The Conserved Role of METTL16 m6A Methyltransferase in Gene Expression Regulation

Mendel, Mateusz

UNIVERSITÉ DE GENÈVE FACULTÉ DES SCIENCES

Section de biologie

Département de biologie moléculaire Professeur Ramesh S. Pillai

The Conserved Role of METTL16 m^{6}A Methyltransferase in Gene Expression Regulation

THÈSE

présentée aux Facultés de médecine et des sciences de l'Université de Genève

pour obtenir le grade de Docteur ès sciences en sciences de la vie,

mention Biosciences moléculaires

par

Mateusz MENDEL

de

Mikołów (Poland)

Thèse N^o 115

GENÈVE

2021

LIFE SCIENCES PhD SCHOOL GENEVA

DOCTORAT ÈS SCIENCES EN SCIENCES DE LA VIE DES
FACULTÉS DE MÉDECINE ET DES SCIENCES
MENTION BIOSCIENCES MOLÉCULAIRES

Thèse de Monsieur Mateusz MENDEL

intitulée :

«The Conserved Role of METTL16 m^{6}A Methyltransferase in Gene Expression Regulation»

Les Facultés de médecine et des sciences, sur le préavis de Monsieur R. PILLAI, professeur ordinaire et directeur de thèse (Département de biologie moléculaire), Monsieur M. KAKSONEN, professeur ordinaire (Département de biochimie), Madame S. VANACOVA, professeure assistante (Central European Institute of Technology, Masaryk University, Brno, Czech Republic), Monsieur J.-Y. ROIGNANT, professeur associé (Centre intégratif de génomique, Faculté de biologie et de médecine, Université de lausanne), autorisent l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 5 septembre 2021

Thèse - 115 -

Le Doyen
Faculté de médecine

Le Doyen
Faculté des sciences

N.B. - La thèse doit porter la déclaration précédente et remplir les conditions énumérées dans les "Informations relatives aux thèses de doctorat à l'Université de Genève".

Table of Contents

1. Abbreviations

3′SS – 3′ splice site

5′SS – 5′ splice site

ac^{4}C – *N*^{4}-acetylcytidine

AML – acute myeloid leukaemia

BP – branch point

C. elegans – Caenorhabditis elegans

CBC – cap-binding complex

CFIm – cleavage factor Im

CHO – chinese hamster ovary

cLEU – cycloleucine

CMTR – Cap-specific mRNA 2 ′O-methyltransferase

CTD – carboxy-terminal domain

D. melanogaster – Drosophila melanogaster

DNA – deoxyribonucleic acid

eIF3 – eukaryotic initiation factor 3

ENE+A – element of nuclear expression with a downstream A-rich

ESCs – embryonic stem cells

FTO – fat mass and obesity-associated protein

hnRNA – heterogenous nuclear RNA

hnRNPC – heterogenous nuclear ribonucleoprotein C

hnRNPs – heterogenous nuclear ribonucleoproteins

I – inosine

IGF2BPs – insulin-like growth factor 2 mRNA binding proteins

Ime4 - Inducer of meiosis 4

IP – immunoprecipitation

KA1 – kinase-associated 1

KD – knock-down

KO – knock-out

LC–MS/MS – Liquid Chromatography with tandem mass spectrometry

lncRNA – long non-coding RNA

m^1A – N^1-methyladenosine

m^5C – 5-methylcytidine

m^6A – N^6-methyladenosine

m^6A-IP-seq – m^6A-immunoprecipitation-sequencing

m^6A_m – N^6,2′-O-dimethyladenosine

m^7G – N^7-methylguanosine

MAT – methionine adenosyltransferase

METT–10 – Methyltransferase homolog 10

METTL3 – Methyltransferase Like 3

METTL14 – Methyltransferase Like 14

METTL16 – Methyltransferase Like 16

mRNA – messenger RNA

MS – mass spectrometry

MTD – methyltransferase domain

N_m – 2′-O-methyl

nt – nucleotide

PN – poikilotderma with neutropenia

polyA – polyadenosine

PPT – polypyrimidine tract

pre–mRNA – precursor messenger RNA

RNA – ribonucleic acid

rRNA – ribosomal RNA

S. cerevisiae – Saccharomyces cerevisiae

S. pombe – Schizosaccharomyces pombe

SAH – S-adenosylcysteine

SAM – S-adenosylmethionine

SCARLET - Site-specific cleavage (RNaseH) and radioactive-labelling followed by ligation-assisted extraction and thin-layer chromatography

SE – exon skipping

snRNA – small nuclear RNAs

snRNPs – small nuclear ribonucleoprotein complexes

SR – serine/arginine rich

Sxl – sex lethal

tRNA – transfer RNA

TUT1 – uridyl transferase 1

U2AF – U2 small nuclear ribonucleoprotein auxillary factor

UHPLC–MS/MS – Ultra-high performance liquid chromatography-MS/MS

VCR – vertebrate conserved domain

WT – wild type

WTAP – Wilms' tumor 1-associating protein

YTH – YT521-B homology

YTHDC – YTH domain-containing

YTHDF – YTH domain-containing family protein

Ψ – pseudouridine

2. Abstract (English)

During recent years, RNA modifications emerged as a new layer of post-transcriptional gene expression regulation, with more than 170 different modifications detected up to date. The most common internal mRNA modification is N^6-methyladenosine (m^6A). It is deposited on pre-mRNA by two methyltransferases: METTL3-METTL14 heterodimer and METTL16. METTL3/14 deposits the majority of m^6A in the cell and is essential for embryonic development in plants and animals. Methylation by METTL3/14 was shown to regulate mRNA degradation, translation, nuclear export, or pre-mRNA splicing. The role of METTL16 methyltransferase is much less understood. METTL16 is conserved from *E. coli* to humans, and in mammals it methylates U6 snRNA and *Mat2a* mRNA, encoding for S-adenosyl methionine (SAM) synthetase. SAM is a methyl group donor essential in the majority of methylation reactions, including m^6A deposition. METTL16 was shown to bind conserved RNA hairpins in the last exon of *Mat2a* pre-mRNA to promote splicing of the last *Mat2a* intron. However, intron splicing is enhanced not by the m^6A methylation of the *Mat2a* transcript, but by the presence of METTL16 itself, which stimulates splicing through so-called vertebrate conserved regions (VCR) located in the C-terminal part of the protein. My research aimed to understand the physiological role of METTL16 in mice, identify its additional targets, and understand its role in gene expression regulation in *C. elegans*.

In my PhD project, we generated the *Mettl16* knock-out (KO) mouse line and showed that METTL16 is essential for embryonic development, with *Mettl16* KO embryos dying between embryonic days 3.5 to 6.5 (during the implantation stage). Analysis of the early embryos isolated at morula (E2.5) and blastocyst (E3.5) stages showed a drastic decrease in *Mat2a* mRNA levels. As *Mat2a* is the only SAM synthetase expressed in the developing embryo, we proposed that the observed lethality was due to the failure in re-establishing DNA methylation levels caused by SAM deficiency. In addition, we showed that METTL16 is also important outside of embryonic

development, with conditional deletion of *Mettl16* in mouse male germline leading to arrested germ cell development and infertility. Finally, we determined RNA substrate requirements of METTL16 and identified the N-terminal part of the protein as necessary for RNA binding.

In the second part of my project, we investigated the role of METTL16 homologue, METT-10, in *C. elegans*. We showed that METT-10 is an m^6A methyltransferase and that its targets are conserved between mice and worms, with U6 snRNA and SAM synthetase transcripts: *sams-3*,*-4*,*-5* being methylated. In worms, *sams* transcripts were methylated at the 3′ splice site (3′SS), leading to splicing inhibition and transcript degradation. This mechanism was active in nutrient-rich conditions and acted as an m^6A-mediated switch to stop SAM production and regulate its homeostasis. In addition, we showed that 3′SS m^6A inhibits splicing by preventing an essential splicing factor U2AF35 binding to the 3′SS. Although in mammals SAM synthetase splicing is regulated differently, the mechanism of splicing inhibition by 3′SS methylation is conserved.

Taken together, my work provided new insights into METTL16 mechanism of action and physiological role in mouse and *C. elegans*. In addition, I identified an evolutionarily conserved mechanism of splicing regulation through m^6A methylation directly inhibiting U2AF35 binding to the 3′SS. This discovery raises a possibility that RNA modifications might be a new layer of splicing regulation.

3. Abstract (French)

Au cours des dernières années, les modifications de l'ARN sont apparues comme un niveau supplémentaire de régulation post-transcriptionnelle de l'expression des gènes, avec plus de 170 modifications différentes détectées à ce jour. La modification interne la plus courante de l'ARN messager (ARNm) est la N6-méthyladénosine (m^6A). Elle est déposée sur le pré-ARNm par deux méthyltransférases : l'hétérodimère METTL3-METTL14 et METTL16. METTL3/14 dépose la majorité de m^6A dans la cellule et est essentielle au développement embryonnaire chez les plantes et les animaux. Il a été démontré que la méthylation par METTL3/14 peut réguler la traduction et la dégradation de l'ARNm, l'export nucléaire ou l'épissage du pré-ARNm. Le rôle de la méthyltransférase METTL16 est beaucoup moins bien compris. METTL16 est conservée d'*E. coli* à l'homme, et chez les mammifères elle méthyle les petits ARN nucléaires (ARNsn) U6 et l'ARNm *Mat2a*, codant pour la S-adénosyl méthionine (SAM) synthétase. La SAM est une donneuse de groupe méthyle essentiel dans la majorité des réactions de méthylation, y compris le dépôt de m^6A. Il a été démontré que METTL16 se lie à des structures en épingle à cheveux conservées dans le dernier exon du pré-ARNm *Mat2a* pour promouvoir l'épissage du dernier intron. Cependant, l'épissage de l'intron est favorisé non pas par la méthylation m^6A du transcrit *Mat2a*, mais par la présence de METTL16 elle-même, qui stimule l'épissage par l'intermédiaire des régions conservées des vertébrés (VCR) situées dans la partie C-terminale de la protéine. Mon projet de thèse vise à comprendre le rôle physiologique de METTL16 chez la souris, à identifier ses autres cibles et à comprendre son rôle dans la régulation de l'expression des gènes chez *C. elegans*.

Dans le cadre de mon projet de thèse, nous avons généré une lignée de souris knock-out (KO) *Mettl16* et montré que METTL16 est essentielle au développement embryonnaire. En effet, les embryons KO *Mettl16* meurent entre les jours embryonnaires 3,5 et 6,5 (durant le stade d'implantation). L'analyse des embryons isolés aux stades morula (E2.5) et blastocyste (E3.5) a

montré une diminution drastique des niveaux d'ARNm *Mat2a*. Comme *Mat2a* est la seule SAM synthétase exprimée dans l'embryon en développement, nous avons proposé que la létalité peut être due à l'échec du rétablissement des niveaux de méthylation de l'ADN causé par la déficience en SAM. En outre, nous avons montré que METTL16 est également importante en dehors du développement embryonnaire. L'élimination conditionnelle de Mettl16 dans la lignée germinale mâle de la souris entraîne l'arrêt du développement des cellules germinales et l'infertilité. Enfin, nous avons déterminé les substrats d'ARN requis pour METTL16 et identifié la partie N-terminale de la protéine comme étant nécessaire pour la liaison à l'ARN.

Dans la deuxième partie de mon projet, nous avons étudié le rôle de l'homologue de METTL16, METT-10, chez *C. elegans*. Nous avons montré que METT-10 est une méthyltransférase dont les cibles sont conservées entre la souris et le ver. En effet, les ARNsn U6 et les transcrits codant pour les SAM synthétases, *sams* 3, 4, 5, sont méthylés. Chez les vers, les transcrits *sams* sont méthylés au niveau du site d'épissage 3′ (3′SS), entraînant une inhibition de l'épissage et une dégradation du transcrit. Ce mécanisme est actif dans des conditions riches en nutriments et agit comme un signal pour arrêter la production de SAM et réguler son homéostasie en réponse à m^6A. De plus, nous avons montré que la modification m^6A sur le 3′SS inhibe l'épissage en empêchant un facteur d'épissage essentiel U2AF35 de s'y lier. Bien que chez les mammifères l'épissage de la SAM synthétase soit régulé différemment, le mécanisme d'inhibition de l'épissage par la méthylation du 3′SS est conservé.

Dans l'ensemble, mes travaux ont permis de mieux comprendre le mécanisme d'action de METTL16 et son rôle physiologique chez la souris et chez *C. elegans*. De plus, j'ai identifié un mécanisme de régulation de l'épissage conservé au cours de l'évolution par la méthylation de m^6A qui inhibe la liaison de U2AF35 au 3′SS. Cette découverte soulève la possibilité que les modifications de l'ARN puissent constituer une nouvelle couche de régulation de l'épissage

4. Acknowledgements

First and foremost, I would like to thank Ramesh for accepting me in the group, for our countless discussions, guidance, and support through all these years. It was a long journey, but a very fruitful one, and I've learned a lot.

Next, I would like to thank members of my TAC committee, André Verdel and Robbie Loewith, for their advice and insights, especially at the beginning of my PhD. I would also like to thank Prof. Štěpánka Vaňáčová and Prof. Jean-Yves Roignant for agreeing to be members of my Thesis Defence Committee and taking time to read my thesis.

I would like to thank Marko Kaksonen for his continuous support, advice, and help during my PhD. I always felt welcomed to discuss and ask for advice.

My PhD would not be possible without all the current and previous members of the Pillai lab. I feel very fortunate that I could work in this stimulating and fun environment. Although everybody was important, I want to especially mention David Homolka, Magda Wojtas and Kuan-Ming Chen. David, thanks for our never-ending discussions, your constructive criticism and advice, all the help with bioinformatics, coffee breaks (before COVID), beers in the evenings, trying to convince me to start climbing or running and much, much more. Magda, thank you for the warm welcome I got, teaching me everything, your patience and your positive approach. You're the best teacher I've ever met. Kuan-Ming, thanks for all your protein purifications, but much more for being my bench-mate during all these years, your patience when I was stealing your Pipetboy, shopping together, helping me out, cool stories, "Cinema Paradiso" quotes and shouting "wa kao" all the time.

Of course, I can't forget about the present and past members of our lab's gang: Michaela, Elena, Mark, Kyrylo, Adriana and Lingyun. You made the last two years plenty of fun and helped

me to open up. I really enjoyed our parties, trips or Nerf gunfights. Also, thanks a lot for keeping my basil alive. I hope it will become your office pet once I'm gone.

I want to thank all the students, especially Asia and Kamila, and post-docs from the Department for your help, advice and creating a great atmosphere. Asia, dzięki za nasze długie rozmowy, wspólne marudzenie, rady i całą pomoc, szczególnie z Career Insights. Kamilko, dzięki za Twoje ogromne zaangażowanie w projekt, ploteczki i pozytywne podejście do życia i pracy.

My PhD would be much more difficult without the wonderful people in our department's secretariat: Fabrice, Cécile, Agnes, Veronica. Thank you for all the last-minute orders, parcels and help with the Swiss bureaucracy. I also can't forget about the department's kitchen and atelier staff, keeping the department running and well-functioning.

Boehringer Ingelheim Fonds PhD Fellowship for the funding, personal support, friendships I found and unforgettable memories from all the meetings in Hirschegg and Mainz.

Last but not least, I would like to thank my family and friends, for constant support, inspiration, managing with my negativity and motivating to strive for more. I wouldn't be here without you. Mamo i Moniko, dziękuję za wsparcie, wiarę we mnie i codzienne rozmowy, bez was nic by się nie udało. Wujku Zbyszku, dziękuję, że zawsze trzymasz za mnie kciuki i że mogę na Ciebie liczyć. Maćku, dzięki za zainspirowanie mnie do studiowania biotechnologii. Być może byłbym dzisiaj lekarzem (!), gdybyś podczas swoich urodzin nie opowiadał o dniach otwartych biotechnologii na UJ. Ingo, mimo że sporo się zmieniło, dziękuję za wszystko co dla mnie zrobiłaś, całą radość, wsparcie, motywację i wszystke lata spędzone razem. Bez Ciebie nie byłbym w tym miejscu. Sacha, thank you for being there for me, for your support during the especially difficult times, surviving my negativity, discussing science, helping me to appreciate France and all the fun we had. Wojtku, dzięki że jesteś i znosisz mnie przez te wszystkie lata.

Pamięci mojej Babci Łucji

10.10.1940 – 29.11.2013

.

5. Introduction

5.1. Discovery of nucleic acids

The year 2021 marks the 60th anniversary of messenger RNA (mRNA) discovery, when on the 13th of May 1961, two articles, one co-authored by Sydney Brenner, second by Jim Watson, described the isolation of mRNA (Brenner et al., 1961; Gros et al., 1961). It was the effect of decades of work, which started in 1871 with the discovery of nucleic acids, more precisely deoxyribonucleic acid (DNA), by a Swiss chemist Friedrich Miescher (Dahm, 2005). Miescher performed a chemical analysis of pus cells, where he showed that the nucleus of human white blood cells contained a phosphorus-rich fraction resistant to proteolysis, which, due to its localization, he termed "nuclein" (Hall and Sankaran, 2021). Further studies have shown that there are two types of nucleic acids, which can be distinguished based on their sugar molecules (either ribose or deoxyribose) and were believed to have distinct localization: ribonucleic acids (RNA or yeast nucleic acid), which can be found only in plant tissues, and deoxyribonucleic acid (DNA or thymonucleic acid or animal nucleic acid), exclusive to animal tissues (Allen, 1941). However, this view was soon debunked with mounting evidence that in all organisms DNA is to be found in the nucleus, while RNA is mainly located in the cytoplasm (Davidson and Waymouth, 1943). In 1952 biochemist Erwin Chargaff published results of sea-urchin DNA analysis composition, where he showed that the ratio of purines to pyrimidines is always approximately 1 to 1. It became known as Chargaff Rule (Chargaff et al., 1952) and paved the way for further research and understanding how information is stored in the DNA. One of the most significant leaps in knowledge came a year later, with the publication of the DNA structure, the result of four scientists efforts: Rosalind Franklin, Maurice Wilkins, James Watson and Francis Crick. Watson and Crick, the authors of the study, showed that DNA forms a double-helix structure, with the nucleobases pairing inside the helix and phosphate groups facing

outside (Watson and Crick, 1953). They also speculated that the DNA sequence could be a basis of genetic information storage: "*it therefore seems likely that the precise sequence of the bases is the code which carries the genetical information*" (Watson and Crick, 1953).

