and Schuger, L. (2008). Effects of the SANT domain of tension-induced/inhibited proteins (TIPs), novel partners of the histone acetyltransferase p300, on p300 activity and TIP-6-induced adipogenesis. Mol. Cell. Biol. *28*, 6358–6372.

Ballou, L.M., and Lin, R.Z. (2008). Rapamycin and mTOR kinase inhibitors. J. Chem. Biol. *1*, 27–36.

Baltz, A.G., Munschauer, M., Schwanhäusser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., et al. (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. Mol. Cell *46*, 674–690.

Barabas, A.J., Lucas, J.R., Erasmus, M.A., Cheng, H.-W., and Gaskill, B.N. (2021). Who's the Boss? Assessing Convergent Validity of Aggression Based Dominance Measures in Male Laboratory Mice, Mus Musculus. Front Vet Sci *8*, 695948.

Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millán-Zambrano, G., Robson, S.C., Aspris, D., Migliori, V., Bannister, A.J., Han, N., et al. (2017). Promoter-bound METTL3 maintains

myeloid leukaemia by m6A-dependent translation control. Nature *552*, 126–131. https://doi.org/10.1038/nature24678.

Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., et al. (2014). m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell *15*, 707–719.

Bavley, C.C., Rice, R.C., Fischer, D.K., Fakira, A.K., Byrne, M., Kosovsky, M., Rizzo, B.K., Del Prete, D., Alaedini, A., Morón, J.A., et al. (2018). Rescue of Learning and Memory Deficits in the Human Nonsyndromic Intellectual Disability Cereblon Knock-Out Mouse Model by Targeting the AMP-Activated Protein Kinase-mTORC1 Translational Pathway. J. Neurosci. *38*, 2780–2795.

Bawankar, P., Lence, T., Paolantoni, C., Haussmann, I.U., Kazlauskiene, M., Jacob, D., Heidelberger, J.B., Richter, F.M., Nallasivan, M.P., Morin, V., et al. (2021). Hakai is required for stabilization of core components of the m6A mRNA methylation machinery. Nat. Commun. *12*, 3778.

Begik, O., Lucas, M.C., Liu, H., Ramirez, J.M., Mattick, J.S., and Novoa, E.M. (2020). Integrative analyses of the RNA modification machinery reveal tissue- and cancer-specific signatures. Genome Biol. *21*, 97.

Behm-Ansmant, I., Gatfield, D., Rehwinkel, J., Hilgers, V., and Izaurralde, E. (2007). A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. EMBO J. *26*, 1591–1601.

Behrens, A., Rodschinka, G., and Nedialkova, D.D. (2021). High-resolution quantitative profiling of tRNA abundance and modification status in eukaryotes by mim-tRNAseq. Mol. Cell *81*, 1802–1815.e7.

Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J., and Nielsen, H. (2015). Profiling of ribose methylations in RNA by high-throughput sequencing. Angew. Chem. Int. Ed Engl. *54*, 451–455.

Bohnsack, K.E., Kleiber, N., Lemus-Diaz, N., and Bohnsack, M.T. (2022). Roles and dynamics of 3-methylcytidine in cellular RNAs. Trends Biochem. Sci. https://doi.org/10.1016/j.tibs.2022.03.004.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

Bowman, M., Oldridge, M., Archer, C., O'Rourke, A., McParland, J., Brekelmans, R., Seller, A., and Lester, T. (2012). Gross deletions in TCOF1 are a cause of Treacher–Collins–Franceschetti syndrome. Eur. J. Hum. Genet. *20*, 769–777.

Buchhaupt, M., Sharma, S., Kellner, S., Oswald, S., Paetzold, M., Peifer, C., Watzinger, P., Schrader, J., Helm, M., and Entian, K.-D. (2014). Partial methylation at Am100 in 18S rRNA of baker's yeast reveals ribosome heterogeneity on the level of eukaryotic rRNA modification.

PLoS One *9*, e89640.

Chang, G., Shi, L., Ye, Y., Shi, H., Zeng, L., Tiwary, S., Huse, J.T., Huo, L., Ma, L., Ma, Y., et al. (2020). YTHDF3 Induces the Translation of mA-Enriched Gene Transcripts to Promote Breast Cancer Brain Metastasis. Cancer Cell *38*, 857–871.e7.

Cheng, Y., Xie, W., Pickering, B.F., Chu, K.L., Savino, A.M., Yang, X., Luo, H., Nguyen, D.T., Mo, S., Barin, E., et al. (2021). N6-Methyladenosine on mRNA facilitates a phase-separated nuclear body that suppresses myeloid leukemic differentiation. Cancer Cell *39*, 958–972.e8.

Choe, J., Lin, S., Zhang, W., Liu, Q., Wang, L., Ramirez-Moya, J., Du, P., Kim, W., Tang, S., Sliz, P., et al. (2018). mRNA circularization by METTL3–eIF3h enhances translation and promotes oncogenesis. Nature *561*, 556–560. https://doi.org/10.1038/s41586-018-0538-8.

Choi, J., Ieong, K.-W., Demirci, H., Chen, J., Petrov, A., Prabhakar, A., O'Leary, S.E., Dominissini, D., Rechavi, G., Soltis, S.M., et al. (2016). N(6)-methyladenosine in mRNA disrupts tRNA selection and translation-elongation dynamics. Nat. Struct. Mol. Biol. *23*, 110–115.

Cooke, A., Schwarzl, T., Huppertz, I., Kramer, G., Mantas, P., Alleaume, A.-M., Huber, W., Krijgsveld, J., and Hentze, M.W. (2019). The RNA-Binding Protein YBX3 Controls Amino Acid Levels by Regulating SLC mRNA Abundance. Cell Rep. *27*, 3097–3106.e5.

Coots, R.A., Liu, X.-M., Mao, Y., Dong, L., Zhou, J., Wan, J., Zhang, X., and Qian, S.-B. (2017). mA Facilitates eIF4F-Independent mRNA Translation. Mol. Cell *68*, 504–514.e7.

Cozen, A.E., Quartley, E., Holmes, A.D., Hrabeta-Robinson, E., Phizicky, E.M., and Lowe, T.M. (2015). ARM-seq: AlkB-facilitated RNA methylation sequencing reveals a complex landscape of modified tRNA fragments. Nat. Methods *12*, 879–884.

Cui, J., Liu, Q., Sendinc, E., Shi, Y., and Gregory, R.I. (2021). Nucleotide resolution profiling of m3C RNA modification by HAC-seq. Nucleic Acids Res. *49*, e27.

Dai, X., Wang, T., Gonzalez, G., and Wang, Y. (2018). Identification of YTH Domain-Containing Proteins as the Readers for N1-Methyladenosine in RNA. Anal. Chem. *90*, 6380–6384.

Danan, C., Manickavel, S., and Hafner, M. (2016). PAR-CLIP: A Method for Transcriptome-Wide Identification of RNA Binding Protein Interaction Sites. Methods Mol. Biol. *1358*, 153–173.