5.2. How genetic information is stored and transfered?

Simultaneously to the discovery and description of nucleic acids, the identity of the genetic information carrier was being intensively discussed. Before that, people for centuries have wondered why children are similar to their parents and how different traits are passed from generation to generation. For example, a Greek philosopher Aristotle proposed that the father, through sperm, provides the "form", while the mother provides the "matter", although with females also being responsible for the development of certain traits (Henry, 2006). In 1868, Darwin, building on the ideas of another Greek philosopher, Hippocrates, proposed the concept of pangenesis, where all cells in a body would be capable of shedding minute particles called gemmules, which then would accumulate in the gonads. That would allow the environment to affect information passed from parents to the offspring (Liu, 2008). At a similar time to Darwin's publication, in 1866, a Moravian monk Johann Gregor Mendel presented his work "*Versuche über Pflanzen hybriden*", where he showed that inherited traits are not blending (as was believed at that time), but are either dominant or recessive. In addition, he proposed the existence of certain factors, which transfer information from each of parents to their progeny, every trait independent from the others. Initially forgotten, his rules were rediscovered 35 years later by Hugo de Vries, Erich von Tschermak and Carl Correns, laying the foundation of modern genetics. In 1909 the term "gene" was coined by a Danish botanist Wilhelm Johannsen to describe the Mendelian unit of heredity. However, the nature of the genes carrier remained unknown (Portin and Wilkins, 2017).

It was the research of three men: Walter Sutton, Theodor Boveri and Thomas Morgan, which showed that nuclear structures visible during the cell division (first described by Walther Flemming

in 1879), called chromosomes, are carriers of the genetic information. Sutton and Boveri first proposed the idea in 1902. However, it gained strong support with further experiments on fruit flies by Thomas Moran, who linked inheritance of a particular trait (white eyes) with the X chromosome and, in 1915, co-authored a seminal book, "*The Mechanism of Mendelian Heredity"*. Nevertheless, it was still not established what inside the chromosomes carries the genetic information. Initially, proteins were believed to be the carriers, and only in the early 1940s, breakthrough research on gene transfer mechanisms in bacteria showed that it is DNA (Avery et al., 1944).

5.3. The Central Dogma of Molecular Biology

The remaining question was the relationship between DNA, RNA, and proteins, and how dynamic it is. In 1957, Francis Crick, during his lecture at the Society of Experiment Biology symposium at University College London, proposed an idea, nowadays known as "Central Dogma of Molecular Biology", which linked all the pieces together and revolutionized biology.

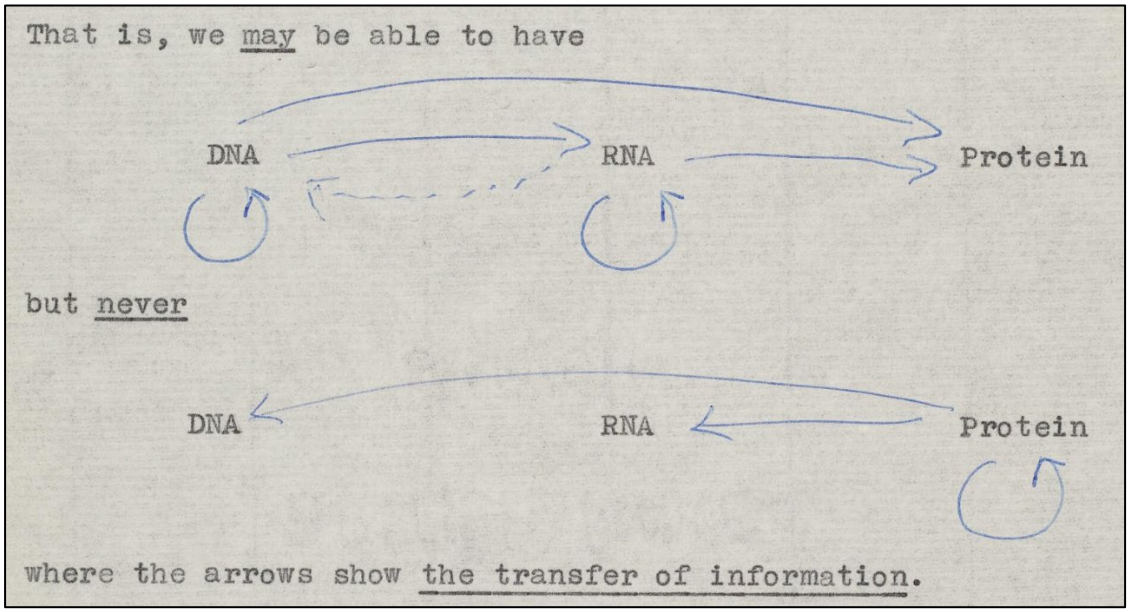


Figure 1. Francis Crick's unpublished notes with the first description of the Central Dogma (1956). Credit: Wellcome collection (ref. PP/CRI/H/2/6).

In the lecture and the following article entitled "*On protein synthesis*", Crick proposed that "[…] *once 'information' has passed into protein it cannot get out again. In more detail, the transfer of*

information from nucleic acid to nucleic acid, or from nucleic acid to protein may be possible, but transfer from protein to protein, or from protein to nucleic acid is impossible." (Crick, 1958). Central dogma brought a logic for the flow of the information, where DNA is responsible for storing all the genetic information. When needed, fragments of this information are being transcribed into messenger RNA (mRNA), which is used as a template to synthesize proteins in the process of translation. Although Crick hypothesised that the information could move from RNA to DNA (rightly so), there is no possibility that any information contained in the protein sequence could be transferred back to RNA or DNA (Figure 1). Over the following years, it was shown that eukaryotic mRNA is not the exact copy of the DNA, but it has to be extensively processed to be translated.

5.4. mRNA processing

The first step of mRNA synthesis is the transcription of DNA template by RNA polymerase II, which ultimately results in the synthesis of a precursor messenger RNA (pre-mRNA). Pre-mRNA subsequently undergoes several maturation steps to form mRNA, that include: 5′ capping, removal of non-coding regions in the process of splicing, decorating of pre-mRNA with chemical modifications and 3′ end processing and polyadenylation (reviewed in (Hocine et al., 2010)). All of these steps are happening co-transcriptionally. The mature mRNA is next bound by specific factors allowing for nuclear export in order to reach the cytoplasm and be translated. One of the requirements for mRNA maturation and export is the successful completion of splicing.

5.4.1. pre-mRNA splicing

During the 1960s and 1970s, the molecular biology field was trying to understand the mechanisms of mRNA synthesis in the mammalian cells. The early studies showed the presence of an unstable RNA fraction that might be the animal equivalent of mRNA previously found in bacteria (Scherrer et al., 1963). This unstable RNA fraction was initially termed heterogeneous nuclear RNA (hnRNA)

and was suspected to be a precursor to cytoplasmic mRNA (Soeiro et al., 1966). Further studies showed that both hnRNA and mRNA are polyadenylated at the 3′ end, suggesting both a direct link between the two and the importance of poly(A) in mRNA processing (Edmonds et al., 1971). However, the relation between hnRNA and mRNA remained unclear until 1977, when two groups combined the method of RNA – DNA hybridization with the electron microscopy analysis, providing the first visualisation of splicing using adenovirus hnRNA (Berget et al., 1977; Chow et al., 1977).

Studies done during the next decades showed that pre-mRNAs are composed of coding and non-coding sequences called exons and introns. To form a mature mRNA, introns have to be removed, and exons joined together in the process of splicing, performed by intricate molecular machinery called spliceosome (Brody and Abelson, 1985). Spliceosome contains five different small nuclear ribonucleoprotein complexes (snRNPs) composed of uridine-rich small nuclear RNAs (snRNA) U1, U2, U4, U5 and U6 snRNAs associated with snRNP-specific proteins. In addition, snRNPs are further associated with many different protein cofactors creating a massive complex of approximately 100 different proteins (reviewed in (Plaschka et al., 2019; Wilkinson et al., 2020)).

5.4.1.1. Splice site recognition

A critical first step in pre-mRNA splicing is the recognition of correct exon/intron boundaries, which are defined by a set of specific short sequences at the 5′ splice site (5′SS), branch point (BP) sequence and the 3′ splice site (3′SS) (revised in (Wilkinson et al., 2020)). In humans, these sequences are: **GU**RAGN for the 5′SS, YNYURAY for BP and the 3′SS is defined by both a stretch of pyrimidines close to the end of the intron called polypyrimidine tract (PPT) as well as Y**AG** sequence exactly at the 3′SS (GU and AG marked in bold are the first and the last nucleotides of the intron) (revised in (Wilkinson et al., 2020)). These sequences are specifically bound by a set of splicing factors: the 5′ splice site by U1 snRNA (Mount et al., 1983), branch point by the branch point binding protein/splicing factor 1 (BBP/SF1) (Berglund et al., 1997) and the 3′ splice site by the U2 small

nuclear ribonucleoprotein auxiliary factor (U2AF) (Ruskin et al., 1988), comprised of two subunits: 65-kDa U2AF^{65} and 35-kDa U2AF^{35} (Zamore and Green, 1989). These sequences are bound by splicing factors leading to the formation of the first spliceosome stage called E complex (Wilkinson et al., 2020). Next, the presence of BBP and U2AF proteins on pre-mRNA recruits U2 snRNP, which binds to the branch point leading to the assembly of the prespliceosome complex (A complex) (Kramer and Utans, 1991). Then, U1 and U2 snRNPs recruit the U4/U6•U5 tri-snRNP, which binding leads to additional rearrangements and the formation of the catalytically active spliceosome. Finally, U1 and U4 snRNAs dissociate, leaving U2/U6•U5 complex together with additional factors to catalyse chemical steps (reviewed in (Wilkinson et al., 2020)).

The two subunits of the U2AF complex bind to distinct sequences at the 3′ splice site. U2AF^{65} directly contacts the polypyrimidine tract, while U2AF^{35} binds exactly to the 3′SS YAG sequence (Wu et al., 1999). First studies showed that while U2AF^{65} is absolutely essential for splicing, both *in vitro* and *in vivo*, U2AF^{35} is required for viability, but not *in vitro* splicing (Zamore and Green, 1991; Zhang et al., 1992). Next, it was reported that binding of U2AF^{35} is critical for splicing of introns having weak polypyrimidine tracts, which are not efficiently bound by U2AF^{65} alone (Wu et al., 1999). Based on that, introns can be divided into two categories, AG-dependent introns, where U2AF^{35} binding is essential for splicing, and AG-independent introns, where intron recognition by U2AF^{65} alone is sufficient for splicing (Wu et al., 1999). The presence of "weak" splice sites allows for modulation of splicing in different conditions in a process known as alternative splicing.

5.4.1.2. Alternative splicing

The complexity of splicing regulation allows for a process of alternative splicing, where one pre-mRNA transcript can be spliced in different ways, resulting in many different mRNA transcripts, potentially encoding a variety of protein variants. Alternative splicing was proposed to be one of the significant sources of species-specific differences and correlated with organism complexity

(Graveley, 2001). While the number of protein-coding genes is similar among invertebrates and vertebrates, the level of alternative splicing is much higher in vertebrates, reaching especially high levels in humans (Barbosa-Morais et al., 2012). Latest genome-wide studies estimated that 90 – 95% of human genes are alternatively spliced (Pan et al., 2008). Alternative splicing differs not only between species, but is also cell-type and tissue-specific, with enrichment of alternatively spliced genes in brain, muscle, testis or embryonic stem cells (ESCs) (Tapial et al., 2017; Yeo et al., 2004). There are several different types of alternative splicing events possible, including alternative 5′ and 3′ splice site usage, exon skipping (SE), mutually exclusive exon splicing, intron retention, alternative promoters and alternative poly(A) (Keren et al., 2010). The majority of alternative splicing events in mammals constitute intron retention and exon skipping events, with a very low percentage of alternative 5′ and 3′ splice site events (Tapial et al., 2017).

Alternative splicing is controlled through cis-acting RNA elements in exons and introns, which are recognized by trans-acting regulators. These factors interact then with the core splicing machinery, increasing (in case of splicing enhancer) or restricting (in case of splicing silencers) access to the splice sites (Furlanis and Scheiffele, 2018). The best-known example of splicing enhancers is the family of serine/arginine (SR)-rich proteins, consisting of 12 members SRSF1 to SRSF12. Proteins involved in splice site usage inhibition (splicing silencers) belong to the family of heterogenous nuclear ribonucleoproteins (hnRNPs) (Busch and Hertel, 2012). Another layer of alternative splicing control has arisen in recent years, based on the interaction between mRNA chemical modifications and splicing machinery.

5.5. RNA modifications

RNA consists of four nucleotides, which are ribose sugars covalently attached to nitrogenous bases: adenine (A), guanine (G), cytosine (C) and uracil (U). The first hint that RNA can be further modified or that other nucleotides exist, came in the 1950s, with a discovery of the so-called “fifth base” in

yeast, which was later identified as pseudouridine (Davis and Allen, 1957). Over the following decades, many more variants and chemical modifications were identified, with over 170 different RNA modifications known today, out of which 71 are present in eukaryotes (Boccaletto et al., 2018). Most of these modifications are found on non-coding RNAs, with transfer RNAs (tRNAs) being the most extensively modified RNA (average 13 modifications per molecule) (Pan, 2018), but also ribosomal RNAs (rRNAs) (Decatur and Fournier, 2002) and small nuclear RNAs (snRNAs) (Morais et al., 2021). However, during recent years it became clear that also mRNA can be chemically modified with several different modifications described so far (reviewed in (Wiener and Schwartz, 2020).

5.5.1. mRNA modifications

One of the first discovered mRNA modifications were methylation at the 5′ end of the transcript forming a cap structure (Wei et al., 1975) and 3′ end polyadenosine (polyA) tail (Darnell et al., 1971; Edmonds et al., 1971), with both modifications needed for pre-mRNA maturation and mRNA export from the nucleus (Shatkin and Manley, 2000). The cap structure formation involves linking N^7-methylguanosine (m^7G) with the first nucleotide via reverse 5′ to 5′ triphosphate linkage (Shatkin, 1976). The two following nucleotides can have their ribose methylated by CMTR1 and CMTR2 enzymes to generate 2′-O-methyl (N_m) modification (Werner et al., 2011). In addition, if the first 2′-O-methyl nucleotide is adenosine, it can be further methylated by PCIF1 to create N^6,2′-O-dimethyladenosine (m^6A_m) (Akichika et al., 2019; Sendinc et al., 2019). The m^7G cap associates with the nuclear cap-binding complex (CBC) and is not only needed to protect mRNA from degradation (Shimotohno et al., 1977), but is also required for splicing (Konarska et al., 1984), nuclear export (Nojima et al., 2007) and translation initiation (Fortes et al., 2000). In addition, ribose methylation deposited by CMTR1 was shown to be essential for self versus non-self discrimination and protection from the interferon response (Devarkar et al., 2016).

Besides the cap structure, several internal mRNA modifications were identified so-far (selected modifications shown in the Figure 2): ribose methylation (N_m), 5-methylcytidine (m^5C), pseudouridine (Ψ), N^1-methyladenosine (m^1A), N^7-methylguanosine (m^7G), N^4-acetylcytidine (ac^4C), inosine (I), N^6,2′-O-dimethyladenosine (m^6A_m) and N^6-methyladenosine (m^6A) (reviewed in (Wiener and Schwartz, 2020).

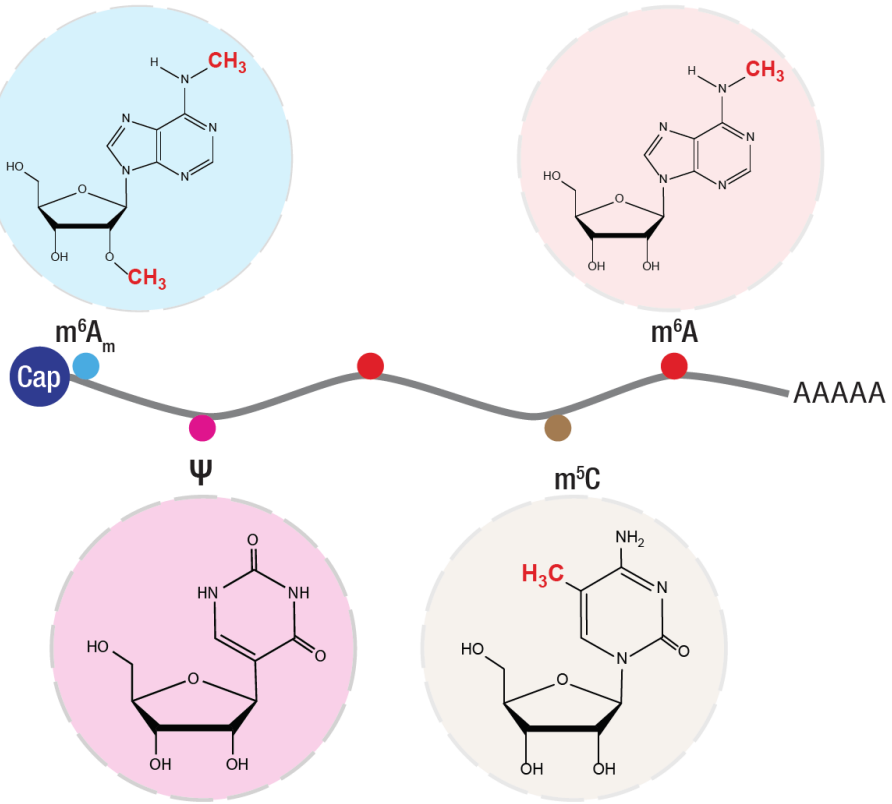


Figure 2. An example of selected chemical modifications present on messenger RNA. The modified sites are marked in red.

As the field is rapidly developing, the presence of some of these modifications on mRNA is being questioned, with conflicting studies concerning m^1A and ac^4C. Initially, m^1A was shown to be enriched in the 5′ UTR and mitochondrial transcripts (Li et al., 2017), while subsequent studies failed to find m^1A on mRNA and attributed the signal to antibody cross-reactivity (Grozhik et al., 2019; Safra et al., 2017). A similar problem concerns mRNA acetylation (ac^4C), which was initially shown to be widespread in mammalian mRNA and affect translation (Arango et al., 2018), while another

study found no ac^{4}C in both mammalian and yeast RNA (Sas-Chen et al., 2020). Nevertheless, the most common and the most studied internal mRNA modification is *N*^{6}-methyladenosine (m^{6}A).

5.5.2. *N*^{6}-methyladenosine

The first surge of interest in mRNA modifications took place in the 1970s, with dozens of articles pointing towards the presence of methylation at sixth nitrogen of adenosine, so called *N*^{6}-methyladenosine (m^{6}A) (Figure 3). It was identified on heterogeneous nuclear RNA (hnRNA) in mice (Perry and Kelley, 1974), rats (Desrosiers et al., 1974), human cells (Wei et al., 1976) and viruses (Krug et al., 1976). These initial studies also had a Geneva-related accent, as one of the initial studies was published by UNIGE professor emeritus Ueli Schibler during his postdoctoral stay in the group of Robert Perry in Philadelphia (Schibler et al., 1977).

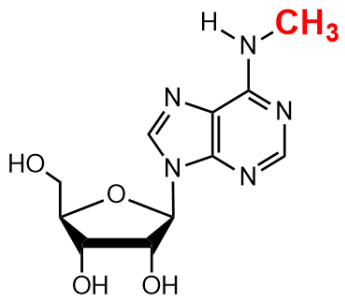


Figure 3. The structure of *N*^{6}-methyladenosine with the methyl group marked in red.

The interest in m^{6}A continued, with the antibodies raised against m^{6}A by Sims lab in 1977 (Munns et al., 1977) and by Luhrmann lab in 1987 (Bringmann and Lührmann, 1987); a demonstration that m^{6}A residues are distributed in a non-random pattern in mammalian mRNA, with enrichment at the 3′ end of mRNA in 1984 (Horowitz et al., 1984) and finally, identification of the enzyme responsible for the deposition of m^{6}A in 1997 (Bokar et al., 1997).

All these studies laid the groundwork for a revolution that came with the arrival of the next-generation sequencing of RNA. In 2012, two groups, one led by Sammie Jaffrey (Meyer et al., 2012) and the second by Gideon Rechavi (Dominissini et al., 2012), used antibodies recognizing m^{6}A

modification to precipitate and sequence mammalian mRNAs containing m^6A (m^6A-IP-seq). They identified thousands of methylation sites across the transcriptome, many of them conserved between mouse and human. The methylated sites overlapped with the RRACH (R = A/G, H = A, T, C) consensus motif, with a clear enrichment towards the 3′ end of transcripts (Dominissini et al., 2012; Meyer et al., 2012). Further studies showed that m^6A is deposited co-transcriptionally on pre-mRNA (Ke et al., 2017; Louloupi et al., 2018) and confirmed that the majority of m^6A is in the last exon, with a very sharp rise within 150 - 400 nucleotides of the start of the last exon (Ke et al., 2015). However, the mechanism driving the 3′ end enrichment of m^6A is not yet understood.

The recent studies showed that m^6A is the most common internal mRNA modification, with approximately 1 m^6A peak per 2000 nucleotides (1.7 peaks per gene) in the HepG2 cells (Dominissini et al., 2012). Further reports confirmed these estimates, with m^6A amounts to be approximately 0.2% of m^6A/A in mouse ESC (Geula et al., 2015) and between 0.11 to 0.23% adenosines methylated in mouse and human tissues (Liu et al., 2020). It is remarkably close to the initial m^6A prevalence assessments from the 1970s, which estimated the amount of m^6A to be around 0.2% of mRNA, with approximately three m^6A sites per mRNA (Perry et al., 1975). The high levels of m^6A mRNA and evolutionary conservation between yeast and humans indicate an important role of m^6A in mRNA biology.

5.5.2.1. The biological role of m^6A

Removal of METTL3/14, which deposits the majority of mRNA m^6A, leads to meiosis defects in yeast *Saccharomyces cerevisiae* (Clancy et al., 2002), sex determination and neuronal defects in *Drosophila melanogaster* (Haussmann et al., 2016; Lence et al., 2016) and early embryonic lethality in plants (Zhong et al., 2008) and mice (Batista et al., 2014; Geula et al., 2015). All these observations underline the essential role of m^6A in early embryonic development and gonad formation. In addition, m^6A was implicated in the control of numerous other biological processes in mammals, including

XIST-mediated transcriptional silencing (Patil et al., 2016); heat shock response (Meyer et al., 2015; Zhou et al., 2015); DNA repair after UV radiation (Xiang et al., 2017); circadian clock control (Fustin et al., 2013, 2018) or cell cycle progression and neurogenesis (Yoon et al., 2017), among others. Alternations in m^{6}A methylation levels are also connected with many different types of cancer (reviewed in (Barbieri and Kouzarides, 2020)), with the most studied role in the development and progression of acute myeloid leukaemia (AML) (Vu et al., 2017).

5.6. The Epitranscriptome

The deposition and action of m^{6}A are controlled by a set of proteins that can specifically deposit, recognize, or remove m^{6}A from the mRNA (Figure 4).

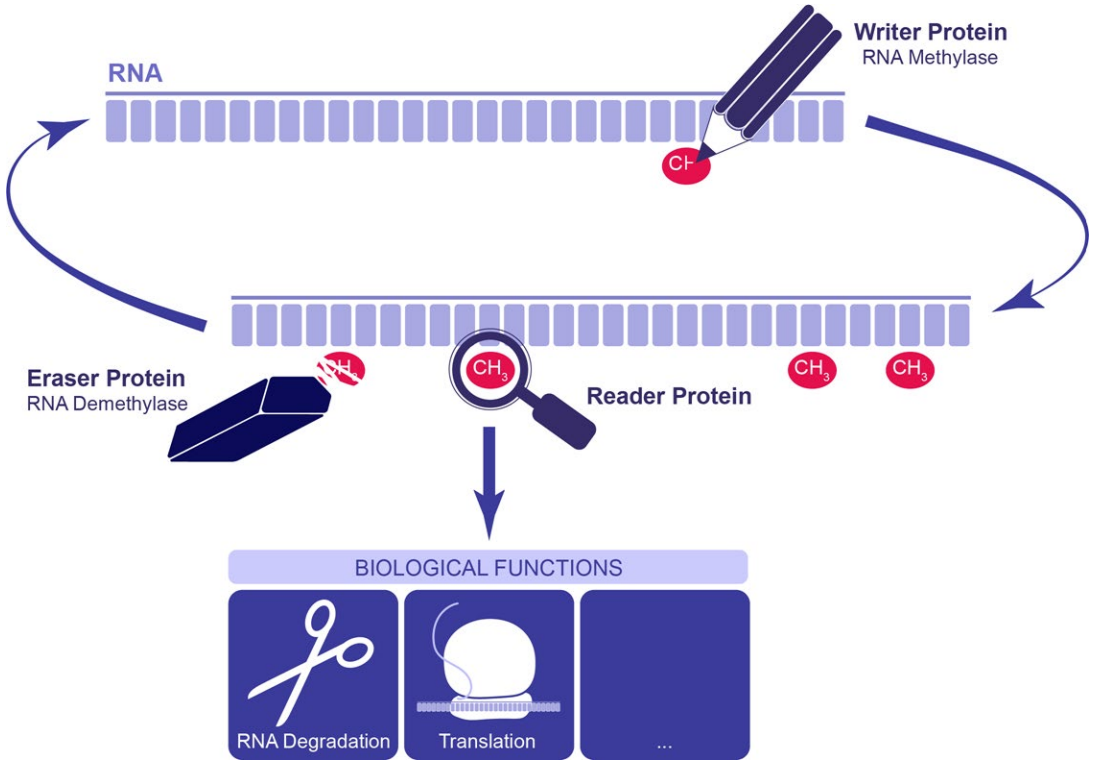


Figure 4. A landscape of factors involved in m^{6}A modification regulation and function. Picture credit: Margot Riggi.

m^{6}A is deposited on particular sites on mRNA by methyltransferase proteins, also referred to as “writers”. After the methylation, m^{6}A can be recognized by proteins called “readers, " which then perform specific biological functions or be removed by m^{6}A demethylases, called “erasers”. All these

factors allow for dynamic regulation of m^{6}A deposition on mRNA and create an entirely new layer of gene expression control, which due to its similarity to epigenetics, was called epitranscriptomics. In the following few chapters, I will discuss the factors involved in m^{6}A regulation in more detail.

5.6.1. m^{6}A Writers

The enzymes depositing *N*^{6}-methyladenosine on RNA belong to the family of methyltransferases, which transfer a methyl group from S-adenosylmethionine (SAM) (donor) to nucleotides (acceptor). So far, four enzymes were described to deposit m^{6}A on RNA. The two mRNA methyltransferases are METTL3/14, which deposits m^{6}A on the RRACH consensus motif and is responsible for the majority of m^{6}A methylation on mRNA (Liu et al., 2014) and METTL16, which methylates structured RNA with a specific consensus motif (UACAGAGAA) and has only two identified targets: U6 snRNA and *MAT2A* pre-mRNA (Pendleton et al., 2017). The two remaining enzymes, METTL5 and ZCCH4, methylate ribosomal rRNA. METTL5 deposits m^{6}A1832 in 18S rRNA and is important for fine-tuning translation (Ignatova et al., 2020; Rong et al., 2020), while ZCCHC4 methylates 28S subunit at position 4220 and is required for global translation activity (Ma et al., 2019). In the next subchapters, I will focus exclusively on mRNA m^{6}A methyltransferases.

5.6.1.1. METTL3/METTL14

Search for the mRNA m^{6}A methyltransferase started together with the discovery of m^{6}A modification on mRNA in the 1970s. However, only in 1994 it was shown that mRNA m^{6}A methylation activity was mediated by a megadalton protein complex comprised of 30 kDa (MT-A1), 200 kDa (MT-A2) and 875 kDa (MT-B) components (Bokar et al., 1994). When the whole complex was crosslinked with ^{3}H-SAM, only the 200 kDa fraction co-purified with a 70 kDa protein, suggesting that it contains the methyltransferase protein, which is part of a bigger, 200 KDa complex (Bokar et al., 1994). The methyltransferase, initially called MT-A70, was cloned soon after, allowing for antibodies

production and more detailed analysis in the subsequent years (Bokar et al., 1997). Further studies showed that MT-A70 is not only present in mammals, but is conserved in yeast and plants, where it also deposits m^6A on mRNA. First, the *S. cerevisiae* homologue of MT-A70, IME4, was described to deposit m^6A on yeast mRNA, which is essential for proper sporulation (Clancy et al., 2002). Later, the plant homologue MTA was shown to methylate mRNA and be critical for embryonic plant development (Zhong et al., 2008).

The next breakthrough came in 2012 when two studies showed widespread m^6A mRNA methylation in human cells (Dominissini et al., 2012; Meyer et al., 2012). Soon after, it was reported that almost all m^6A mRNA methylation in mammals is deposited by MT-A70 (renamed METTL3), which forms a complex with METTL14 protein, and that m^6A methylation is essential for embryonic stem cells differentiation (Liu et al., 2014; Wang et al., 2014). In 2016 the first structures of the catalytic core of METTL3/14 heterodimer were published, showing that METTL3 is a catalytic subunit, while METTL14 lacks catalytic activity and is needed for the complex stabilisation and mRNA binding (Śledź and Jinek, 2016; Wang et al., 2016a). METTL3/14 adds m^6A co-transcriptionally, with the lower speed of RNA polymerase II increasing the methylation rate (Slobodin et al., 2017). It preferentially methylates GGACU and GGACA sequences, with 50% lower efficiency for GAACU and GGAUU (Wang et al., 2016b), explaining why almost all m^6A sites are found within the RRACH motives (Dominissini et al., 2012; Meyer et al., 2012).

Although METTL3/14 alone is active *in vitro*, it requires a much bigger complex to methylate its targets *in vivo*. The first identified component of the complex was Wilms' tumour 1-associating protein (WTAP), which is required for METTL3/14 localization in the nuclear speckles (Ping et al., 2014). The current view is that METTL3 interacts with WTAP through its N-terminal domain (Schöller et al., 2018), and WTAP is needed for linking METTL3/14 with the rest of the complex. Analysis of WTAP proteome revealed, among many, interactions with VIRMA, RBM15/15B,

ZC3H13 and HAKAI proteins (Horiuchi et al., 2013). VIRMA interacts with CPSF5 and CPSF6 polyadenylation and cleavage factors and is important for preferential mRNA methylation close to 3′ UTRs. Its depletion led to a significant drop in mRNA m^6A levels (Yue et al., 2018). Recently it was proposed that VIRMA forms a scaffold for the whole complex, facilitating interaction between WTAP, RBM15/15B and HAKAI (Bawankar et al., 2021). RBM15/15B was shown to interact with WTAP and drive specificity of the METTL3/14 complex (Patil et al., 2016), while ZC3H13 to bridge WTAP with RBM15/15B and to be required for m^6A levels maintenance as well as sex determination in *Drosophila* (Knuckles et al., 2018). Finally, HAKAI is important for the stability of WTAP, RBM15/15B and VIRMA complex (Bawankar et al., 2021).

All the insights gathered over the years showed that despite the simple methods, initial biochemical characterization successfully purified components of the mammalian mRNA m^6A methyltransferase complex: the 200 kDa fraction contained METTL3/14 heterodimer complex, while the bigger fraction was most likely composed of the other components of the METTL3/14 complex: WTAP, VIRMA, RBM15/15B, ZC3H13 and HAKAI.

5.6.1.2. METTL16

Early biochemical experiments hinted at the existence of the second m^6A methyltransferase in mammalian cells, which would be specific to U6 snRNA, with methylation activity dependent on the U6 snRNA secondary structure (Shimba et al., 1995). It took over 20 years to show that U6 snRNA is methylated by METTL16 methyltransferase (Pendleton et al., 2017). Before that, METTL16 was shown to be a nuclear protein, binding (although not methylating) specific secondary structures in long non-coding RNA *MALAT1* (Brown et al., 2016). It was only later identified to methylate not only U6 snRNA, but also *Mat2a* pre-mRNA, encoding for SAM synthetase (Pendleton et al., 2017). In mammals, METTL16 comprises two domains: the N-terminal methyltransferase domain and the C-terminal vertebrate-specific regions (VCR) (Pendleton et al., 2017). Initially, it was proposed that

VCR domains act as splicing enhancers (Pendleton et al., 2017). However, recently it was shown that they are rather RNA binding domains, essential for U6 snRNA (Aoyama et al., 2020) and *MALAT1* binding (Ruszkowska et al., 2018).

5.6.1.2.1. Evolutionary conservation of METTL16

Analysis of METTL16 evolutionary conservation shows that while the methyltransferase domain of METTL16 is conserved from bacteria to humans, the C-terminal vertebrate conserved regions (VCR) seems to be unique for vertebrates (Figure 5) (Pendleton et al., 2017).

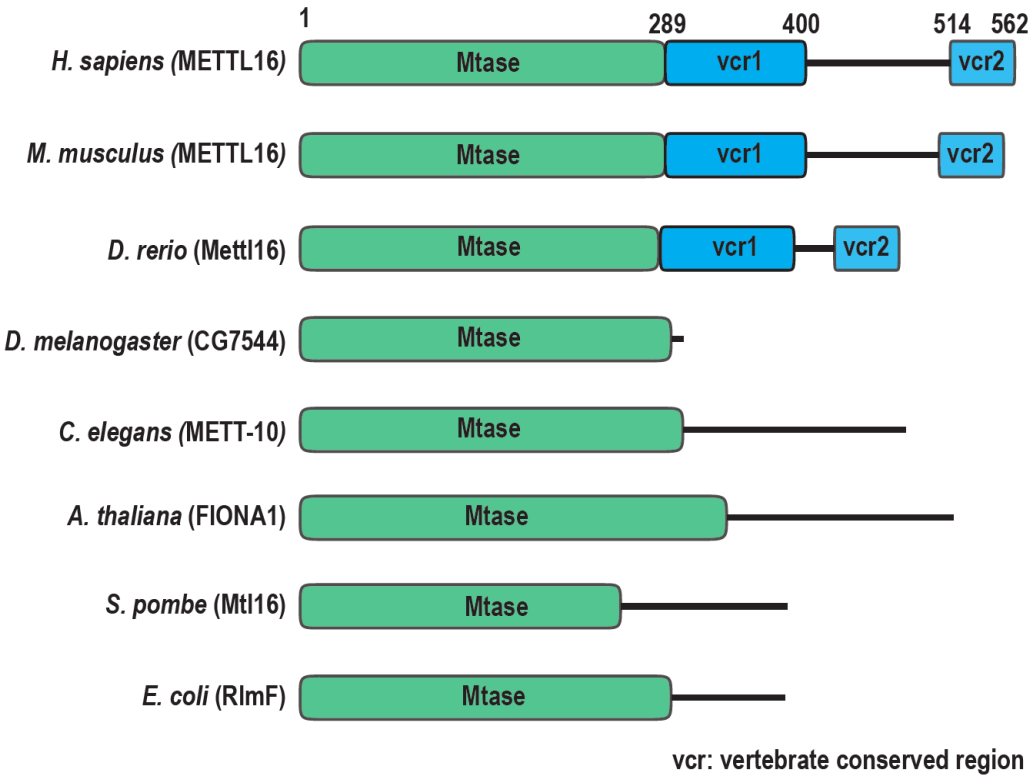


Figure 5. Evolutionary conservation of METTTL16 protein and its domain composition. (modified from (Pendleton et al., 2017)).

METTL16 homologue in *Escherichia coli,* YbiN, deposit N^6-methyladenosine at position A1618 in 23S rRNA (Sergiev et al., 2008). Changes in *ybiN* gene expression (both downregulation and overexpression) lead to moderate growth retardation and loss of cell fitness compared to the

parental strain (Sergiev et al., 2008). In yeast, METTL16 homologue is present in *Schizosaccharomyces pombe,* while it's not conserved in *S. cerevisiae,* which lacks U6 snRNA methylation. Deletion of *S. pombe* homologue, Duf890, leads to slower growth (Pendleton et al., 2017), and it was recently shown that U6 snRNA m^6A methylation by Duf890 is essential for stabilization of weak 5′ splice sites and proper splicing (Ishigami et al., 2021). In plants (*Arabidopsis thaliana*), FIONA1 is essential for the circadian clock regulation, but its targets remain unknown (Kim et al., 2008). While in *Caenorhabditis elegans,* METT-10 is needed for proper germline development (Dorsett et al., 2009), its role in *D. melanogaster* remains unknown.

5.6.1.2.2. METTL16 targets

While METTL3/14 methylates thousands of different sites in the cell characterized by a short consensus motif RRACH (Liu et al., 2014; Wang et al., 2016a), METTL16 is much more specific, recognizing much longer consensus motif UACAGAGAA placed in a context of structured RNA (Pendleton et al., 2017). A recent study showed that METTL16 binds over 400 different RNAs in HEK293T cells, including 355 mRNAs with 93% of peaks located within introns, 68 lncRNAs and nine ncRNAs (Warda et al., 2017). It is important to underline that out of these 400 proposed targets, only three targets were experimentally validated: *MALAT1* lncRNA (Brown et al., 2016), *Mat2a* mRNA and U6 snRNA (Pendleton et al., 2017). METTL16 binds, but not methylates, the triple helix sequence of *MALAT1* called an element of nuclear expression with a downstream A-rich tract (ENE+A); however, the functional role of the METTL16-*MALAT1* interaction is not known (Brown et al., 2016). When it comes to the other targets, *Mat2a* encodes for methionine adenosyltransferase (MAT) enzyme, which is essential for S-adenosylmethionine (SAM) production, while U6 snRNA is a core component of splicing machinery. Both targets will be described in more detail below.

5.6.1.2.2.1. *Mat2a* mRNA

S-adenosylmethionine (SAM, also known as AdoMet) is one of the most essential molecules in the cell, being the primary methyl donor in all living organisms as well as the precursor of aminopropyl groups in polyamine synthesis and glutathione precursor in the liver (Finkelstein, 1990). SAM is synthesized by enzymes belonging to the methionine adenosyltransferase (MAT) family, also known as S-adenosylmethionine (SAM) synthetases. There are two SAM synthetases in mammals: MAT2A is expressed in all tissues except the liver, where another isoform, MAT1A, is present (Lu and Mato, 2012). The majority of methionine from the diet is processed by MAT1A in the liver, while the rest is processed in tissues by MAT2A (Lu and Mato, 2012). While *Mat1a* KO mice are viable (Lu et al., 2001), *Mat2a* KO leads to early embryonic lethality in mice (Dickinson et al., 2016).