Darnell, R.B. (2010). HITS-CLIP: panoramic views of protein-RNA regulation in living cells. Wiley Interdiscip. Rev. RNA *1*, 266–286.

Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y.S., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W., et al. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. Cell *146*, 247–261.

Decatur, W.A., and Fournier, M.J. (2002). rRNA modifications and ribosome function. Trends Biochem. Sci. *27*, 344–351.

Dixit, D., Prager, B.C., Gimple, R.C., Poh, H.X., Wang, Y., Wu, Q., Qiu, Z., Kidwell, R.L., Kim, L.J.Y., Xie, Q., et al. (2021). The RNA m6A Reader YTHDF2 Maintains Oncogene Expression and Is a Targetable Dependency in Glioblastoma Stem Cells. Cancer Discov. *11*, 480–499.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature *485*, 201–206.

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M.S., Dai, Q., Di Segni, A., Salmon-Divon, M., Clark, W.C., et al. (2016). The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. Nature *530*, 441–446.

Dostie, J., Ferraiuolo, M., Pause, A., Adam, S.A., and Sonenberg, N. (2000). A novel shuttling protein, 4E-T, mediates the nuclear import of the mRNA 5’ cap-binding protein, eIF4E. EMBO J. *19*, 3142–3156.

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m(6)A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. Nat. Commun. *7*, 12626.

Dumelie, J.G., and Jaffrey, S.R. (2017). Defining the location of promoter-associated R-loops at near-nucleotide resolution using bisDRIP-seq. Elife *6*. https://doi.org/10.7554/eLife.28306.

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.

Edupuganti, R.R., Geiger, S., Lindeboom, R.G.H., Shi, H., Hsu, P.J., Lu, Z., Wang, S.-Y., Baltissen, M.P.A., Jansen, P.W.T.C., Rossa, M., et al. (2017). N6-methyladenosine (m6A) recruits and repels proteins to regulate mRNA homeostasis. Nat. Struct. Mol. Biol. *24*, 870–878.

Elliott, B.A., Ho, H.-T., Ranganathan, S.V., Vangaveti, S., Ilkayeva, O., Assi, H.A., Choi, A.K., Agris, P.F., and Holley, C.L. (2019). Modification of messenger RNA by 2′-O-methylation regulates gene expression in vivo. Nature Communications *10*. https://doi.org/10.1038/s41467-019-11375-7.

Erales, J., Marchand, V., Panthu, B., Gillot, S., Belin, S., Ghayad, S.E., Garcia, M., Laforêts, F., Marcel, V., Baudin-Baillieu, A., et al. (2017). Evidence for rRNA 2’-O-methylation plasticity: Control of intrinsic translational capabilities of human ribosomes. Proc. Natl. Acad. Sci. U. S. A. *114*, 12934–12939.

Esteve-Puig, R., Bueno-Costa, A., and Esteller, M. (2020). Writers, readers and erasers of RNA modifications in cancer. Cancer Lett. *474*, 127–137.

Evdokimova, V., Ruzanov, P., Imataka, H., Raught, B., Svitkin, Y., Ovchinnikov, L.P., and Sonenberg, N. (2001). The major mRNA-associated protein YB-1 is a potent 5' cap-dependent mRNA stabilizer. EMBO J. *20*, 5491–5502.

Farley-Barnes, K.I., Ogawa, L.M., and Baserga, S.J. (2019). Ribosomopathies: Old Concepts, New Controversies. Trends Genet. *35*, 754–767.

Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. *47*, D766–D773.

Freude, K., Hoffmann, K., Jensen, L.-R., Delatycki, M.B., des Portes, V., Moser, B., Hamel, B., van Bokhoven, H., Moraine, C., Fryns, J.-P., et al. (2004). Mutations in the FTSJ1 gene coding for a novel S-adenosylmethionine-binding protein cause nonsyndromic X-linked mental retardation. Am. J. Hum. Genet. *75*, 305–309.

Frye, M., Harada, B.T., Behm, M., and He, C. (2018). RNA modifications modulate gene expression during development. Science *361*, 1346–1349.

Fu, Y., and Zhuang, X. (2020). m6A-binding YTHDF proteins promote stress granule formation. Nat. Chem. Biol. *16*, 955–963.

Fu, D., Brophy, J.A.N., Chan, C.T.Y., 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.

Fuhrmann, D.C., and Brüne, B. (2017). Mitochondrial composition and function under the control of hypoxia. Redox Biol *12*, 208–215.

Genuth, N.R., and Barna, M. (2018). The Discovery of Ribosome Heterogeneity and Its Implications for Gene Regulation and Organismal Life. Mol. Cell *71*, 364–374.

Go, C.D., Knight, J.D.R., Rajasekharan, A., Rathod, B., Hesketh, G.G., Abe, K.T., Youn, J.-Y., Samavarchi-Tehrani, P., Zhang, H., Zhu, L.Y., et al. (2021). A proximity-dependent biotinylation map of a human cell. Nature *595*, 120–124.

Gu, H., Do, D.V., Liu, X., Xu, L., Su, Y., Nah, J.M., Wong, Y., Li, Y., Sheng, N., Tilaye, G.A., et al. (2018). The STAT3 Target Mettl8 Regulates Mouse ESC Differentiation via Inhibiting the JNK Pathway. Stem Cell Reports *10*, 1807–1820.

Haimovich, G., Choder, M., Singer, R.H., and Trcek, T. (2013). The fate of the messenger is pre-determined: a new model for regulation of gene expression. Biochim. Biophys. Acta *1829*, 643–653.

Han, J., Kim, D., and Morris, K.V. (2007). Promoter-associated RNA is required for RNA-directed transcriptional gene silencing in human cells. Proc. Natl. Acad. Sci. U. S. A. *104*, 12422–12427.

Hao, Z., Wu, T., Cui, X., Zhu, P., Tan, C., Dou, X., Hsu, K.-W., Lin, Y.-T., Peng, P.-H., Zhang, L.-S., et al. (2020). N6-Deoxyadenosine Methylation in Mammalian Mitochondrial DNA. Mol. Cell *78*, 382–395.e8.

He, P.C., Cody He, P., and He, C. (2021). m ^{6} A RNA methylation: from mechanisms to therapeutic potential. The EMBO Journal *40*. https://doi.org/10.15252/embj.2020105977.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell *38*, 576–589.

Hsu, P.J., Shi, H., and He, C. (2017). Epitranscriptomic influences on development and disease. Genome Biol. *18*, 1–9.

Hsu, P.J., Shi, H., Zhu, A.C., Lu, Z., Miller, N., Edens, B.M., Ma, Y.C., and He, C. (2019). The RNA-binding protein FMRP facilitates the nuclear export of N 6-methyladenosine-containing mRNAs. J. Biol. Chem. *294*, 19889–19895.