METTL16 was shown to be essential for the regulation of *Mat2a* mRNA splicing and stability in response to changes in SAM levels (Pendleton et al., 2017; Shima et al., 2017). METTL16 specifically recognizes and methylates six conserved hairpins (hp1 to hp6) in the last exon of *Mat2a* pre-mRNA (Pendleton et al., 2017; Shima et al., 2017). The binding of METTL16 to hp1 is essential for induction of splicing of the last intron of *Mat2a* and transcript stabilization in low SAM conditions (Pendleton et al., 2017), while methylation of hp2 to hp6 controls *Mat2a* stability, leading to nuclear degradation by YTHDC1 in high SAM conditions (Shima et al., 2017).

In low SAM conditions, METTL16 binds to the hp1on the *Mat2a* transcript and stays there waiting for the methyl group donor SAM (Pendleton et al., 2017). The increased occupancy of METTL16 on the transcript stimulates splicing through the VCR domains, which were recently proposed to attract the cleavage factor I_m complex 25 KDa subunit (CFI_m25, Nudt21, CPSF5), which then drives splicing (Scarborough et al., 2021). However, the exact mechanism of this regulation, as well as the relation between METTL16 and CFI_m25, are not clear. On the other hand, when the levels of SAM are high, METTL16 has enough substrate to methylate the hairpins quickly, there is no

splicing stimulation, and the last intron is retained (Pendleton et al., 2017). In addition, m^6A deposited on hp1 – hp6 hairpins is recognized by m^6A reader YTHDC1, which then targets the transcript for degradation (Shima et al., 2017).

5.6.1.2.2.2. U6 snRNA

The second target of METTL16 is U6 snRNA, which is the core component of splicing machinery, recruited by U1 and U2 snRNPs as a part of U4/U6•U5 tri-snRNP to form the catalytically active spliceosome (Will and Luhrmann, 2011). U6 snRNA forms the catalytic core of the spliceosome, where it coordinates the magnesium ions required for splicing chemistry (Yean et al., 2000) and, together with U2 and U5 snRNA, positions pre-mRNA for the splicing reaction. U6 snRNA basepairs with the 5′SS intron sequence through the ACAGA motif, which is conserved between yeast and humans and is essential for splicing (Kandels-Lewis and Séraphin, 1993; Lesser and Guthrie, 1993).

U6 snRNA is the most conserved of the five snRNA involved in splicing (Madhani et al., 1990). While *S. cerevisiae* and *S. pombe* have only one copy of U6 snRNA, there are over 900 copies in humans (Doucet et al., 2015). However, only four were shown to be transcriptionally active. Unlike other snRNAs, which are transcribed by RNA polymerase II and exported to the cytoplasm for maturation, U6 snRNA is transcribed by RNA polymerase III, and all the maturation steps take place in the nucleus (Reddy et al., 1987; Vankan et al., 1990). Additionally, U6 snRNA doesn't share the 2, 2, 7-trimethylguanosine cap with other snRNAs or m^7G cap with mRNAs, but instead, it possesses γ-monomethyl phosphate 5′ modification (Singh and Reddy, 1989).

Although all snRNAs are extensively chemically modified, m^6A is not commonly found, with single m^6A sites present only on U4 and U6 snRNA (Morais et al., 2021). While the enzyme modifying U4 snRNA has not yet been identified (Morais et al., 2021), METTL16 was shown to methylate U6 snRNA in humans and yeast *S. pombe* (Pendleton et al., 2017; Shimba et al., 1995).

The methylated nucleotide is located in the asymmetric bulge region of U6 snRNA (Montemayor et al., 2014) and in both organisms, it is in the middle of the conserved AC**A**GA motif (position A43 in humans, A37 in *S. pombe*) which is essential for splicing (Pendleton et al., 2017; Shimba et al., 1995).

The role of the m^6A methylation of ACAGA motif is not well understood. In *S. cerevisiae,* there is no METTL16, and ACAGA motif of U6 snRNA is not methylated. At the same time, deletion of Duf890, *S. pombe* homologue of METTL16, leads to delayed growth (Pendleton et al., 2017). It was recently shown that U6 snRNA m^6A methylation in *S. pombe* is essential for stabilising weak 5′ splice sites and proper splicing (Ishigami et al., 2021). Although the role of U6 snRNA m^6A methylation in mammals is not known, mutation of A43 position into G43 leads to lower efficiency of *in vitro* splicing in human splicing extracts, while mutations to C or U completely block splicing (Datta and Weiner, 1993).

5.6.2. m^6A readers

Methylation deposited on mRNA can be specifically recognized by "reader" proteins, which can either directly bind to the methylated residues (direct readers like YTH proteins or IGF2BPs) or bind to RNA because of m^6A-mediated changes to the RNA structure and opening of RNA-binding motifs (indirect readers like HNRNP proteins). The best-known family of proteins able to recognize m^6A modification are YTH proteins. There are five proteins belonging to the family, cytoplasmic YTHDF1 – YTHDF3 and YTHDC2, as well as nuclear YTHDC1 (reviewed in (Patil et al., 2018)).

YTHDC1, initially named YT521-B, was first identified as a nuclear protein involved in splicing modulation and localizing to specific nuclear loci (Hartmann et al., 1999). Soon after, a new domain was identified in human YT521-B and called YTH (for YT521-B homology) (Stoilov et al., 2002). The YTH domain turned out to be highly conserved across a wide species range and was shown to be an RNA-binding domain (Zhang et al., 2010). The breakthrough came with the discovery

that the YTH domain specifically recognizes m^6A methylated RNA, which was first showed using m^6A binding assays (Dominissini et al., 2012) and shortly after was confirmed by several structural studies (Luo and Tong, 2014; Theler et al., 2014; Xu et al., 2014). Over the years, YTHDC1 was shown to be involved in several different processes, including splicing regulation by facilitating SRSF3 and blocking SRSF10 binding to mRNA (Xiao et al., 2016), control of alternative polyadenylation through interaction with the 3′ processing factors SRSF3, SRSF7 and CPSF6 (CFI_m68) (Kasowitz et al., 2018) as well as export of methylated mRNA to the cytoplasm through binding to SRSF3 and NXF1 (Roundtree et al., 2017). All these functions are essential for life, as removal of YTHDC1 in mice leads to embryonic lethality past early post-implantation stages (Kasowitz et al., 2018).

The second YTHDC protein, YTHDC2, is highly enriched in testes and ovaries (Wojtas et al., 2017). It is a very unusual m^6A reader as, besides the YTH domain, it contains several RNA binding domains as well as a helicase domain (Meyer and Jaffrey, 2017). YTHDC2 was shown to be essential for the control of meiosis in flies germline (Chen et al., 2014) and mice, both males and females (Bailey et al., 2017; Hsu et al., 2017; Wojtas et al., 2017). *Ythdc2* knock-out male germ cells enter meiosis, but because of persistent cyclin A2 expression, the cells attempt an abnormal mitotic-like division, which results in cell death (Bailey et al., 2017). Analysis of *Ythdc2* knock-out testes showed upregulation of highly methylated transcripts, indicating that YTHDC2 might be involved in the degradation of m^6A-methylated transcripts (Wojtas et al., 2017). Additionally, YTHDC2 was shown to interact with XRN1 5′ - 3′ exonuclease (Kretschmer et al., 2018; Wojtas et al., 2017) and MEIOC protein, with *Meioc* KO mouse phenocopying the *Ythdc2* KO phenotype (Soh et al., 2017). However, the exact role of XRN1 and MEIOC interaction with YTHDC2 remains not understood.

Except for the YTHDC family of proteins, there is also the YTHDF family consisting of three highly similar cytoplasmic proteins: YTHDF1, YTHDF2 and YTHDF3 (also called DF1 – DF3)

(reviewed in (Meyer and Jaffrey, 2017)). YTHDF proteins consist only of the YTH domain and an extended low-complexity region, which was recently reported to drive liquid-liquid phase separation after binding to m^{6}A-methylated transcripts (Ries et al., 2019). While mouse knock-outs of *Ythdf1* (Shi et al., 2018) and *Ythdf3* (Zhang et al., 2019) are viable and fertile, lack of *Ythdf2* leads to oocyte defects and sub-lethality (Ivanova et al., 2017). There is conflicting evidence about the role and mechanism of action of different YTHDF proteins. Initially, it was reported that DF1 enhance the translation of m^{6}A methylated mRNAs (Wang et al., 2015), DF2 promotes degradation through the interaction with the CCR4-NOT deadenylase complex (Du et al., 2016), and DF3 participates in both roles, translation regulation and degradation (Shi et al., 2017). However, recently it was proposed that all the YTHDF proteins recognize the same targets and act redundantly to mediate mRNA degradation (Zaccara and Jaffrey, 2020). The model was further confirmed in mice and mouse embryonic stem cells, where it was shown that YTHDF proteins are redundant and any specific defects reported previously for various *Ythdf* mutants are caused by different tissue expression profiles of YTHDF protein family members (Lasman et al., 2020). In conclusion, based on the recent evidence, the primary role of YTHDF proteins is the degradation of m^{6}A-methylated transcripts and the individual members of the family act redundantly.

Besides the YTH family of proteins, recently proteins belonging to the insulin-like growth factor 2 mRNA binding proteins family (IGF2BPs) were shown to specifically bind m^{6}A methylated mRNAs. IGF2BPs recognize m^{6}A through the K homology domains, promoting the storage and stability of their target mRNAs (Huang et al., 2018). Another example is the eIF3 protein, which binds m^{6}A sites in the 5′ UTR and recruits the 43S complex to initiate cap-independent translation (Meyer et al., 2015). Furthermore, eIF3 was shown to interact with METTL3, leading to RNA circularization and enhanced translation (Choe et al., 2018).

Another type of readers are so-called indirect RNA readers, which binding to RNA depends on the structural changes mediated through m^6A influence on base pairing and RNA secondary structures formation. Due to steric contact, m^6A destabilizes A – U base pairing and RNA duplexes by 0.5 – 1.7 kcal/mol (Roost et al., 2015). However, if A – U base pairing is neighboured by a 5′ bulge, then m^6A stabilizes the paring by approximately 1 kcal/mol. In addition, m^6A leads to destabilisation of double-stranded RNAs having A – G and A – C pairing, although to smaller extend (Roost et al., 2015). At the same time, it stabilises A – A pairs by approximately 0.7 kcal/mol and increases stacking of adenine in unpaired context by 0.4 – 0.6 kcal/mol (Roost et al., 2015). An example of protein sensitive to m^6A-driven structural changes is heterogenous nuclear ribonucleoprotein C (HNRNPC), which regulates alternative splicing of target mRNAs based on the presence of m^6A in the vicinity of its binding motif (Liu et al., 2015).

Finally, the list of mRNA m^6A binding proteins might not be complete. A recent mass-spectrometry-based screen of proteins binding to m^6A-methylated and unmethylated RNA, detected many proteins preferentially binding to methylated RNA. At the same time, they also showed that m^6A repels certain RNA-binding proteins like G3BP1, G3BP2 or USP10 (Edupuganti et al., 2017).

5.6.3. m^6A erasers

In addition to “writer” and “reader” proteins, there are proteins capable of removing m^6A methylation from RNA called demethylases or “erasers”. In mammals, two enzymes were shown to demethylate m^6A RNA and convert it back to adenosine: fat mass and obesity-associated protein (FTO) (Jia et al., 2011) and ALKBH5 (Zheng et al., 2013). Both of the enzymes have nuclear localisation (Gerken et al., 2007; Zheng et al., 2013) and both belong to the AlkB family of Fe(II)/α-ketoglutarate-dependent dioxygenases, which remove modifications in the reaction of oxidative dealkylation (Fedeles et al., 2015). While both enzymes demethylate m^6A *in vitro*, their *in vivo* targets and roles are still debated.

The role of FTO in removing m^6A from mRNA is being continuously discussed. Although it was initially shown to demethylate m^6A mRNA (Jia et al., 2011), a subsequent study showed a very modest increase in m^6A levels in FTO-deficient mice (Hess et al., 2013). Finally, it was demonstrated that although FTO demethylates m^6A, it has over 100-times higher affinity to m^6A_m (Mauer et al., 2017). FTO-mediated demethylation of m^6A_m adjacent to m^7G cap was found to reduce the stability of mRNAs (Mauer et al., 2017). However, these results are in disagreement with the other studies showing no influence of cap m^6A_m on transcript stability, but rather on translation efficiency (Akichika et al., 2019; Sendinc et al., 2019). Finally, FTO was found to demethylate m^6A_m at the first nucleotide of snRNAs, with depletion of FTO causing a dramatic increase in m^6A_m levels at snRNAs caps, correlated with alternative splicing patterns (Mauer et al., 2019). FTO role in m^6A_m demethylation is challenged by another study showing that although *in vitro*, FTO has a higher preference for m^6A_m than m^6A, *in vivo,* it demethylates both m^6A and m^6A_m, with a higher affinity towards m^6A (Wei et al., 2018).

The other eraser, ALKBH5, was shown to be essential for proper spermatogenesis in mice with *Alkbh5* KO animals having a strongly reduced number of spermatozoa due to apoptosis of pachytene and metaphase-stage spermatocytes (Zheng et al., 2013). LC-MS/MS analysis of mRNA modifications showed increased m^6A levels in *Alkbh5* KO testes, while RNA sequencing of WT and KO testes showed thousands of differentially expressed genes (Zheng et al., 2013). ALKBH5 might also function outside of the germ cell development. It was recently demonstrated that overexpression of ALKBH in acute myeloid leukaemia (AML) promotes leukaemia stem cells self-renewal and is important for disease progression (Shen et al., 2020).

To summarize, besides the role of ALKBH5 in murine testes, the broad influence of both enzymes on m^6A levels in most cell types remains disputable. In agreement with that, a recent study analysed pre-mRNA and mRNA m^6A methylation levels in HeLa cells and found no changes in methylation

status, suggesting that demethylation of m^{6}A mRNA might not be widespread and instead be limited to specific conditions or tissues (Ke et al., 2017).

5.6.4. The role of m^{6}A in splicing

Since its discovery, it was speculated that m^{6}A might play a role in RNA processing and splicing regulation (Kane and Beemon, 1985). Studies in the 1980s and 1990s showed that general inhibition of methylation leads to accumulation of unspliced pre-mRNA both in chicken embryo fibroblasts (Stoltzfus and Dane, 1982) as well as Chinese Hamster Ovary (CHO) cells (Carroll et al., 1990). More recent studies brought further evidence of the role of m^{6}A methylation in splicing regulation. One of the first studies of transcriptome-wide m^{6}A distribution reported that silencing of m^{6}A methyltransferase METTL3 led to changes in splicing on hundreds of genes in HepG2 cells (Dominissini et al., 2012). The same effect was later shown in *Mettl3* KO mouse embryonic stem cells (ESCs) and preimplantation epiblasts, where hundreds of alternative splicing events, mostly exon skipping and intron retention, were reported (Geula et al., 2015). Another study used nascent pre-mRNA m^{6}A-seq to show that the majority of m^{6}A (over 50%) is deposited on pre-mRNA inside introns, having a splicing modulatory function. The presence of m^{6}A close to splice junctions increases splicing kinetics, while high levels of intronic m^{6}A were correlated with alternative splicing (Louloupi et al., 2018). The other studies reported much lower intronic m^{6}A levels, between 6 -10% (Ke et al., 2017; Wei et al., 2021), however still leading to perturbation of splicing events following acute depletion of METTL3 (Wei et al., 2021).

Additionally, both readers and erasers were described to be involved in the regulation of splicing. The best-studied example is the role of the YTHDC1 reader in splicing. In mice, YTHDC1 was shown to promote exon inclusion through recruiting pre-mRNA splicing factor SRSF3 and blocking SRSF10 factor binding (Xiao et al., 2016), and its deficiency in oocytes caused massive alternative splicing defects (Kasowitz et al., 2018). In *Drosophila*, binding of YTHDC1 homologue,

YT521-B, to the m^6A methylated *Sex-lethal (Sxl)* transcript leads to alternative splicing and is required for sex determination (Haussmann et al., 2016; Lence et al., 2016). Another example of m^6A reader implicated in splicing is hnRNPG, which binds to m^6A residues near splice sites and promotes exon inclusion through interaction with carboxy-terminal domain (CTD) of RNA polymerase II (Zhou et al., 2019).

When it comes to erasers, both FTO and ALKBH5 were reported to affect pre-mRNA splicing. The exact role and mechanism of FTO regulation are not yet clear, with studies reporting that FTO binds to introns, preventing exon skipping and controlling the 3′UTR length (Bartosovic et al., 2017), that FTO promotes exon skipping (Zhao et al., 2014) or that FTO regulates splicing though control of m^6A_m levels in adenosine adjacent to the snRNA cap (Mauer et al., 2019). At the same time, ALKBH5 was shown to be essential for correct splicing and production of longer 3′-UTR mRNAs in testis (Tang et al., 2017).

Finally, one of the potential ways of m^6A-driven alternative splicing could be the regulation of splicing factors binding to pre-mRNA. A recent study showed that m^6A modestly inhibits some splicing enhancers binding, while promoting splicing silencers binding (Edupuganti et al., 2017).

5.7. *C. elegans* as a model for m^6A research

Caenorhabditis elegans are microscopic (~1 mm in length), free-living nematodes (worms) found around the world, which feeds on bacteria and reproduces rapidly, with approximately 300 offspring per reproductive cycle, which take 3.5 days to mature (Meneely et al., 2019). *C. elegans* are mostly hermaphroditic and able to self-fertilize, which together with a fast developmental cycle, makes them a very easy model for genetic manipulations and screens (Meneely et al., 2019). However, for a long time, *C. elegans* seemed to be an unsuitable model for mRNA m^6A research as it lacks all the main proteins involved in m^6A biology. Worms lack homologues of METTL3/14

methyltransferase complex, FTO and ALKBH5 demethylases, as well as YTH reader proteins except of YTHDC2 homologue, F52B5.3 which, however, lacks the YTH domain (Harris et al., 2020; Howe et al., 2021). Nevertheless, analysis of *C. elegans* RNA showed that many different RNA modifications, including m^6A, were present in both short (< 200 nt) and long (> 200 nt) RNA fractions (van Delft et al., 2017). However, m^6A presence in the long RNA fraction could be attributed to both mRNA and rRNA. Indeed, another study, which analysed the influence of 13 different *C. elegans* methyltransferases on ribosomal RNA methylation, identified METL-5, a homolog of mammalian METTL5, as m^6A methyltransferase methylating adenosine 1717 on 18*S* ribosomal RNA in *C. elegans* (Liberman et al., 2020). Furthermore, a broader screen of 22 different potential RNA methyltransferases identified F33A8.4 (an ortholog of the human ZCCHC4) as the large rRNA subunit m^6A methyltransferase (Sendinc et al., 2020).