Hu, H., Kahrizi, K., Musante, L., Fattahi, Z., Herwig, R., Hosseini, M., Oppitz, C., Abedini, S.S., Suckow, V., Larti, F., et al. (2019). Genetics of intellectual disability in consanguineous families. Mol. Psychiatry *24*, 1027–1039.

Hu, L., Liu, S., Peng, Y., Ge, R., Su, R., Senevirathne, C., Harada, B.T., Dai, Q., Wei, J., Zhang, L., et al. (2022). mA RNA modifications are measured at single-base resolution across the mammalian transcriptome. Nat. Biotechnol. https://doi.org/10.1038/s41587-022-01243-z.

Huang, D.W., Sherman, B.T., Tan, Q., Collins, J.R., Alvord, W.G., Roayaei, J., Stephens, R., Baseler, M.W., Lane, H.C., and Lempicki, R.A. (2007). The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. Genome Biol. *8*, R183.

Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B.S., Mesquita, A., Liu, C., Yuan, C.L., et al. (2018). Recognition of RNA N6-methyladenosine by IGF2BP Proteins Enhances mRNA Stability and Translation. Nat. Cell Biol. *20*, 285.

Huang, H., Weng, H., Zhou, K., Wu, T., Zhao, B.S., Sun, M., Chen, Z., Deng, X., Xiao, G., Auer, F., et al. (2019). Histone H3 trimethylation at lysine 36 guides mA RNA modification co-transcriptionally. Nature *567*, 414–419.

Hubstenberger, A., Courel, M., Bénard, M., Souquere, S., Ernoult-Lange, M., Chouaib, R., Yi, Z., Morlot, J.-B., Munier, A., Fradet, M., et al. (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. Mol. Cell *68*, 144–157.e5.

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. (2020a). The rRNA mA methyltransferase METTL5 is involved in pluripotency and developmental programs. Genes Dev. *34*, 715–729.

Ignatova, V.V., Kaiser, S., Ho, J.S.Y., Bing, X., Stolz, P., Tan, Y.X., Lee, C.L., Gay, F.P.H., Lastres, P.R., Gerlini, R., et al. (2020b). METTL6 is a tRNA m3C methyltransferase that regulates pluripotency and tumor cell growth. Sci Adv *6*, eaaz4551.

Jakkaraju, S., Zhe, X., Pan, D., Choudhury, R., and Schuger, L. (2005). TIPs are tension-responsive proteins involved in myogenic versus adipogenic differentiation. Dev. Cell *9*, 39–49.

Jeong, H., Lee, Y., and Kim, J. (2022). Structural and functional characterization of TrmM in m A modification of bacterial tRNA. Protein Sci. *31*, e4319.

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.-G., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat. Chem. Biol. *7*, 885–887.

Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. *37*, 907–915.

Kirkpatrick, L.L., McIlwain, K.A., and Nelson, D.L. (2001). Comparative Genomic Sequence Analysis of the FXR Gene Family: FMR1, FXR1, and FXR2. Genomics *78*, 169–177. https://doi.org/10.1006/geno.2001.6667.

Kleiber, N., Lemus-Diaz, N., Stiller, C., Heinrichs, M., Mai, M.M.-Q., Hackert, P., Richter-Dennerlein, R., Höbartner, C., Bohnsack, K.E., and Bohnsack, M.T. (2022). The RNA methyltransferase METTL8 installs m3C32 in mitochondrial tRNAsThr/Ser(UCN) to optimise tRNA structure and mitochondrial translation. Nat. Commun. *13*, 209.

Kusevic, D., Kudithipudi, S., and Jeltsch, A. (2016). Substrate Specificity of the HEMK2 Protein Glutamine Methyltransferase and Identification of Novel Substrates. J. Biol. Chem. *291*, 6124–6133.

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

Lasman, L., Krupalnik, V., Viukov, S., Mor, N., Aguilera-Castrejon, A., Schneir, D., Bayerl, J., Mizrahi, O., Peles, S., Tawil, S., et al. (2020). Context-dependent functional compensation between Ythdf m6A reader proteins. Genes Dev. *34*, 1373–1391.

Lee, Y., Choe, J., Park, O.H., and Kim, Y.K. (2020). Molecular Mechanisms Driving mRNA Degradation by m6A Modification. Trends Genet. *36*, 177–188.

Leetsi, L., Õunap, K., Abroi, A., and Kurg, R. (2019). The Common Partner of Several Methyltransferases TRMT112 Regulates the Expression of N6AMT1 Isoforms in Mammalian Cells. Biomolecules *9*. https://doi.org/10.3390/biom9090422.

Legrand, C., Tuorto, F., Hartmann, M., Liebers, R., Jacob, D., Helm, M., and Lyko, F. (2017). Statistically robust methylation calling for whole-transcriptome bisulfite sequencing reveals distinct methylation patterns for mouse RNAs. Genome Res. *27*, 1589–1596.

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 A methyltransferase Mettl5 is required for normal walking behavior in Drosophila. EMBO Rep. *21*, e49443.

Lentini, J.M., Alsaif, H.S., Faqeih, E., Alkuraya, F.S., and Fu, D. (2020). DALRD3 encodes a protein mutated in epileptic encephalopathy that targets arginine tRNAs for 3-methylcytosine modification. Nat. Commun. *11*, 2510.

Lentini, J.M., Bargabos, R., Chen, C., and Fu, D. (2022). Methyltransferase METTL8 is required for 3-methylcytosine modification in human mitochondrial tRNAs. J. Biol. Chem. *298*, 101788.

Leppek, K., Das, R., and Barna, M. (2018). Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. Nat. Rev. Mol. Cell Biol. *19*, 158–174.

Lesbirel, S., and Wilson, S.A. (2019). The m6A-methylase complex and mRNA export. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms *1862*, 319–328. https://doi.org/10.1016/j.bbagrm.2018.09.008.

Li, F., Zhao, D., Wu, J., and Shi, Y. (2014). Structure of the YTH domain of human YTHDF2 in complex with an m6A mononucleotide reveals an aromatic cage for m6A recognition. Cell Research *24*, 1490–1492. https://doi.org/10.1038/cr.2014.153.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics *25*, 2078–2079.

Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., Mao, Y., Lv, J., Yi, D., Chen, X.-W., et al. Base-resolution mapping reveals distinct m^1A methylome in nuclear- and mitochondrial-encoded transcripts. https://doi.org/10.1101/202747.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics *30*, 923–930.

Liberman, N., O'Brown, Z.K., Earl, A.S., Boulias, K., Gerashchenko, M.V., Wang, S.Y., Fritsche, C., Fady, P.-E., Dong, A., Gladyshev, V.N., et al. (2020). N6-adenosine methylation of ribosomal RNA affects lipid oxidation and stress resistance. Sci Adv *6*, eaaz4370.