The remaining question was the presence of m^6A on mRNA in *C. elegans*. Measurements of mRNA m^6A levels by UHPLC-MS/MS showed that m^6A makes 0.0008% of *C. elegans* mRNA (Liberman et al., 2020), approximately 250-times less than 0.2% present in mammalian mRNA (Geula et al., 2015). The second study performed m^6A-IP-seq of *C. elegans* mRNA, showing no m^6A enrichment over the gene bodies, lack of specific consensus motif and no significant m^6A signal in HPLC-MS/MS, concluding that *C. elegans* mRNA lacks m^6A modification (Sendinc et al., 2020).

Nevertheless, the *C. elegans* genome has homologues of two recently described RNA methyltransferases: METT-10, being an ortholog of human METTL16, which was shown to methylate U6 snRNA and *Mat2a* mRNA (Pendleton et al., 2017) and METL-4, an ortholog of human METTL4, which was reported to methylate U2 snRNA (Chen et al., 2020; Goh et al., 2020). While METL-4 was shown to be DNA 6mA methyltransferase (Greer et al., 2015), the study was later questioned as the 6mA signal likely resulted from bacterial DNA contamination (O'Brown et al., 2019), so its function remains unknown. On the other hand, METT-10 has high homology to

METTL16 protein (Dorsett et al., 2009), making it a good candidate for worm mRNA m^6A methyltransferase.

5.7.1. METT-10 role in *C. elegans*

As mentioned previously, METTL16 is very well conserved, ranging from *E. coli* to humans. METTL16 homologue is also found in *C. elegans*, where it's called METT-10 (Harris et al., 2020). METT-10 is a nuclear protein with a well-conserved and most likely active methyltransferase domain (Dorsett et al., 2009). It was reported to play a role in worm germline development and inhibits specification of germ-cell proliferative fate (Dorsett et al., 2009). Germ cells in m*ett-10* KO worms could not correctly progress through the mitotic cell cycle, with enlarged, diffuse nuclei in proliferative zone indicating cell cycle arrest, leading to a sterile phenotype when animals are raised at 25°C (Dorsett et al., 2009). METT-10 is expressed in many different cell types, including intestine, vulva, spermatheca, somatic gonad, germline, oocytes and pachytene germ cells (Dorsett et al., 2009). The nuclear localization of METT-10 is driven by both a nuclear localization signal (NLS) and an interaction with dynein light chain 1 (DLC-1), which promotes its nuclear accumulation as well as protein expression (Dorsett and Schedl, 2009). Interestingly, expression of *mett-10* transgenic construct with mutations in NLS and DLC-1 binding motif almost completely rescued *mett-10* KO phenotype, whereas expression of *mett-10* construct having a mutation in the methyltransferase catalytic domain failed to do so. Although these results underline the importance of METT-10 methylation activity (Dorsett and Schedl, 2009), it is still unknown why METT-10 methylation is important and what are the methylated targets.

5.7.2. Splicing in *C. elegans*

Although the entire splicing machinery and the mechanism of splice site recognition are conserved between worms and humans, there are some important differences. In *C. elegans*, unlike mammals,

there are two types of splicing: *trans*-splicing, which joins exons from two different transcripts, and *cis*-splicing, joining exons from the same primary transcript (Blumenthal, 2012). Trans-splicing takes place only for the first intron, while all the subsequent ones are splicing in *cis.* In addition, *C. elegans* introns lack BP sequences and have different 3′SS sequences, with a highly conserved U_4CAG/R sequence instead of a long polypyrimidine tract and YAG sequences found in mammals (Blumenthal and Thomas, 1988). Similarly to mammals, 3′SS is recognized by $U2AF^{65}$ and $U2AF^{35}$ proteins, where $U2AF^{65}$ binds to the short U stretch and $U2AF^{35}$ contacts the 3′SS directly (Hollins et al., 2005). Mutation of the uridine residues upstream of the 3′SS reduce proper splice site recognition, whereas a strong uridine stretch was driving splicing even in the case of 3′SS YAG mutations (Zhang and Blumenthal, 1996). These results suggest that, similarly to mammals, two types of splice sites, strong and weak, exist in C. elegans, potentially allowing for regulation and alternative splicing (Hollins et al., 2005; Zhang and Blumenthal, 1996).

6. Aim of the study

The initial aim of my thesis was to understand the biological role of a recently identified m^6A methyltransferase METTL16 as well as to understand the structural features, both at the level of protein and the RNA substrate, which allow METTL16 to specifically recognize U6 snRNA and *Mat2a* mRNA. In the second part of my work, I intended to understand the role of the METTL16 methyltransferase domain alone, using *C. elegans* as a research model as non-vertebrates lack the VCR region. The initial observation that METTL16 homologue in *C. elegans*, METT-10, methylates 3′SS of SAM synthetase transcripts leading to splicing inhibition had encouraged me to expand my goals further to understand the mechanisms of m^6A-driven splicing inhibition and conservation of this mechanism in mammals.

7. Results

7.1. Chapter I – The role of METTL16 in Mouse Embryonic Development

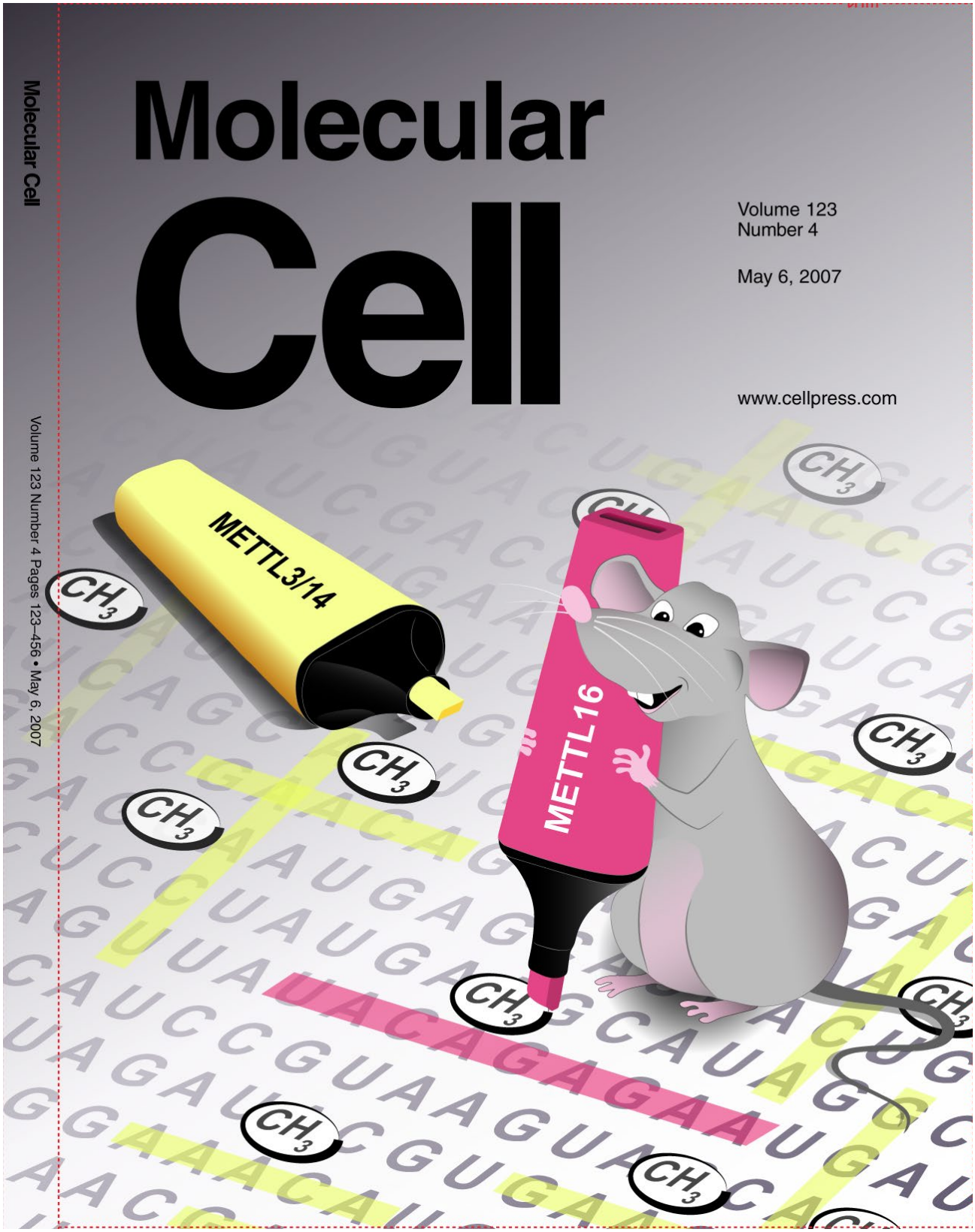
This chapter consist of a peer-reviewed article entitled "Methylation of Structured RNA by the m6A Writer METTL16 Is Essential for Mouse Embryonic Development", published in Molecular Cell journal in September 2018. In this study, we showed that METTL16 is required for early embryonic development in mice due to its role in S-adenosylmethionine (SAM) levels regulation. In addition, we solved the crystal structure of the METTL16 methyltransferase domain. We showed that the N-terminal region (1 – 78 aa) is important for RNA binding, while the loop region (187 – 223 aa) is necessary for RNA methylation, but not binding. Finally, we analysed RNA substrate requirements of METTL16 and showed that it methylates structured RNA, having METTL16 consensus motif inside the loop region. While mutations of the loop region are detrimental for methylation activity, the exact sequence of the stem region is not important as long as the pairing is preserved.

I contributed to this project by starting this project with Prof. Pillai and being its lead author. I maintained the mouse colony and organised all mouse experiments. E6.5 embryos were isolated with the help of Leonardo Beccari, E8.5 embryos were isolated by myself. I planned and performed the isolation of E2.5 and E3.5 embryos with help from Pascal Gos and Olivier Fazio. E2.5 and E3.5 RNA-seq libraries were prepared in Genomics Core Facility at EMBL Heidelberg. I did IP-MS to identify METTL16 complexes as well as almost all biochemical assays (except Fig. 2A) using recombinant proteins produced by Kuan-Ming Chen. The *Mettl16* KO mouse line and RNA-seq libraries was generated by Raman Radha Pandey, METTL16 protein and crystals were obtained by Kuan-Ming Chen, the crystal structure was solved by Andrew McCarthy, all bioinformatics analysis

was done by David Homolka. The manuscript was written by Ramesh Pillai, with my and other authors input. I was involved in editing the manuscript at every stage of the publication process.

The manuscript was not modified for the purpose of this thesis, and thus, the figure numeration and bibliography are separate from the rest of the thesis.

The graphic on the next page is our cover design proposal originally submitted to the journal. The cover was created by Margot Riggi (Twitter: @MargotRiggi).

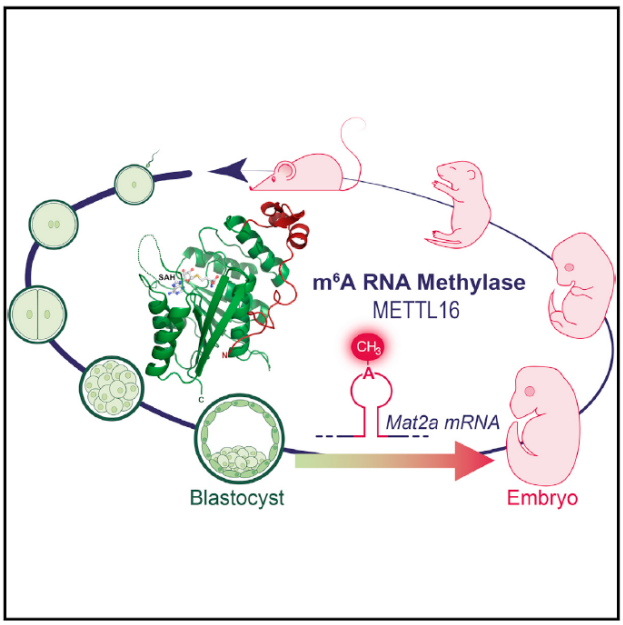

Molecular Cell
Molecular
Cell
Volume 123
Number 4
May 6, 2007
www.cellpress.com
Volume 123 Number 4 Pages 123–456 • May 6, 2007
METTL3/14
METTL16
CH3

Article

Molecular Cell

Methylation of Structured RNA by the m^{6}A Writer METTL16 Is Essential for Mouse Embryonic Development

Graphical Abstract



Authors

Mateusz Mendel, Kuan-Ming Chen, David Homolka, Pascal Gos, Radha Raman Pandey, Andrew A. McCarthy, Ramesh S. Pillai

Correspondence

ramesh.pillai@unige.ch

In Brief

Mendel et al. reveal the structural basis for structured RNA recognition by the mammalian m^{6}A writer METTL16 and demonstrate its essential role in mouse early embryonic development via regulation of the SAM synthetase *Mat2a* mRNA.

Highlights

- Structure of the METTL16 m^{6}A writer domain with a unique N-terminal module
- N-terminal module of METTL16 is essential for charge-based binding to RNA
- METTL16 preferentially methylates adenosines within structured RNAs
- Regulation of *Mat2a* mRNA by *Mettl16* is essential for mouse embryonic development

Mendel et al., 2018, Molecular Cell *71*, 1–15
September 20, 2018
https://doi.org/10.1016/j.molcel.2018.08.004

CellPress

Please cite this article in press as: Mendel et al., Methylation of Structured RNA by the m^{6}A Writer METTL16 Is Essential for Mouse Embryonic Development, Molecular Cell (2018), https://doi.org/10.1016/j.molcel.2018.08.004

Molecular Cell
Article

Methylation of Structured RNA by the m^{6}A Writer METTL16 Is Essential for Mouse Embryonic Development

Mateusz Mendel,[1,3] Kuan-Ming Chen,[1,3] David Homolka,[1] Pascal Gos,[1] Radha Raman Pandey,[1] Andrew A. McCarthy,[2] and Ramesh S. Pillai[1,4,*]

[1]Department of Molecular Biology, Science III, University of Geneva, 30 Quai Ernest-Ansermet, CH-1211 Geneva 4, Switzerland
[2]European Molecular Biology Laboratory, Grenoble Outstation, 71 Avenue des Martyrs, 38042 Grenoble, France
[3]These authors contributed equally
[4]Lead Contact
*Correspondence: ramesh.pillai@unige.ch
https://doi.org/10.1016/j.molcel.2018.08.004

SUMMARY

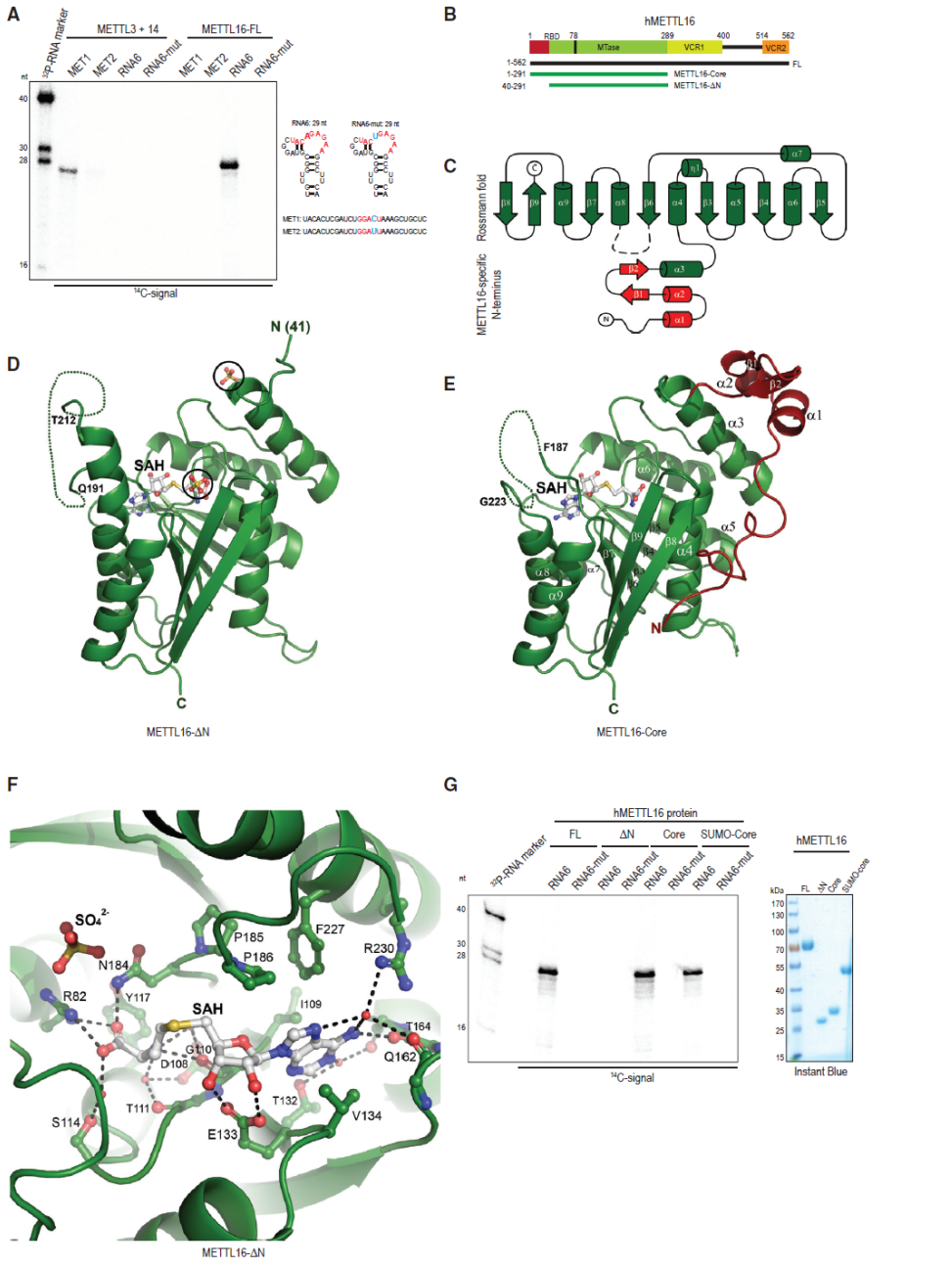
Internal modification of RNAs with N^6-methyladenosine (m^{6}A) is a highly conserved means of gene expression control. While the METTL3/METTL14 heterodimer adds this mark on thousands of transcripts in a single-stranded context, the substrate requirements and physiological roles of the second m^{6}A writer METTL16 remain unknown. Here we describe the crystal structure of human METTL16 to reveal a methyltransferase domain furnished with an extra N-terminal module, which together form a deep-cut groove that is essential for RNA binding. When presented with a random pool of RNAs, METTL16 selects for methylation-structured RNAs where the critical adenosine is present in a bulge. Mouse 16-cell embryos lacking *Mettl16* display reduced mRNA levels of its methylation target, the SAM synthetase *Mat2a*. The consequence is massive transcriptome dysregulation in ~64-cell blastocysts that are unfit for further development. This highlights the role of an m^{6}A RNA methyltransferase in facilitating early development via regulation of SAM availability.