Lin, S., Choe, J., Du, P., Triboulet, R., and Gregory, R.I. (2016). The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. Mol. Cell *62*, 335–345.

Linder, B., Grozhik, A.V., Olarerin-George, A.O., Meydan, C., Mason, C.E., and Jaffrey, S.R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat. Methods *12*, 767–772.

Liu, B., Merriman, D.K., Choi, S.H., Schumacher, M.A., Plangger, R., Kreutz, C., Horner, S.M., Meyer, K.D., and Al-Hashimi, H.M. (2018a). A potentially abundant junctional RNA motif stabilized by m6A and Mg2. Nat. Commun. *9*, 2761.

Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat. Chem. Biol. *10*, 93–95.

Liu, J., Eckert, M.A., Harada, B.T., Liu, S.-M., Lu, Z., Yu, K., Tienda, S.M., Chryplewicz, A., Zhu, A.C., Yang, Y., et al. (2018b). mA mRNA methylation regulates AKT activity to promote the proliferation and tumorigenicity of endometrial cancer. Nat. Cell Biol. *20*, 1074–1083.

Liu, J., Dou, X., Chen, C., Chen, C., Liu, C., Xu, M.M., Zhao, S., Shen, B., Gao, Y., Han, D., et al. (2020a). -methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. Science *367*, 580–586.

Liu, J., Huang, T., Chen, W., Ding, C., Zhao, T., Zhao, X., Cai, B., Zhang, Y., Li, S., Zhang, L., et al. (2022). Developmental mRNA m5C landscape and regulatory innovations of massive m5C modification of maternal mRNAs in animals. Nat. Commun. *13*, 2484.

Liu, K., Santos, D.A., Hussmann, J.A., Wang, Y., Sutter, B.M., Weissman, J.S., and Tu, B.P. (2021). Regulation of translation by methylation multiplicity of 18S rRNA. Cell Rep. *34*, 108825.

Liu, L., Zhang, S.-W., Huang, Y., and Meng, J. (2017a). QNB: differential RNA methylation analysis for count-based small-sample sequencing data with a quad-negative binomial model. BMC Bioinformatics *18*, 387.

Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N6-methyladenosine-dependent RNA structural switches regulate RNA–protein interactions. Nature *518*, 560–564. https://doi.org/10.1038/nature14234.

Liu, N., Zhou, K.I., Parisien, M., Dai, Q., Diatchenko, L., and Pan, T. (2017b). N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. Nucleic Acids Res. *45*, 6051–6063.

Liu, Q., Shvarts, T., Sliz, P., and Gregory, R.I. (2020b). RiboToolkit: an integrated platform for analysis and annotation of ribosome profiling data to decode mRNA translation at codon resolution. Nucleic Acids Res. *48*, W218–W229.

Lomakin, I.B., Stolboushkina, E.A., Vaidya, A.T., Zhao, C., Garber, M.B., Dmitriev, S.E., and Steitz, T.A. (2017). Crystal Structure of the Human Ribosome in Complex with DENR-MCT-1. Cell Rep. *20*, 521–528.

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.

Luo, Y., Na, Z., and Slavoff, S.A. (2018). P-Bodies: Composition, Properties, and Functions.

Biochemistry *57*, 2424–2431.

Ma, C.-J., Ding, J.-H., Ye, T.-T., Yuan, B.-F., and Feng, Y.-Q. (2019a). AlkB Homologue 1 Demethylates N3-Methylcytidine in mRNA of Mammals. ACS Chem. Biol. *14*, 1418–1425.

Ma, H., Wang, X., Cai, J., Dai, Q., Natchiar, S.K., Lv, R., Chen, K., Lu, Z., Chen, H., Shi, Y.G., et al. (2019b). Publisher Correction: N-Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. Nat. Chem. Biol. *15*, 549.

Ma, H., Wang, X., Cai, J., Dai, Q., Natchiar, S.K., Lv, R., Chen, K., Lu, Z., Chen, H., Shi, Y.G., et al. (2019c). N6-Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. Nat. Chem. Biol. *15*, 88–94.

Mao, S., Haruehanroengra, P., Ranganathan, S.V., Shen, F., Begley, T.J., and Sheng, J. (2021). Base Pairing and Functional Insights into N3-Methylcytidine (m3C) in RNA. ACS Chem. Biol. *16*, 76–85.

Mao, Y., Dong, L., Liu, X.-M., Guo, J., Ma, H., Shen, B., and Qian, S.-B. (2019). m6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. Nat. Commun. *10*, 5332.

Mapelli, S.N., Napoli, S., Pisignano, G., Garcia-Escudero, R., Carbone, G.M., and Catapano, C.V. (2019). Deciphering the complexity of human non-coding promoter-proximal transcriptome. Bioinformatics *35*, 2529–2534.

Marchand, V., Bourguignon-Igel, V., Helm, M., and Motorin, Y. (2021). Mapping of 7-methylguanosine (mG), 3-methylcytidine (mC), dihydrouridine (D) and 5-hydroxycytidine (hoC) RNA modifications by AlkAniline-Seq. Methods Enzymol. *658*, 25–47.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal *17*, 10–12.

Mathlin, J., Le Pera, L., and Colombo, T. (2020). A Census and Categorization Method of Epitranscriptomic Marks. Int. J. Mol. Sci. *21*. https://doi.org/10.3390/ijms21134684.

Melanson, G., Timpano, S., and Uniacke, J. (2017). The eIF4E2-Directed Hypoxic Cap-Dependent Translation Machinery Reveals Novel Therapeutic Potential for Cancer Treatment. Oxid. Med. Cell. Longev. *2017*, 6098107.

Meyer, K.D. (2019). mA-mediated translation regulation. Biochim. Biophys. Acta Gene Regul. Mech. *1862*, 301–309.

Meyer, K.D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C.E., and Jaffrey, S.R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3’ UTRs and near stop codons. Cell *149*, 1635–1646.

Meyer, K.D., Patil, D.P., Zhou, J., Zinoviev, A., Skabkin, M.A., Elemento, O., Pestova, T.V., Qian, S.-B., and Jaffrey, S.R. (2015). 5’ UTR m(6)A Promotes Cap-Independent Translation.

Cell *163*, 999–1010.

Natchiar, S.K., Myasnikov, A.G., Kratzat, H., Hazemann, I., and Klaholz, B.P. (2017). Visualization of chemical modifications in the human 80S ribosome structure. Nature *551*, 472–477.

Nicholson, A.L., and Pasquinelli, A.E. (2019). Tales of Detailed Poly(A) Tails. Trends Cell Biol. *29*, 191–200.

Olivier, M., Hollstein, M., and Hainaut, P. (2010). TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb. Perspect. Biol. *2*, a001008.