INTRODUCTION

Methylation of adenosines at the N^6 position (N^6-methyladenosine or m^{6}A) is a highly conserved internal RNA modification with a huge impact on gene regulation (Fu et al., 2014). The modification is added by methyltransferase "writers" and can be removed by RNA demethylase "erasers," and a major part of its functions is mediated by YTH domain "reader" proteins that can recognize the m^{6}A mark. Readers of the m^{6}A modification are shown to modulate mRNA splicing, RNA export, RNA stability, and translation (Patil et al., 2018). Alterations in RNA structure are also a consequence of m^{6}A methylation (Liu et al., 2015). The m^{6}A pathway is physiologically important, as mutations in the writer protein METTL3 in mice lead to embryonic lethality (Batista et al., 2014; Geula et al., 2015), while in flies it affects sex determination (Haussmann et al., 2016; Lence et al., 2016). The only nuclear reader protein YTHDC1 is essential for early embryonic development, and its conditional deletion causes infertility in the germline where it acts via modulation of splicing and alternative polyadenylation site usage (Kasowitz et al., 2018). Loss of the cytoplasmic reader YTHDF2 in fish impairs embryonic development as a result of defective maternal RNA clearance during maternal-zygotic transition (Zhao et al., 2017), while in mice loss of YTHDF2 results in defective maternal RNA metabolism during oocyte maturation, leading to female-specific infertility (Ivanova et al., 2017). In contrast, mouse YTHDC2 is essential for proper progression through meiosis and fertility in both sexes (Bailey et al., 2017; Hsu et al., 2017; Jain et al., 2018; Wojtas et al., 2017). Thus, gene regulation by m^{6}A plays a critical role in a variety of developmental processes.

The heterodimeric m^{6}A writer complex METTL3/METTL14 co-transcriptionally (Knuckles et al., 2017; Slobodin et al., 2017) installs this mark on thousands of transcripts in the cell (Dominissini et al., 2012; Schwartz et al., 2013). While METTL3 is the active component, METTL14 facilitates substrate RNA binding (Śledź and Jinek, 2016; Wang et al., 2016a, 2016b). METTL16 is the second m^{6}A methyltransferase identified, and its known substrates include U6 snRNA and the human *MAT2A* mRNA that encodes for *S*-adenosylmethionine (SAM) synthetase (Pendleton et al., 2017). SAM is a methyl donor for methylation reactions in the cell, including those of DNA, RNA, and protein. While METTL3 prefers to methylate single-stranded RNAs (ssRNAs) in a sequence context RR<u>A</u>CH (R = A or G; H = A,C or U), METTL16 uses structured RNAs carrying a specific nonamer sequence (UAC<u>A</u>GAGAA; methylated adenosine is underlined) (Pendleton et al., 2017). Methylation of *MAT2A* mRNA within specific hairpin structures in the 3′ UTR is proposed to be used by YTHDC1 to mediate downregulation of the mRNA under high-SAM conditions (Shima et al., 2017). Apart from this enzymatic role, METTL16 is also reported to act as a splicing enhancer during low-SAM conditions when it occupies its binding site on the six *MAT2A* hairpins (hp) to promote splicing of a 3′ terminal intron that is frequently retained. This results in increased mature *MAT2A* mRNA production and acts as a feedback loop ensuring optimal production of the SAM synthetase in

Please cite this article in press as: Mendel et al., Methylation of Structured RNA by the m^{6}A Writer METTL16 Is Essential for Mouse Embryonic Development, Molecular Cell (2018), https://doi.org/10.1016/j.molcel.2018.08.004



(legend on next page)

response to low SAM levels (Pendleton et al., 2017). Unlike the METTL3/METTL14 complex which mainly methylates exonic sequences (Ke et al., 2017), METTL16 was shown to have binding sites on several intronic sequences in pre-mRNAs and structured noncoding RNAs, some of which carry m^{6}A marks (Brown et al., 2016; Warda et al., 2017). How METTL16 recognizes its RNA substrates and the physiological importance of having a second m^{6}A methyltransferase is currently not known.

METTL16 is a highly conserved enzyme with orthologs found in *E.coli* (Sergiev et al., 2008) to human (Figure S1A). Here, we examine the crystal structure of the methyltransferase (MTase) domain from human METTL16 and identify key features that are essential for RNA binding and methylation activity. We define the RNA substrate requirements *in vitro* using a randomized RNA library to find that structured RNAs with a bulged adenosine are preferred. Finally, we generate a knockout *Mettl16* mouse mutant to show that the protein is essential for early embryonic development. Our studies show that METTL16 is essential for embryonic development around implantation stage and acts via regulation of the *Mat2a* mRNA which encodes the SAM synthetase.

RESULTS

Crystal Structures of the Human m^{6}A Methyltransferase METTL16

We produced the recombinant full-length (FL) human METTL3/METTL14 heterodimeric complex and FL human METTL16 (1–562 aa) in a eukaryotic expression system (Figures S1B and S1C; Star Methods). Together with the methyl donor *S*-adenosylmethionine (SAM), the enzymes were presented with either a single-stranded RNA (ssRNA, MET1) carrying the RR**A**CH consensus site or a 29 nt hairpin RNA (RNA6) derived from the human *MAT2A* mRNA, carrying the nonamer methylation site for METTL16 (UAC**A**GAGAA) (Table S1) (Pendleton et al., 2017). While the METTL3/METTL14 complex efficiently methylated the ssRNA, it did not use the hairpin RNA as a substrate (Figure 1A). On the contrary, METTL16-FL methylated only the hairpin substrate, but not the ssRNA. Both enzymes also sensed the sequence context of their respective substrates, as single nucleotide mutations within the RNA consensus sites either reduced (for METTL3/14 complex) or abolished (for METTL16) the methylation activity (Figure 1A). The METTL16-FL protein was also capable of using U6 snRNA and the full-length *MAT2A* hairpin (hp) 1 as substrates for methylation (Figure S1D). Thus, the purified m^{6}A methyltransferases are able to discriminate their respective RNA substrates *in vitro*.

To obtain structural information on METTL16, we identified stable protein domains by limited proteolysis (Figure S1E). Two constructs (core, 1–291 aa; and ΔN, 40–291 aa) encompassing the methyltransferase domain (MTase) were expressed in *E. coli* and crystallized (Star Methods) (Figures 1B and S1B). Consistent with the SAM-dependent methyltransferase activity of METTL16, both structures reveal a Rossmann fold composed of a central seven-stranded β sheet (strands β3–β9) flanked by three α helices each (helices α4–α6 and α7–α9) (Figures 1C–1E). The complexed byproduct of the SAM-dependent methylation reaction, *S*-adenosyl homocysteine (SAH), is coordinated by a hydrogen bond interaction network with the highly conserved (Figure S1A) amino acid residues R82, D108, G110, T111, S114, E133, Q162, N184, and R230 (Figures 1F, S1G, and S2A). The methyl donor SAM is presumed to fit into the same surface-exposed pocket. The catalytic residues NPPF (184–187 aa) are present in a loop positioned in close proximity to the SAH molecule (Figures 1E, 1F, and S2A). Point mutations (PP185–186AA or F187G) of these residues abolish m^{6}A RNA methyltransferase activity on a *MAT2A* hairpin substrate when tested *in vitro* (Figures 2D and S2D). All these key features define METTL16 as a SAM-dependent methyltransferase.

Three additional observations can be made from our structures. First, and most striking, is the presence of an N-terminal module (1–78 aa) in the core structure that is appended to the α4 of the Rossmann fold and that consists of three helices (α1–α3) and two short β strands (β1 and β2) (Figures 1C and 1E). This module is flexible, as shown by proteolysis, with only α3 remaining in the ΔN structure (Figures 1D and S1E). Second, the loop containing the catalytic residues NPPF (N184–F187) is part of a larger stretch of 35 amino acids that links β6 to α8, and it is disordered in both structures (dotted lines in Figures 1D and 1E). While 20 residues between Q191 and T212 are not visible in the ΔN structure, a much larger region (35 residues) between F187 and G223 lacks density in the core structure (Figures

Figure 1. Structure of Human METTL16 Reveals an N-Terminal Module Essential for Activity

(A) *In vitro* methylation assays of indicated full-length (FL) human m^{6}A methyltransferases with ^{14}C-SAM and different RNA substrates (right). Predicted structure of a short hairpin RNA (RNA6) derived from the longer *MAT2A* hairpin 1 (Pendleton et al., 2017) and its mutant (RNA6-mut) with A→U mutation of the methylated adenosine are shown. The MET1 RNA has the consensus site for methylation by the METTL3+METTL14 complex, while the MET2 RNA has a point mutation (C→U) of a conserved residue in the methylation consensus site (see Table S1). Single-stranded RNA markers (length in nucleotides, nt) are ^{32}P-end-labeled. See also Figure S1D.

(B) Domain architecture of human METTL16. RBD, RNA-binding domain (1–78 aa); MTase, methyltransferase domain; VCR, vertebrate conserved region. Boundaries of the two protein constructs crystallized in this study are indicated (in green). The ΔN version has an N-terminal deletion.

(C) Schematic view of the MTase domain. Cylinders represent α helices, and arrows represent β strands. Regions shaded in red (α1-2 and β1-2) are seen only in the METTL16-core structure and together with α3 form a separate N-terminal module.

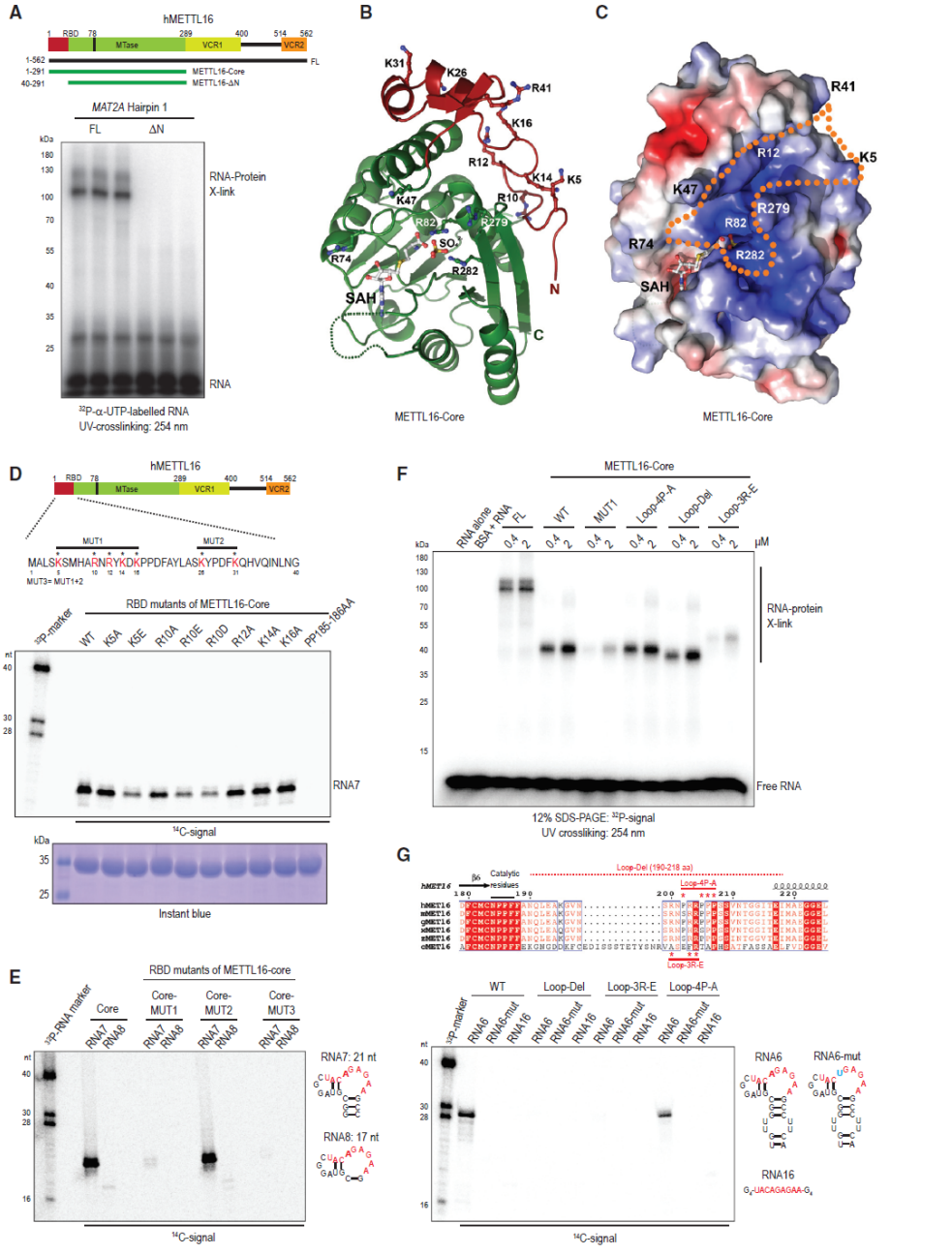
(D) Model of the METTL16-ΔN construct (PDB 6GFK). Two-sulfate (SO_4^{2-}) ions visualized in the crystal structure are circled. A disordered loop between α8 and β6 is connected by a dotted line. SAH, S-adenosyl homocysteine.

(E) Model of the METTL16-core construct (PDB 6GFN). The additional regions at the N terminus seen in this structure are shown in red. See also Figure S1G.

(F) A zoom of the catalytic pocket in the METTL16-ΔN structure showing coordination of SAH. Catalytic residues N184, P185, P186, and F187 and position of a sulfate (SO_4^{2-}) ion are indicated. See also Figure S2A.

(G) *In vitro* methylation assays showing that METTL16-ΔN protein is inactive. The METTL16-Core protein was used as untagged or tagged (SUMO) versions. See also Figure S2C. Quality of proteins used is shown on the right. Protein molecular weight markers (in kilo Daltons, kDa) are indicated.

Please cite this article in press as: Mendel et al., Methylation of Structured RNA by the m^{6}A Writer METTL16 Is Essential for Mouse Embryonic Development, Molecular Cell (2018), https://doi.org/10.1016/j.molcel.2018.08.004



(legend on next page)

1D and 1E). This loop in *C. elegans* is even longer (48 aa) and shows poor overall sequence conservation with its vertebrate orthologs (Figure S1A). Given its strategic location, it may be involved in contacting the bound RNA substrate during catalysis. Third, the ΔN structure reveals the presence of two sulfate ions (SO_4^{2-}) (Figure 1D), one of which is next to R82 (one of the residues coordinating SAH), likely mimicking how an RNA substrate might access the catalytic pocket (Figure 1F).

To test whether these structures are representative of catalytically active versions of the METTL16 MTase domain, we incubated the recombinant proteins with ^{14}C-SAM and a 29 nt RNA (RNA6) derived from the *MAT2A* hp 1. The full-length METTL16 and the core domain versions were able to methylate this RNA (RNA6) at a specific adenosine residue (A17) within the nonamer UAC**A**GAGAA motif (Figure 1G). In contrast, a mutant hairpin (RNA6-mut) with A17U mutation was not methylated. To our surprise, even though the ΔN-truncated version (40–291 aa) has a similar conformation in terms of the Rossmann fold and catalytic residues, it was inactive (Figure 1G). A similar situation is seen even when the full-length *MAT2A* hairpin 1 is used (Figure S2C). Taken together, our two structures reveal an architecture where the MTase domain is attached to a METTL16-specific N-terminal module that is essential for activity.

The N-Terminal Module of METTL16 Is Essential for RNA Binding, while a Disordered Loop Is Required for Catalysis

Given that the METTL16-ΔN domain is inactive as an MTase, we examined whether it can bind RNA substrates using UV crosslinking experiments. Incubation of METTL16-FL with a body-labeled *MAT2A* hp 1 RNA gave an RNA-protein crosslink consistent with RNA binding, while the ΔN version did not reveal such an interaction (Figure 2A). This aligns with the observed lack of methylation activity of the ΔN protein when the same RNA substrate was used (Figure S2C). Examination of the primary sequence of the N-terminal module revealed the presence of several highly conserved positively charged residues that can potentially be involved in mediating interaction with RNAs (Figure S1A). Indeed, when mapped onto the METTL16-core structure, these reveal a positively charged cluster (K5, R10, R12, K14, K16, and R41) that forms the entrance of a wide deep-cut groove (Figures 2B and 2C). With additional contributions from the N-terminal module (K47 and R74) and those from within the Rossmann fold (R82, R279, and R282), the groove runs all the way to the catalytic pocket containing the bound SAH (Figures 2B and 2C). The residues K47 and R279 serve to constrict the space within this groove, while R74 overlooks the ridge that surrounds the SAH-binding pocket. Crucially, R82 and R282, that are centrally located close to the SAH molecule itself, coordinate one of the negatively charged sulfate ions (SO_4^{2-}) that we found in the ΔN structure (Figures 1D and 2C), potentially mimicking how an RNA substrate might position itself on the enzyme.

To directly examine the role of these N-terminal residues in RNA-binding and hence catalytic activity, we individually converted positively charged residues to neutral alanine. Interestingly, mutant METTL16-core versions carrying the single point mutations K5A, R10A, R12A, K14A, and K16A did not affect RNA methylation activity (Figures 2D and S2D). However, individual mutations into a negatively charged residue (K5E, R10E, and R10D) had a more discernible impact by reducing RNA methylation activity (Figure 2D). These individual mutations did not completely abolish activity as seen in the catalytic-dead mutant PP185-186AA, indicating that these might merely reduce RNA binding. Strikingly, a combined mutant with all five residues converted to alanine (MUT1: K5A, R10A, R12A, K14A, and K16A) completely eliminated RNA methylation activity (Figure 2E). In contrast, a combined mutation of charged residues not lining the potential RNA-binding groove (MUT2: K26A and K31A) (Figure 2B) did not affect RNA methylation activity (Figure 2E). Consistently, a further mutant (MUT3) containing all the mutations made in MUT1 and MUT2 did not show any activity. To examine RNA binding, we performed UV crosslinking experiments (Figure 2F). As expected, the full-length METTL16 and the METTL16-core version bound the 5′ end-labeled RNA. In contrast, the METTL16-core version carrying the combined MUT1 mutations showed highly reduced binding. These results provide a structural rationale for absence of RNA binding and RNA methylation activity in the METTL16-ΔN protein.

Figure 2. The N-Terminal Module of Human METTL16 Is Required for Substrate RNA Binding

(A) Domain architecture of human METTL16. UV crosslinking assay (triplicate reactions) showing RNA-protein crosslinks (X-link) between human full-length (FL) METTL16 and *MAT2A* hairpin 1 RNA. See also Figure S2C for *in vitro* methylation with the same proteins and RNA.