Park, O.H., Ha, H., Lee, Y., Boo, S.H., Kwon, D.H., Song, H.K., and Kim, Y.K. (2019). Endoribonucleolytic Cleavage of m6A-Containing RNAs by RNase P/MRP Complex. Mol. Cell *74*, 494–507.e8.

de la Parra, C., Ernlund, A., Alard, A., Ruggles, K., Ueberheide, B., and Schneider, R.J. (2018). A widespread alternate form of cap-dependent mRNA translation initiation. Nat. Commun. *9*, 3068.

Patil, D.P., Chen, C.-K., Pickering, B.F., Chow, A., Jackson, C., Guttman, M., and Jaffrey, S.R. (2016). m(6)A RNA methylation promotes XIST-mediated transcriptional repression. Nature *537*, 369–373.

Pazo, A., Pérez-González, A., Oliveros, J.C., Huarte, M., Chavez, J.P., and Nieto, A. (2019). hCLE/RTRAF-HSPC117-DDX1-FAM98B: A New Cap-Binding Complex That Activates mRNA Translation. Front. Physiol. *10*, 92.

Pendleton, K.E., Chen, B., Liu, K., Hunter, O.V., Xie, Y., Tu, B.P., and Conrad, N.K. (2017). The U6 snRNA mA Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. Cell *169*, 824–835.e14.

Picard-Jean, F., Brand, C., Tremblay-Létourneau, M., Allaire, A., Beaudoin, M.C., Boudreault, S., Duval, C., Rainville-Sirois, J., Robert, F., Pelletier, J., et al. (2018). 2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. PLoS One *13*, e0193804.

Piekna-Przybylska, D., Decatur, W.A., and Fournier, M.J. (2008). The 3D rRNA modification maps database: with interactive tools for ribosome analysis. Nucleic Acids Res. *36*, D178–D183.

Ping, X.-L., Sun, B.-F., Wang, L., Xiao, W., Yang, X., Wang, W.-J., Adhikari, S., Shi, Y., Lv, Y., Chen, Y.-S., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. *24*, 177–189.

Pinto, R., Vågbø, C.B., Jakobsson, M.E., Kim, Y., Baltissen, M.P., O'Donohue, M.-F., Guzmán, U.H., Małecki, J.M., Wu, J., Kirpekar, F., et al. (2020). The human methyltransferase ZCCHC4 catalyses N6-methyladenosine modification of 28S ribosomal RNA. Nucleic Acids Res. *48*, 830–846.

Potapov, V., Fu, X., Dai, N., Corrêa, I.R., Jr, Tanner, N.A., and Ong, J.L. (2018). Base modifications affecting RNA polymerase and reverse transcriptase fidelity. Nucleic Acids Res. *46*, 5753–5763.

Qu, J., Yan, H., Hou, Y., Cao, W., Liu, Y., Zhang, E., He, J., and Cai, Z. (2022). RNA demethylase ALKBH5 in cancer: from mechanisms to therapeutic potential. J. Hematol. Oncol. *15*, 8.

Ramanathan, A., Robb, G.B., and Chan, S.-H. (2016). mRNA capping: biological functions and applications. Nucleic Acids Res. *44*, 7511–7526.

Ramser, J., Winnepenninckx, B., Lenski, C., Errijgers, V., Platzer, M., Schwartz, C.E., Meindl, A., and Kooy, R.F. (2004). A splice site mutation in the methyltransferase gene FTSJ1 in Xp11.23 is associated with non-syndromic mental retardation in a large Belgian family (MRX9). J. Med. Genet. *41*, 679–683.

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.

Räsch, F., Weber, R., Izaurralde, E., and Igreja, C. (2020). 4E-T-bound mRNAs are stored in a silenced and deadenylated form. Genes Dev. *34*, 847–860.

Rauch, S., He, C., and Dickinson, B.C. (2018). Targeted mA Reader Proteins To Study Epitranscriptomic Regulation of Single RNAs. J. Am. Chem. Soc. *140*, 11974–11981.

Richard, E.M., Polla, D.L., Assir, M.Z., Contreras, M., Shahzad, M., Khan, A.A., Razzaq, A., Akram, J., Tarar, M.N., Blanpied, T.A., et al. (2019). Bi-allelic Variants in METTL5 Cause Autosomal-Recessive Intellectual Disability and Microcephaly. Am. J. Hum. Genet. *105*, 869–878.

Ries, R.J., Zaccara, S., Klein, P., Olarerin-George, A., Namkoong, S., Pickering, B.F., Patil, D.P., Kwak, H., Lee, J.H., and Jaffrey, S.R. (2019). m6A enhances the phase separation potential of mRNA. Nature *571*, 424–428.

Rintala-Dempsey, A.C., and Kothe, U. (2017). Eukaryotic stand-alone pseudouridine synthases - RNA modifying enzymes and emerging regulators of gene expression? RNA Biol. *14*, 1185–1196.

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 m6A Methyltransferase METTL5 Promotes Translation Initiation and Breast Cancer Cell Growth. Cell Rep. *33*, 108544.

Rosa, B.A., Ahmed, M., Singh, D.K., Choreño-Parra, J.A., Cole, J., Jiménez-Álvarez, L.A., Rodríguez-Reyna, T.S., Singh, B., Gonzalez, O., Carrion, R., Jr, et al. (2021). IFN signaling and neutrophil degranulation transcriptional signatures are induced during SARS-CoV-2 infection. Commun Biol *4*, 290.

Roundtree, I.A., Evans, M.E., Pan, T., and He, C. (2017a). Dynamic RNA Modifications in Gene

Expression Regulation. Cell *169*, 1187–1200.

Roundtree, I.A., Luo, G.-Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., Sha, J., Huang, X., Guerrero, L., Xie, P., et al. (2017b). YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. Elife *6*. https://doi.org/10.7554/eLife.31311.

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.

Safran, M., Solomon, I., Shmueli, O., Lapidot, M., Shen-Orr, S., Adato, A., Ben-Dor, U., Esterman, N., Rosen, N., Peter, I., et al. (2002). GeneCards 2002: towards a complete, object-oriented, human gene compendium. Bioinformatics *18*, 1542–1543.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nature Methods *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

Schöller, E., Marks, J., Marchand, V., Bruckmann, A., Powell, C.A., Reichold, M., Mutti, C.D., Dettmer, K., Feederle, R., Hüttelmaier, S., et al. (2021). Balancing of mitochondrial translation through METTL8-mediated m3C modification of mitochondrial tRNAs. Mol. Cell *81*, 4810–4825.e12.

Schrodinger (2010). The PyMOL molecular graphics system. Version.

Schwartz, S., Bernstein, D.A., Mumbach, M.R., Jovanovic, M., Herbst, R.H., León-Ricardo, B.X., Engreitz, J.M., Guttman, M., Satija, R., Lander, E.S., et al. (2014). Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell *159*, 148–162.

Schwenk, F., Baron, U., and Rajewsky, K. (1995). A cre-transgenic mouse strain for the ubiquitous deletion of loxP-flanked gene segments including deletion in germ cells. Nucleic Acids Res. *23*, 5080–5081.

Seimetz, J., Arif, W., Bangru, S., Hernaez, M., and Kalsotra, A. (2019). Cell-type specific polysome profiling from mammalian tissues. Methods *155*, 131–139.

Sendinc, E., Valle-Garcia, D., Jiao, A., and Shi, Y. (2020). Analysis of m6A RNA methylation in Caenorhabditis elegans. Cell Discov *6*, 47.

Sepich-Poore, C., Zheng, Z., Schmitt, E., Wen, K., Zhang, Z.S., Cui, X.-L., Dai, Q., Zhu, A.C., Zhang, L., Sanchez Castillo, A., et al. (2022). The METTL5-TRMT112 N-methyladenosine methyltransferase complex regulates mRNA translation via 18S rRNA methylation. J. Biol. Chem. *298*, 101590.

Sergiev, P.V., Aleksashin, N.A., Chugunova, A.A., Polikanov, Y.S., and Dontsova, O.A. (2018). Structural and evolutionary insights into ribosomal RNA methylation. Nat. Chem. Biol. *14*, 226–235.

Shi, H., Wang, X., Lu, Z., Zhao, B.S., Ma, H., Hsu, P.J., Liu, C., and He, C. (2017a). YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. Cell Res. *27*, 315–328.

Shi, H., Wei, J., and He, C. (2019). Where, When, and How: Context-Dependent Functions of RNA Methylation Writers, Readers, and Erasers. Mol. Cell *74*, 640–650.

Shi, Z., Fujii, K., Kovary, K.M., Genuth, N.R., Röst, H.L., Teruel, M.N., and Barna, M. (2017b). Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide. Mol. Cell *67*, 71–83.e7.

Shima, H., Matsumoto, M., Ishigami, Y., Ebina, M., Muto, A., Sato, Y., Kumagai, S., Ochiai, K., Suzuki, T., and Igarashi, K. (2017). S-Adenosylmethionine Synthesis Is Regulated by Selective N6-Adenosine Methylation and mRNA Degradation Involving METTL16 and YTHDC1. Cell Reports *21*, 3354–3363. https://doi.org/10.1016/j.celrep.2017.11.092.

Shimazu, T., Barjau, J., Sohtome, Y., Sodeoka, M., and Shinkai, Y. (2014). Selenium-based S-adenosylmethionine analog reveals the mammalian seven-beta-strand methyltransferase METTL10 to be an EF1A1 lysine methyltransferase. PLoS One *9*, e105394.

Silverman, J.L., Yang, M., Lord, C., and Crawley, J.N. (2010). Behavioural phenotyping assays for mouse models of autism. Nat. Rev. Neurosci. *11*, 490–502.

Siveen, K.S., Sikka, S., Surana, R., Dai, X., Zhang, J., Kumar, A.P., Tan, B.K.H., Sethi, G., and Bishayee, A. (2014). Targeting the STAT3 signaling pathway in cancer: role of synthetic and natural inhibitors. Biochim. Biophys. Acta *1845*, 136–154.

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.

Slobodin, B., Han, R., Calderone, V., Vrielink, J.A.F.O., Loayza-Puch, F., Elkon, R., and Agami, R. (2017). Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-adenosine Methylation. Cell *169*, 326–337.e12.

Smith, T., Heger, A., and Sudbery, I. (2017). UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. Genome Res. *27*, 491–499.

Speakman, J.R. (2013). Measuring energy metabolism in the mouse - theoretical, practical, and analytical considerations. Front. Physiol. *4*, 34.

Splendore, A., Silva, E.O., Alonso, L.G., Richieri-Costa, A., Alonso, N., Rosa, A., Carakushanky, G., Cavalcanti, D.P., Brunoni, D., and Passos-Bueno, M.R. (2000). High mutation detection rate in TCOF1 among Treacher Collins syndrome patients reveals clustering of mutations and 16 novel pathogenic changes. Hum. Mutat. *16*, 315–322.

Su, R., Ma, J., Zheng, J., Liu, X., Liu, Y., Ruan, X., Shen, S., Yang, C., Wang, D., Cai, H., et al. (2020). PABPC1-induced stabilization of BDNF-AS inhibits malignant progression of glioblastoma cells through STAU1-mediated decay. Cell Death Dis. *11*, 81.

Sun, F., Palmer, K., and Handel, M.A. (2010). Mutation of Eif4g3, encoding a eukaryotic translation initiation factor, causes male infertility and meiotic arrest of mouse spermatocytes. Development *137*, 1699–1707.

Sun, S., Fei, K., Zhang, G., Wang, J., Yang, Y., Guo, W., Yang, Z., Wang, J., Xue, Q., Gao, Y., et al. (2020). Construction and Comprehensive Analyses of a METTL5-Associated Prognostic Signature With Immune Implication in Lung Adenocarcinomas. Front. Genet. *11*, 617174.

Tahmasebi, S., Amiri, M., and Sonenberg, N. (2018). Translational Control in Stem Cells. Front. Genet. *9*, 709.

Taoka, M., Nobe, Y., Yamaki, Y., Yamauchi, Y., Ishikawa, H., Takahashi, N., Nakayama, H., and Isobe, T. (2016). The complete chemical structure of Saccharomyces cerevisiae rRNA: partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. Nucleic Acids Res. *44*, 8951–8961.

Teo, G., Liu, G., Zhang, J., Nesvizhskii, A.I., Gingras, A.-C., and Choi, H. (2014). SAINTexpress: improvements and additional features in Significance Analysis of INTeractome software. J. Proteomics *100*, 37–43.

Tooley, C.E.S., Petkowski, J.J., Muratore-Schroeder, T.L., Balsbaugh, J.L., Shabanowitz, J., Sabat, M., Minor, W., Hunt, D.F., and Macara, I.G. (2010). NRMT is an alpha-N-methyltransferase that methylates RCC1 and retinoblastoma protein. Nature *466*, 1125–1128.

Topisirovic, I., Svitkin, Y.V., Sonenberg, N., and Shatkin, A.J. (2011). Cap and cap-binding proteins in the control of gene expression. Wiley Interdiscip. Rev. RNA *2*, 277–298.

Torabi, S.-F., Vaidya, A.T., Tycowski, K.T., DeGregorio, S.J., Wang, J., Shu, M.-D., Steitz, T.A., and Steitz, J.A. (2021). RNA stabilization by a poly(A) tail 3'-end binding pocket and other modes of poly(A)-RNA interaction. Science *371*. https://doi.org/10.1126/science.abe6523.

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.

Tsao, N., Brickner, J.R., Rodell, R., Ganguly, A., Wood, M., Oyeniran, C., Ahmad, T., Sun, H., Bacolla, A., Zhang, L., et al. (2021). Aberrant RNA methylation triggers recruitment of an alkylation repair complex. Mol. Cell *81*, 4228–4242.e8.