(B) Overview of the METTL16-core MTase domain. Key positively charged residues that create the RNA-binding groove are indicated. Note that residues K26 and K31 when mutated (MUT2) do not affect activity. SAH, S-adenosyl homocysteine; SO_4^{2-}, position of a sulfate ion as seen in METTL16-ΔN is shown. The disordered loop with catalytic residues is shown as a dotted line.

(C) Surface charge representation of the METTL16-core MTase domain showing a positively charged (blue) groove (outlined) leading from the N terminus to the catalytic pocket.

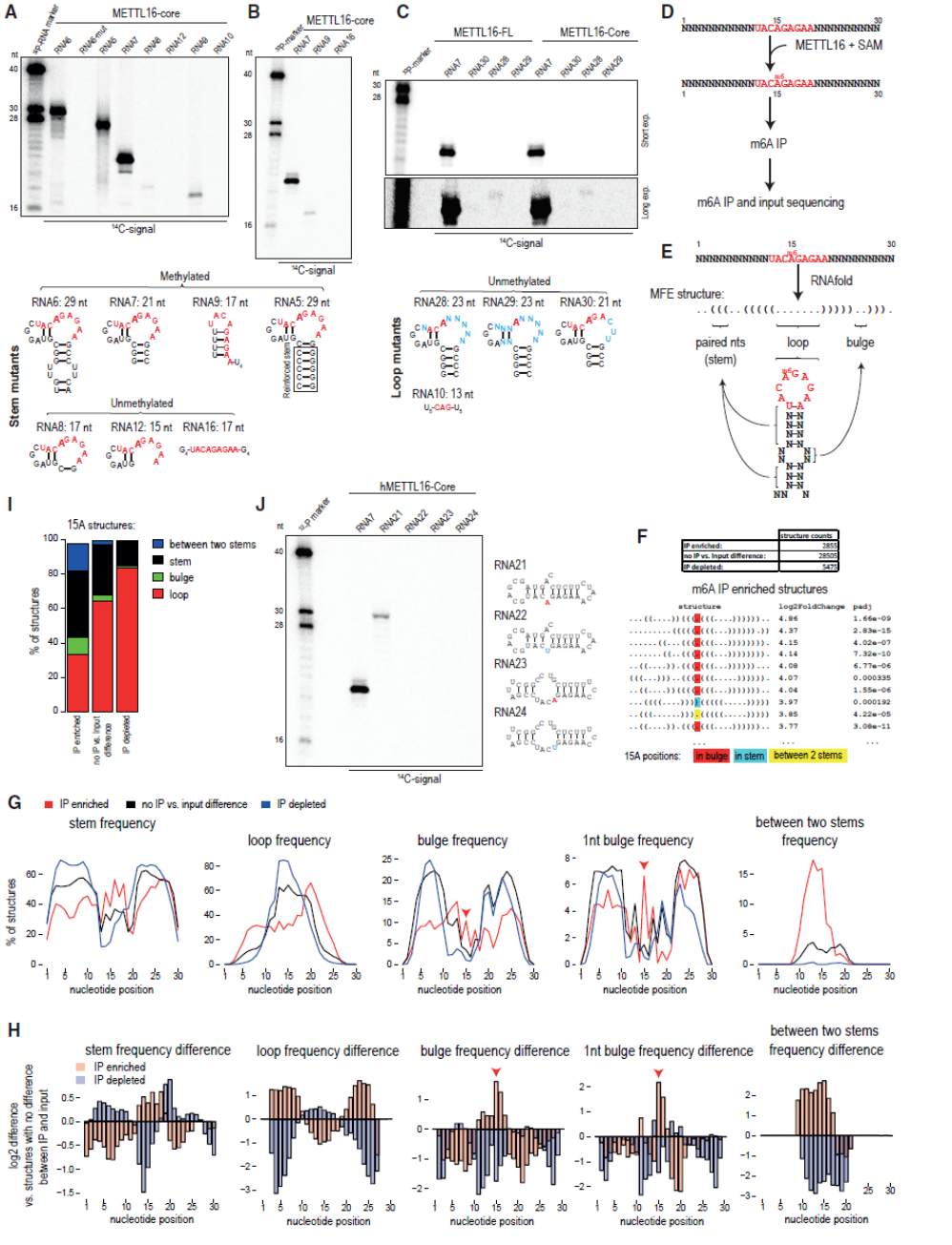
(D) Cartoon showing the N-terminal 40 amino acids of human METTL16, with the highlighted positively charged residues that were mutated (red, with asterisks). Gel shows the *in vitro* methylation assay with wild-type (WT) or indicated point mutant METTL16-core proteins. Quality of recombinant proteins used is shown below the gel. RNA7 was used as an RNA substrate (see Table S1). Single-stranded RNA markers (length in nucleotides, nt) are ^{32}P-end-labeled. See also Figure S2D.

(E) *In vitro* methylation assay with RNAs indicated and mutants carrying multiple point mutations on the N-terminal module (see D). See also Figure S2E for additional mutations within the RNA binding groove.

(F) UV crosslinking assay with METTL16 proteins indicated and ^{32}P-end-labeled RNA6. The positions of the free RNA and RNA-protein crosslinks (X-link) are shown. Control binding reactions are carried out without any protein (RNA alone) or with bovine serum albumin (BSA).

(G) Sequence alignment of METTL16 orthologs showing the catalytic residues and disordered loop region. See also Figure S1A. Deletions and mutations introduced into the loop in the context of the METTL16-core construct are indicated. *In vitro* methylation assay with indicated proteins and RNAs is shown below. Quality of proteins used is shown in Figure S2F.

Please cite this article in press as: Mendel et al., Methylation of Structured RNA by the m^{6}A Writer METTL16 Is Essential for Mouse Embryonic Development, Molecular Cell (2018), https://doi.org/10.1016/j.molcel.2018.08.004



(legend on next page)

We extended the mutational analyses to the other positively charged residues lining the putative RNA-binding groove. Mutation of the residues K47 and R279 that form a claw-like constriction of the groove either reduces (in the case of K47E) or abolishes (in R279E or double mutant K47E+R279E) methylation activity (Figure S2E). Mutation of other residues R82E, R282E, and R74E also abolishes activity, confirming their involvement in construction of the putative RNA-binding groove.

Next, we probed the importance of the disordered loop containing the catalytic residues (Figure 2G). Confirming its critical role, deletion of most of the loop (190–218 aa) abolishes *in vitro* methylation. In fact, loss of methylation can be reproduced by just three point mutations (Loop-3R-E: RRR-200-203-204-EEE) converting positive charges to negative residues, while mutation of four prolines (Loop-4P-A) within the loop did not affect methylation activity (Figure 2G). Interestingly, deletion of the disordered loop did not have an adverse effect on RNA binding, as measured by UV crosslinking (Figure 2F). However, the Loop3R-E mutant displayed highly reduced RNA binding, probably due to repulsion of RNA. These results indicate that the loop, per se, is not required for RNA binding, but has a potential catalytic role by directly contacting the RNA for proper positioning within the catalytic pocket.

In conclusion, our structure-informed mutagenesis study traces an RNA-binding groove lined by positively charged residues contributed by the N-terminal module and the MTase domain itself. This facilitates RNA binding, and thus promotes RNA methylation activity. In addition, we identify a disordered loop that is essential for catalysis.

METTL16 Prefers Structured RNAs as Substrates for m^{6}A Methylation *In Vitro*

The two known methylation targets of METTL16 are structured RNAs: the U6 snRNA and *MAT2A* hp 1 (Pendleton et al., 2017; Warda et al., 2017). To identify the RNA features that can allow *in vitro* methylation by METTL16, we carried out truncations/mutations of the *MAT2A* hp1. The *MAT2A* hp RNAs with reductions in the stem region beyond three base pairs fail to get methylated (Figures 3A, S2B, S3A, and S3B) (Pendleton et al., 2017). Activity can be restored by an artificial six base pair C/G stem (RNA5), indicating the critical requirement for any stem region (Figures 3A and S1D). On the other hand, presence of the nonamer motif in a single-stranded context, when flanked by a run of Gs (RNA16), did not support methylation (Figure 3B). Interestingly, an RNA where the nonamer motif is flanked by a run of Us (RNA9) is a substrate (albeit weak) for METTL16, likely because it has the potential to form a structured feature (Figures 3A, 3B, and S3B). To test the importance of the nonamer itself, we introduced three mutations within this motif (RNA30), which abolished activity (Figure 3C). Additional mutations, in the form of randomizations of the motif sequences (RNA28 and RNA29), also abolished activity (Figure 3C). Lastly, an RNA with only the central CAG flanked by Us (RNA10) was inactive, reaffirming the importance of this nonamer motif for methylation (Figure 3A). Taken together, these studies indicate that the nonamer sequence in the context of a secondary structure feature is essential for METTL16 methylation activity *in vitro*.

To probe the secondary structure requirements in an unbiased manner, we carried out *in vitro* methylation reactions with recombinant full-length METTL16 and a library of randomized 30 nt single-stranded RNAs carrying a central nonamer motif (Figure 3D). Subsequently, a part was retained as input while the rest was used for immunoprecipitation (IP) of methylated RNAs with the anti-m^{6}A antibody (Figure 3D). Sequences were identified by deep sequencing and preferred secondary structures of these sequences were examined using RNAfold (STAR Methods; Table S2) (Figure 3E). We identified ~2,800 predicted structures (referring to unique dot bracket notations) to be enriched in the m^{6}A-IP library (Figure 3F).

To find out whether specific structural features are important for methylation, we focused on selected secondary structural features (i.e., stems, loops, bulges, and nucleotides lying between two stems) and compared the frequencies of these features for individual nucleotide positions between the m^{6}A-enriched, non-enriched, and m^{6}A-depleted structures (Figures 3G and 3H), or directly between the sequenced oligos (Figures S3C and S3D). Analysis indicates that the nonamer sequence motif of m^{6}A-enriched structures occurs with a higher frequency in a paired stem region. However, the 15th nt adenosine that is expected to be methylated within the motif is often not paired

Figure 3. METTL16 Requires Structured RNA for m^{6}A Methylation

(A) *In vitro* methylation assay with RNAs carrying truncations of the stem region. See also Figure S3A. The predicted structures of RNAs used are shown below.

(B) *In vitro* methylation assay.

(C) *In vitro* methylation assay with RNAs carrying mutations in the nonamer consensus sequence (shown below). Short and long exposures of the gel are shown.

(D) Scheme of an *in vitro* methylation experiment using a library of randomized (N = any of the four nucleotides) RNA oligos.

(E) For each sequence, we predicted the minimum free energy (MFE) secondary structure using RNAfold (STAR Methods). A model structure is shown in dot bracket notation.

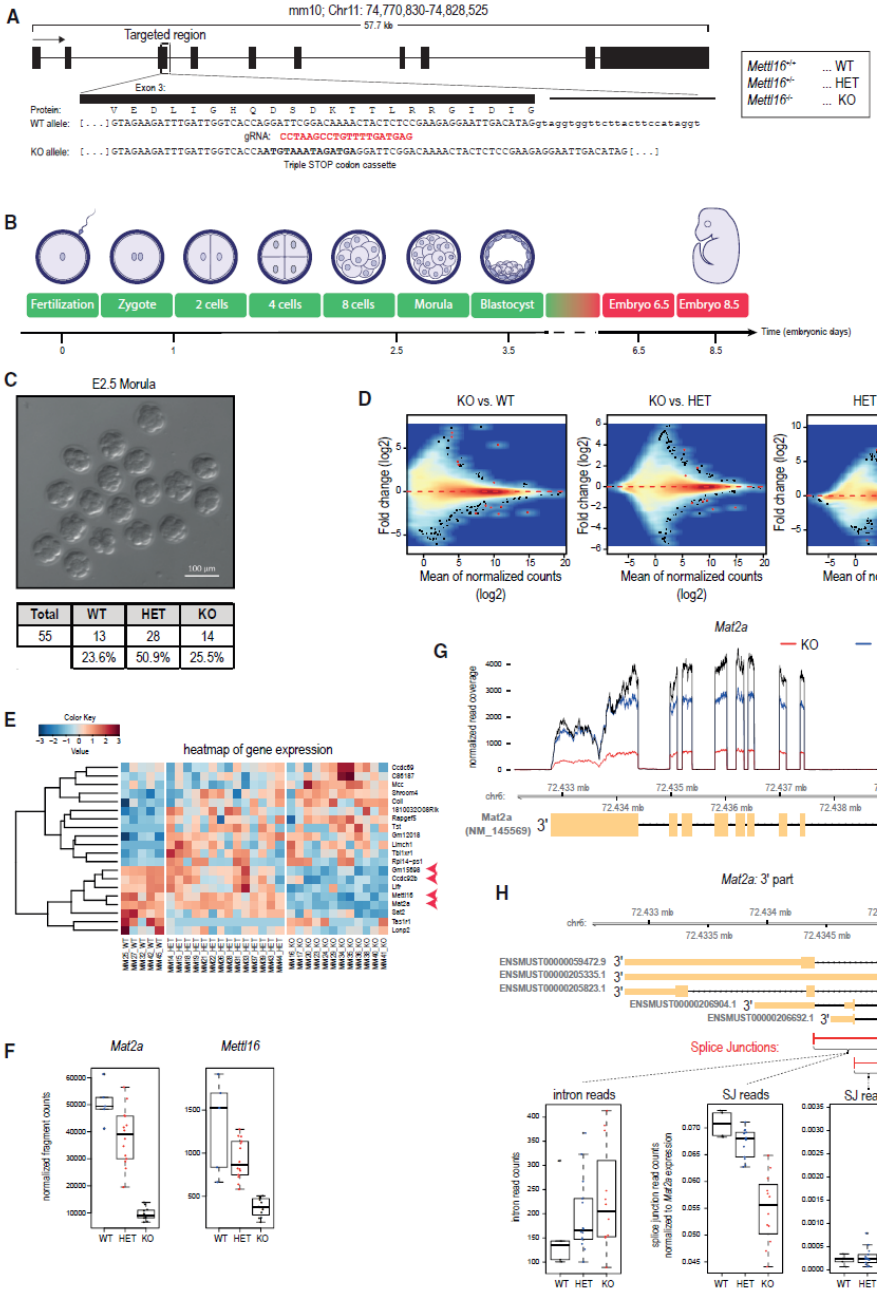
(F) The representation of individual structures (corresponding to unique dot bracket notation) was compared between m^{6}A-IP samples and input samples. Top ten IP-enriched structures are shown. The 15th position adenosine (A) that is in the consensus nonamer sequence is highlighted.

(G) Frequency of structures forming stem, loop and other selected features at individual positions is shown. The IP-enriched structures have increased frequency of 15A (red arrowhead) in a bulge and surrounded by double-stranded regions (stems), pointing to specific structural requirements of RNA substrates for METTL16.

(H) Structures enriched or depleted in m^{6}A IP were compared to those that do not show such difference (between IP and input). While the IP-enriched structures have higher proportion of 15A (red arrowhead) forming a bulge or lying between two stems, the IP-depleted structures show the opposite trend, with lack of structures with 15A in a bulge or in between two stems.

(I) The barplot shows the proportion of structural features in which the 15A was found. Note the high proportion of structures with 15A in the bulge and between two stems, among the m^{6}A IP-enriched structures. See also Figure S3.

(J) *In vitro* methylation assay with METTL16-core protein and RNAs (RNA21 and RNA23) selected from randomized library methylation experiment (D). This confirms the specific methylation of 15A which is in a 1 nt bulge (in RNA21, but not in mutant RNA22).



(legend on next page)

and more frequently present in a 1 nt bulge. There is also a higher frequency for the unpaired 15A to lie between two stems (Figures 3G–3I). m^{6}A IP-enriched and IP-depleted sequences show an opposing trend affirming the importance of an unpaired adenosine surrounded by local double-strands for methylation (Figure 3H). Indeed, an RNA (RNA 21) representative of the structure enriched in the m^{6}A-IP (15th position adenosine in a bulge between two stems) supports methylation by METTL16 (Figure 3J). This is specific as methylation is abolished when the 15th position A is mutated to U (RNA22) (Figure 3J). However, another RNA (RNA23) did not show any methylation activity (Figure 3J). These results show that the nonamer motif does not necessarily have to adopt a loop structure for activity but that the target adenosine must be unpaired and surrounded by stems. Taken together, we reveal that besides the sequence motif there is a structural requirement for a nonamer to serve as a substrate for METTL16 methylation.

Mouse METTL16 Regulates Embryonic *Mat2a* mRNA Levels and This Is Essential for Embryonic Development

To uncover the endogenous targets of METTL16, we created a knockout allele (*Mettl16*$^{-}$) by inserting a triple-stop codon cassette into exon 3 of the mouse *Mettl16* genomic locus (Figures 4A, S4A, and S4B; STAR Methods). Heterozygous (HET) *Mettl16*$^{+/-}$ mice of both sexes are viable and fertile. Intercrosses between them provided litters that were completely devoid of any homozygous (KO) *Mettl16*$^{-/-}$ animals, indicating potential embryonic lethality (Figure S4C).

After fertilization of the oocyte by the sperm, the 1-cell zygote goes through mitotic divisions to take it through a totipotent 2-cell stage, followed later by the pluripotent 16-cell morula seen at embryonic day 2.5 (E2.5) and 32- to 64-cell blastocyst at E3.5. Subsequently, the embryo becomes implanted in the uterine wall and proceeds into post-implantation development (Figure 4B). To examine the embryonic arrest in the *Mettl16* mutant, we first isolated E2.5 morula from superovulated *Mettl16*$^{+/-}$ females crossed with *Mettl16*$^{+/-}$ males (Figure 4C). Visual examination revealed no apparent differences in the embryos at this stage. Examination of E3.5 blastocysts resulted in a similar conclusion (Figure 5A). Indeed, genotyping of individual embryos confirmed the presence of all genotypes in the expected Mendelian ratios at both E2.5 and E3.5 (Figures 4C and 5A). However, examination of post-implantation embryos at E8.5 and E12.5 indicated a total absence of the knockout genotype, but E6.5 KO embryos (1.9%) were detected at sub-Mendelian ratios (Figures S4D–S4F). These results indicate that the *Mettl16*$^{-/-}$ knockout mutation allows embryonic development until blastocyst stage but causes developmental arrest around the time of implantation.

To evaluate the impact of loss of METTL16, we sequenced the transcriptomes of individual 16-cell morulas at E2.5 originating from heterozygous *Mettl16*$^{+/-}$ crosses. Embryos were genotyped based on presence or absence of specific *Mettl16* reads (STAR Methods), and gene expression levels were compared between different genotypes (Figure 4D). We find 20 genes to be differentially expressed between the different genotypes (WT, HET, and KO) (Figure 4E; Table S3). However, only four genes are consistently different in the KO embryos when compared to both WT and HET (marked with red arrowheads in Figures 4E and S5B). Examination of transcript changes in the individual embryos reveals an expected and consistent downregulation of *Mettl16* in the KO embryos (Figure 4F). Strikingly, the most significantly dysregulated transcript was *Mat2a*, which displays a 5-fold downregulation in the KO embryos. Two additional transcripts *Ccdc92b* and *Gm15698* also display significant downregulation in the KO embryos (Figure S5C).