Ueda, Y., Ooshio, I., Fusamae, Y., Kitae, K., Kawaguchi, M., Jingushi, K., Hase, H., Harada, K., Hirata, K., and Tsujikawa, K. (2017). AlkB homolog 3-mediated tRNA demethylation promotes protein synthesis in cancer cells. Sci. Rep. *7*, 42271.

Uniacke, J., Holterman, C.E., Lachance, G., Franovic, A., Jacob, M.D., Fabian, M.R., Payette, J., Holcik, M., Pause, A., and Lee, S. (2012). An oxygen-regulated switch in the protein synthesis machinery. Nature *486*, 126–129.

Uttam, S., Wong, C., Price, T.J., and Khoutorsky, A. (2018). eIF4E-Dependent Translational

Control: A Central Mechanism for Regulation of Pain Plasticity. Front. Genet. *9*, 470.

Walkley, C.R., and Li, J.B. (2017). Rewriting the transcriptome: adenosine-to-inosine RNA editing by ADARs. Genome Biol. *18*, 205.

Wang, C., Schmich, F., Srivatsa, S., Weidner, J., Beerenwinkel, N., and Spang, A. (2018). Context-dependent deposition and regulation of mRNAs in P-bodies. Elife *7*. https://doi.org/10.7554/eLife.29815.

Wang, L., Liang, Y., Lin, R., Xiong, Q., Yu, P., Ma, J., Cheng, M., Han, H., Wang, X., Wang, G., et al. (2022). Mettl5 mediated 18S rRNA N6-methyladenosine (m6A) modification controls stem cell fate determination and neural function. Genes Dis. *9*, 268–274.

Wang, P., Doxtader, K.A., and Nam, Y. (2016). Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. Mol. Cell *63*, 306–317.

Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., et al. (2014a). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature *505*, 117–120.

Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell *161*, 1388–1399.

Wang, Y., Li, Y., Toth, J.I., Petroski, M.D., Zhang, Z., and Zhao, J.C. (2014b). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nat. Cell Biol. *16*, 191–198.

Wang, Y., Zhang, Z., Sepich-Poore, C., Zhang, L., Xiao, Y., and He, C. (2021a). LEAD-m6A-seq for Locus-Specific Detection of N6-Methyladenosine and Quantification of Differential Methylation. Angew. Chem. Int. Ed. *60*, 873–880.

Wang, Y.-J., Yang, B., Lai, Q., Shi, J.-F., Peng, J.-Y., Zhang, Y., Hu, K.-S., Li, Y.-Q., Peng, J.-W., Yang, Z.-Z., et al. (2021b). Reprogramming of m6A epitranscriptome is crucial for shaping of transcriptome and proteome in response to hypoxia. RNA Biol. *18*, 131–143.

Warda, A.S., Kretschmer, J., Hackert, P., Lenz, C., Urlaub, H., Höbartner, C., Sloan, K.E., and Bohnsack, M.T. (2017). Human METTL16 is a N6-methyladenosine (m6A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep. *18*, 2004–2014.

Weber, R., Chung, M.-Y., Keskeny, C., Zinnall, U., Landthaler, M., Valkov, E., Izaurralde, E., and Igreja, C. (2020). 4EHP and GIGYF1/2 Mediate Translation-Coupled Messenger RNA Decay. Cell Rep. *33*, 108262.

Wei, J., Liu, F., Lu, Z., Fei, Q., Ai, Y., Cody He, P., Shi, H., Cui, X., Su, R., Klungland, A., et al. (2018). Differential m6A, m6Am, and m1A Demethylation Mediated by FTO in the Cell Nucleus and Cytoplasm. Molecular Cell *71*, 973–985.e5. https://doi.org/10.1016/j.molcel.2018.08.011.

Wei, J., Yu, X., Yang, L., Liu, X., Gao, B., Huang, B., Dou, X., Liu, J., Zou, Z., Cui, X.-L., et al. (2022). FTO mediates LINE1 m ^{6} A demethylation and chromatin regulation in mESCs and mouse development. Science *376*, 968–973. https://doi.org/10.1126/science.abe9582.

Weisser, M., Voigts-Hoffmann, F., Rabl, J., Leibundgut, M., and Ban, N. (2013). The crystal structure of the eukaryotic 40S ribosomal subunit in complex with eIF1 and eIF1A. Nat. Struct. Mol. Biol. *20*, 1015–1017.

Weisser, M., Schäfer, T., Leibundgut, M., Böhringer, D., Aylett, C.H.S., and Ban, N. (2017). Structural and Functional Insights into Human Re-initiation Complexes. Mol. Cell *67*, 447–456.e7.

Wen, J., Lv, R., Ma, H., Shen, H., He, C., Wang, J., Jiao, F., Liu, H., Yang, P., Tan, L., et al. (2018). Zc3h13 Regulates Nuclear RNA mA Methylation and Mouse Embryonic Stem Cell Self-Renewal. Mol. Cell *69*, 1028–1038.e6.

Werner, M.S., and Ruthenburg, A.J. (2015). Nuclear Fractionation Reveals Thousands of Chromatin-Tethered Noncoding RNAs Adjacent to Active Genes. Cell Rep. *12*, 1089–1098.

Wilkinson, E., Cui, Y.-H., and He, Y.-Y. (2021). Context-Dependent Roles of RNA Modifications in Stress Responses and Diseases. Int. J. Mol. Sci. *22*. https://doi.org/10.3390/ijms22041949.

Wollen, K.L., Hagen, L., Vågbø, C.B., Rabe, R., Iveland, T.S., Aas, P.A., Sharma, A., Sporsheim, B., Erlandsen, H.O., Palibrk, V., et al. (2021). ALKBH3 partner ASCC3 mediates P-body formation and selective clearance of MMS-induced 1-methyladenosine and 3-methylcytosine from mRNA. J. Transl. Med. *19*, 287.

Wu, R., Li, A., Sun, B., Sun, J.-G., Zhang, J., Zhang, T., Chen, Y., Xiao, Y., Gao, Y., Zhang, Q., et al. (2019). A novel mA reader Prrc2a controls oligodendroglial specification and myelination. Cell Res. *29*, 23–41.

Xia, T.-L., Li, X., Wang, X., Zhu, Y.-J., Zhang, H., Cheng, W., Chen, M.-L., Ye, Y., Li, Y., Zhang, A., et al. (2021). N(6)-methyladenosine-binding protein YTHDF1 suppresses EBV replication and promotes EBV RNA decay. EMBO Rep. *22*, e50128.

Xiao, W., Adhikari, S., Dahal, U., Chen, Y.-S., Hao, Y.-J., Sun, B.-F., Sun, H.-Y., Li, A., Ping, X.-L., Lai, W.-Y., et al. (2016). Nuclear m6A Reader YTHDC1 Regulates mRNA Splicing. Molecular Cell *61*, 925. https://doi.org/10.1016/j.molcel.2016.03.004.

Xiao, Y., Wang, Y., Tang, Q., Wei, L., Zhang, X., and Jia, G. (2018). An Elongation- and Ligation-Based qPCR Amplification Method for the Radiolabeling-Free Detection of Locus-Specific N -Methyladenosine Modification. Angew. Chem. Int. Ed Engl. *57*, 15995–16000.

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 A methyltransferase METTL5 promotes mouse embryonic stem cell differentiation. EMBO Rep. *21*, e49863.

Xu, C., Wang, X., Liu, K., Roundtree, I.A., Tempel, W., Li, Y., Lu, Z., He, C., and Min, J. (2014). Structural basis for selective binding of m6A RNA by the YTHDC1 YTH domain. Nat. Chem. Biol. *10*, 927–929.

Xu, C., Liu, K., Ahmed, H., Loppnau, P., Schapira, M., and Min, J. (2015). Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. J. Biol. Chem. *290*, 24902–24913.

Xu, L., Liu, X., Sheng, N., Oo, K.S., Liang, J., Chionh, Y.H., Xu, J., Ye, F., Gao, Y.-G., Dedon, P.C., et al. (2017). Three distinct 3-methylcytidine (mC) methyltransferases modify tRNA and mRNA in mice and humans. J. Biol. Chem. *292*, 14695–14703.

Xue, C., Zhao, Y., and Li, L. (2020). Advances in RNA cytosine-5 methylation: detection, regulatory mechanisms, biological functions and links to cancer. Biomark Res *8*, 43.

Yang, Y., Fan, X., Mao, M., Song, X., Wu, P., Zhang, Y., Jin, Y., Yang, Y., Chen, L.-L., Wang, Y., et al. (2017). Extensive translation of circular RNAs driven by N6-methyladenosine. Cell Res. *27*, 626–641.

Yelick, P.C., and Trainor, P.A. (2015). Ribosomopathies: Global process, tissue specific defects. Rare Dis *3*, e1025185.

Yildirim, I., Kierzek, E., Kierzek, R., and Schatz, G.C. (2014). Interplay of LNA and 2'-O-methyl RNA in the structure and thermodynamics of RNA hybrid systems: a molecular dynamics study using the revised AMBER force field and comparison with experimental results. J. Phys. Chem. B *118*, 14177–14187.

Youn, J.-Y., Dunham, W.H., Hong, S.J., Knight, J.D.R., Bashkurov, M., Chen, G.I., Bagci, H., Rathod, B., MacLeod, G., Eng, S.W.M., et al. (2018). High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. Mol. Cell *69*, 517–532.e11.

Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., Cheng, T., Gao, M., Shu, X., Ma, H., et al. (2018). VIRMA mediates preferential m6A mRNA methylation in 3′UTR and near stop codon and associates with alternative polyadenylation. Cell Discovery *4*. https://doi.org/10.1038/s41421-018-0019-0.

Zaccara, S., and Jaffrey, S.R. (2020). A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA. Cell *181*, 1582–1595.e18.

Zhang, F., Kang, Y., Wang, M., Li, Y., Xu, T., Yang, W., Song, H., Wu, H., Shu, Q., and Jin, P. (2018). Fragile X mental retardation protein modulates the stability of its m6A-marked messenger RNA targets. Hum. Mol. Genet. *27*, 3936–3950.

Zhang, L.-H., Zhang, X.-Y., Hu, T., Chen, X.-Y., Li, J.-J., Raida, M., Sun, N., Luo, Y., and Gao, X. (2020a). The SUMOylated METTL8 Induces R-loop and Tumorigenesis via m3C. iScience *23*, 100968.

Zhang, L.-S., Liu, C., Ma, H., Dai, Q., Sun, H.-L., Luo, G., Zhang, Z., Zhang, L., Hu, L., Dong, X., et al. (2019). Transcriptome-wide Mapping of Internal N-Methylguanosine Methylome in Mammalian mRNA. Mol. Cell *74*, 1304–1316.e8.

Zhang, L.-S., Xiong, Q.-P., Peña Perez, S., Liu, C., Wei, J., Le, C., Zhang, L., Harada, B.T., Dai, Q., Feng, X., et al. (2021). ALKBH7-mediated demethylation regulates mitochondrial polycistronic RNA processing. Nat. Cell Biol. *23*, 684–691.

Zhang, Z., Luo, K., Zou, Z., Qiu, M., Tian, J., Sieh, L., Shi, H., Zou, Y., Wang, G., Morrison, J., et al. (2020b). Genetic analyses support the contribution of mRNA N-methyladenosine (mA) modification to human disease heritability. Nat. Genet. *52*, 939–949.

Zheng, G., Qin, Y., Clark, W.C., Dai, Q., Yi, C., He, C., Lambowitz, A.M., and Pan, T. (2015). Efficient and quantitative high-throughput tRNA sequencing. Nat. Methods *12*, 835–837.

Zhong, S., Li, H., Bodi, Z., Button, J., Vespa, L., Herzog, M., and Fray, R.G. (2008). MTA is an Arabidopsis messenger RNA adenosine methylase and interacts with a homolog of a sex-specific splicing factor. Plant Cell *20*, 1278–1288.

Zhou, H., Rauch, S., Dai, Q., Cui, X., Zhang, Z., Nachtergaele, S., Sepich, C., He, C., and Dickinson, B.C. (2019a). Evolution of a reverse transcriptase to map N-methyladenosine in human messenger RNA. Nat. Methods *16*, 1281–1288.

Zhou, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.-B. (2015). Dynamic m6A mRNA methylation directs translational control of heat shock response. Nature *526*, 591–594. https://doi.org/10.1038/nature15377.

Zhou, J., Wan, J., Shu, X.E., Mao, Y., Liu, X.-M., Yuan, X., Zhang, X., Hess, M.E., Brüning, J.C., and Qian, S.-B. (2018). N6-Methyladenosine Guides mRNA Alternative Translation during Integrated Stress Response. Mol. Cell *69*, 636–647.e7.

Zhou, K.I., Parisien, M., Dai, Q., Liu, N., Diatchenko, L., Sachleben, J.R., and Pan, T. (2016). N6-Methyladenosine Modification in a Long Noncoding RNA Hairpin Predisposes Its Conformation to Protein Binding. Journal of Molecular Biology *428*, 822–833. https://doi.org/10.1016/j.jmb.2015.08.021.

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019b). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523.

Zhu, T., Roundtree, I.A., Wang, P., Wang, X., Wang, L., Sun, C., Tian, Y., Li, J., He, C., and Xu, Y. (2014). Crystal structure of the YTH domain of YTHDF2 reveals mechanism for recognition of N6-methyladenosine. Cell Res. *24*, 1493–1496.

Zou, Z., Sepich-Poore, C., Zhou, X., Wei, J., and He, C. (2022). The mechanism underlying redundant functions of the YTHDF proteins.