Identification of *Mat2a* as a downregulated transcript in the *Mettl16* knockout embryos is interesting, as it is already an established target of METTL16 in human cell lines (Pendleton et al., 2017). METTL16 was proposed to promote splicing of the terminal intron, failure of which leads to intron-retention and transcript degradation (Pendleton et al., 2017). Examination of the read count distribution over the exons and introns of *Mat2a* reveals that while the exonic reads are consistently decreased in the KO, we did not observe any dramatic change in intronic reads (Figures 4G and S5D). The same was true for the two other transcripts downregulated in the KO embryos (Figure S5E). Nevertheless, a closer examination around the terminal intron of the *Mat2a* indicates a differential usage of splice junctions in the KO accompanied by a slight increase in the terminal intronic

Figure 4. Reduced *Mat2a* mRNA Levels and Embryonic Lethality around Implantation Stage in the *Mettl16* Mutant Mice

(A) Generation of a *Mettl16* knockout (KO) allele. See also Figure S4A and STAR Methods.

(B) Timeline of mouse embryogenesis. Embryonic day 2.5 (E2.5) embryos referring to 16-cell morula stage, E3.5 blastocysts, and E6.5 and E8.5 embryos were collected for genotyping. KO embryos were detected in expected Mendelian ratios till E3.5 (colored in green), but at sub-Mendelian ratios at E6.5 or none beyond (colored in red). See also Figures S4C–S4F.

(C) Genotyping of E2.5 embryos from *Mettl16*$^{+/-}$ x *Mettl16*$^{+/-}$ crosses confirmed the expected Mendelian ratios among the genotypes. Scale bar in μm is indicated.

(D) Transcriptome of individual isolated E2.5 embryos of *Mettl16*$^{-/-}$ (KO), *Mettl16*$^{+/-}$ (HET), and *Mettl16*$^{+/+}$ (WT) was sequenced and compared between the genotypes. The MA plots show a very limited number of differentially expressed genes (red dots, adjusted $p \leq 0.1$). See also Figure S5.

(E) Heatmap shows the expression of genes with significant differential expression between any two genotypes (adjusted $p \leq 0.1$). Genes differentially expressed in *Mettl16*$^{-/-}$(KO) when compared to both *Mettl16*$^{+/-}$ (HET) and *Mettl16*$^{+/+}$ (WT) are marked by red arrowhead.

(F) The boxplots show the expected downregulation of the targeted gene (*Mettl16*) in KO samples, as well as the downregulation of *Mat2a*. Transcript levels of individual samples are shown as dots. See also Figure S5C.

(G) Normalized read coverage along the *Mat2a* locus demonstrates the overall depletion in the KO. Note that the gene is on the Crick strand, so it goes from right to left.

(H) Lack of METTL16 results in aberrant splicing of the last intron. The reads spanning the splice junction (SJ) of last *Mat2a* (ENSMUST00000059472.9) intron are significantly depleted in the KO even when normalized to overall *Mat2a* transcript levels. This is accompanied by slight increase for intron reads and increased usage of alternative 3′ splice-site characteristic for the ENSMUST00000206904.1 and ENSMUST00000206692.1 variants. See also Figure S5D.

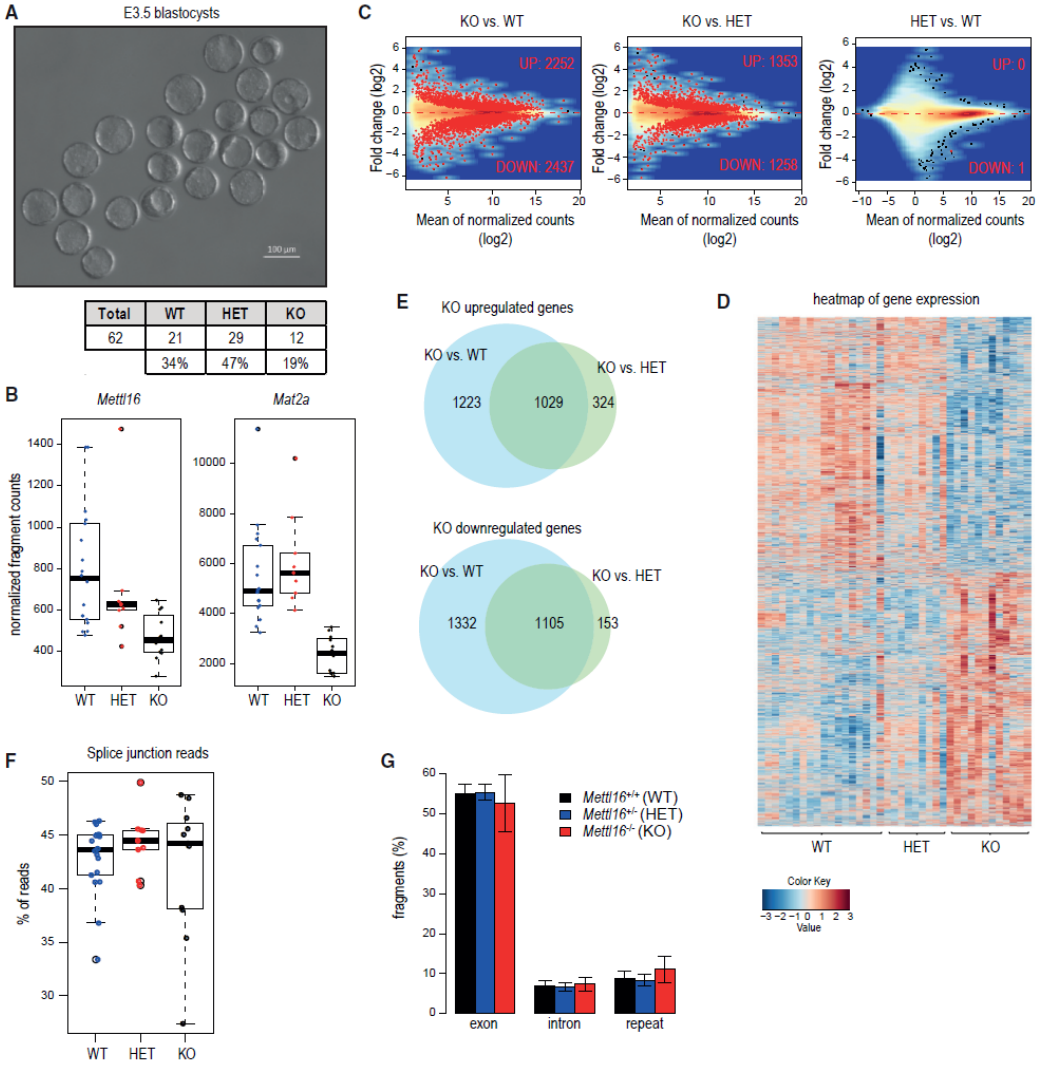


Figure 5. E3.5 *Mettl16*$^{-/-}$ Blastocysts Display Normal Morphology but Vast Transcriptome Dysregulation

(A) E3.5 *Mettl16*$^{-/-}$ KO embryos display normal morphology and their counts from *Mettl16*$^{+/-}$ x *Mettl16*$^{+/-}$ crosses correspond to expected Mendelian ratios among the genotypes. Scale bar in μm is indicated.

(B) The boxplots show the expected downregulation of the targeted gene (*Mettl16*) in KO samples, as well as the downregulation of *Mat2a*. Transcript levels of individual samples are shown as dots. See also Figure S6A.

(C) MA plots comparing the expression between the genotypes reveal that the vast number of genes are dysregulated in the *Mettl16*$^{-/-}$ KO embryos. The genes with significantly different expression are shown as red dots (adjusted $p \leq 0.1$).

(D) Heatmap shows the expression of 5,166 genes with significant differential expression between any two genotypes (adjusted $p \leq 0.1$). Note the massive dysregulation in the KO embryos. See also Figures S6B–S6D.

(E) Venn diagrams compare the lists of dysregulated genes when *Mettl16*$^{-/-}$ expression is compared to *Mettl16*$^{+/-}$ or to *Mettl16*$^{+/+}$.

(F) Comparison of proportion of reads encompassing splice junctions does not reveal a difference in splicing between individual genotypes.

(G) Global transcription from exons, introns, and repeats is not affected in *Mettl16*$^{-/-}$. Error bars refer to SD.

reads (Figure 4H). In conclusion, we demonstrate that *Mettl16* is essential for viability of early mouse embryos where it regulates the levels of *Mat2a* mRNA.

Loss of METTL16 Leads to Dramatic Alterations in the E3.5 Blastocyst Transcriptome

To examine whether the downregulation of very few transcripts in E2.5 embryos has further consequences in the E3.5 blastocysts, we collected such embryos from superovulated *Mettl16*$^{+/-}$ females crossed with *Mettl16*$^{+/-}$ males (Figure 5A). Sequencing of single embryos revealed the expected downregulation of *Mettl16* and *Mat2a* (Figure 5B). Strikingly, ~5,000 other transcripts were either upregulated or downregulated in the KO, when compared to the WT embryos, while up to half that number was altered in the KO versus HET comparison (Figures 5C, 5D, and S6D). Examination of these altered-gene lists indicates that up to 1,000 genes are either commonly up- or downregulated in the KO when compared to both WT and HET embryos (Figure 5E). A previous study identified key transcription and chromatin factors that define specific developmental stage transcriptomes (Mohammed et al., 2017). Examination of these factors in our datasets did not reveal any altered expression profile between the genotypes (Figures S6B and S6C). A Gene Ontology (GO) term analysis of the altered transcripts revealed an upregulation in splicing-related factors (Table S3), but analysis of splice junction reads did not reveal any changes in the KO embryos (Figure 5F). We also did not observe any dramatic changes in the representation of exon, intron, and repeat reads in the different libraries (Figure 5G). Taken together, even though the molecular effect of the loss of METTL16 is already seen in E2.5 embryos in the form of reduced mRNA levels of its methylation target *Mat2a*, its consequences are amplified in the E3.5 KO embryos. Here, a massive dysregulation of gene expression is observed, such that mutant embryos undergoing implantation are doomed to fail in further development.

DISCUSSION

Crystal structures now reveal how two RNA methyltransferases are built to recognize distinct RNA targets and install the same m^{6}A mark. The two methyltransferase (MTase) domains in the heterodimeric METTL3/METTL14 complex interact to create a narrow groove lined with conserved positively charged residues into which single-stranded RNAs can fit (Śledź and Jinek, 2016; Wang et al., 2016a, 2016b). This interaction facilitates stabilization of a large "interface loop" in METTL3 that contributes to the catalytic activity. Indeed, this ensures that METTL3, which binds SAM, is not active on its own, requiring at least the MTase domain of METTL14 to complete the creation of a functional catalytic complex (Śledź and Jinek, 2016; Wang et al., 2016a). In contrast, we show here that METTL16 is active as a monomer (Figures 1 and S1B), and it contains a large deep-cut groove that can accommodate structured RNAs (Figure 2). Interestingly, the METTL3/METTL14 crystal complex with the two MTase domains is inactive and requires the two N-terminal CCCH zinc finger motifs of METTL3 to recover methylation activity (Śledź and Jinek, 2016; Wang et al., 2016a), presumably because it aids in substrate RNA binding. Similarly, here we demonstrate that the N-terminal module attached to the MTase of METTL16 is essential for RNA-binding and catalysis (Figures 1 and 2). We note that the recently reported crystal structure of the human METTL16 core MTase domain (PDB 6B92) (Ruszkowska et al., 2018) shows a high degree of overlap with the one studied here (Figure S1F; Star Methods).

A structural comparison of the human METTL3/METTL14 complex (PDB 5IL2) with that of our human METTL16-core (PDB 6GFN) reveals similarity to METTL3 in the overall Rossmann fold (Figure 6A). It also shows how the disordered loop in METTL16 (Figures 1E and 2G) is very similar to the "gate loop 1" in METTL3, as both harbor the catalytic residues and are likely involved in contacting the bound RNA during enzymatic reaction. Our mutational studies indicate that the disordered loop in METTL16 is not required for RNA binding (Loop-Del in Figure 2F) but is essential for catalytic activity (Figure 2G). Thus, its role might be to contact the substrate bound via the RNA-binding groove and orient it for catalysis. This is supported by our finding that mutation of positively charged arginine (R) residues in the loop to glutamic acid (E) abolishes RNA binding (Figure 2F), perhaps via charge repulsions. We modeled a structured RNA (tRNA from PDB 2ZZM) into this groove, and it shows how an unpaired adenosine in the loop region might reach into the catalytic pocket for methylation (Figure 6B). To get better insight into the catalytic mechanism, we modeled a methyl-acceptor adenosine (from PDB 4ZCF, chain B) (Gupta et al., 2015) into the SAH binding site of METTL16-core (PDB 6GFN) (Figure 6C). Superimposition of the METTL16-core structure with that of the m^{6}A DNA MTase, EcoP1GI (PDB 4ZCF, chain B), reveals how the adenosine is favorably positioned by coordination with catalytic residues N184 and P185 for the methyl transfer from SAM (represented by SAH in Figure 6C). However, our experiments do not reveal how METTL16 might be able to recognize an adenosine within a specific nonamer sequence for m^{6}A methylation. This information will be forthcoming only when structures with bound RNA become available.

Regulation of gene expression by m^{6}A is essential at multiple steps during mouse embryonic development. The writer *Mettl3* is essential for embryonic development, with *Mettl3*-deficient embryonic stem cells (ESCs) failing to exit pluripotency despite differentiation cues (Batista et al., 2014; Geula et al., 2015; Wang et al., 2014). Now we show that the writer *Mettl16* is also essential for embryonic development around implantation stage (Figures 4 and 5). Our biochemical studies and *in vivo* transcriptome profiling reveals severe sequence and structural constraints on potential RNA targets of METTL16 (Figures 3 and 4). Although a few hundred transcripts carrying the nonamer sequence motif exist in the mouse genome, we did not find any differences in their levels in *Mettl16* knockout E2.5 embryos (Figure 4). This reinforces our finding that a combination of sequence and structural features define the target set for METTL16. The fact that *Mat2a* is the sole main target of METTL16 in pre-implantation embryos is interesting, as it encodes for the SAM synthetase, which produces SAM, the main methyl donor required for many methylation reactions (including DNA, protein, and RNA methylation) with huge regulatory potential. Before implantation, the embryonic genome undergoes massive erasure of DNA methylation marks, while during post-implantation development,

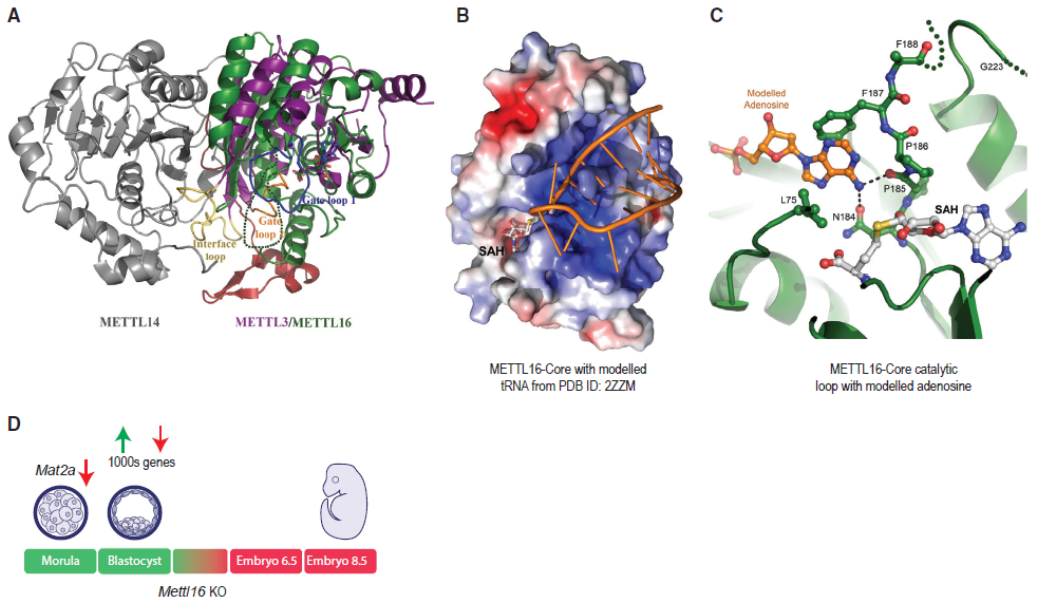


Figure 6. A Model for METTL16 Function during Early Embryonic Development

(A) Structural comparison of METTL16 core and METTL3/METTL14 complex. METTL3 (PDB 5IL2), colored purple, was superimposed on the METTL16-core-SAH (PDB 6GFN), colored in green (core) and red (N-terminal). Gate loops 1 and 2, and the interface loop of METTL3, are colored in blue, orange, and yellow, respectively. The disordered loop in our METTL16-core is shown as a dotted line.

(B) Surface charge representation of human METTL16-core domain with a modeled tRNA (from PDB: 2ZZM). See STAR Methods. The SAH bound in the catalytic pocket is shown.

(C) A methyl-acceptor adenosine (orange) was modeled into the SAH binding site of METTL16 core (PDB 6GFN) by superimposition of an m^{6}A DNA MTase, EcoP151 (PDB: 4ZCF, chain B). The sulfate binding site (as in Figure 1F) overlaps with the adenosine base moiety.

(D) A model summarizing the physiological role of METTL16 during early mouse development. The downregulation of the SAM synthetase *Mat2a* mRNA in *Mettl16* KO E2.5 morula is potentially a trigger for subsequent massive alteration in gene expression in the E3.5 blastocysts. Such mutant embryos fail to proceed further in development (indicated in red).

DNA methylation increases and is restored back to normal levels (Reik et al., 2001). Given our finding that the transcriptome in the E3.5 embryos is massively dysregulated (Figure 5), we propose that it is a snowballing effect of the initial downregulation of *Mat2a*. The low levels of *Mat2a* mRNA will mean that downstream epigenetic reprogramming events are also bound to fail. Such E3.5 mutant blastocysts are unfit for continuing in development (Figure 6D). In this context, it is interesting to point out that a homozygous *Mat2a* knockout mutation results in embryonic lethality in mice (International Mouse Phenotyping Consortium [IMPC]). Furthermore, chemical inhibition of bovine MAT2A enzyme in cultured bovine pre-implantation embryos also reduced blastocyst development (Ikeda et al., 2017).

How might METTL16 function to stabilize *Mat2a* mRNA in the mouse embryos? METTL16-mediated methylation of the hairpin structures in the 3′ UTR in *Mat2a* mRNA is used by YTHDC1 to promote its decay in high-SAM conditions (Shima et al., 2017). Thus in the absence of *Mettl16* we would have expected a stabilization of the transcript. Perhaps an explanation might come from the proposed non-canonical function of METTL16 as a splicing enhancer (Pendleton et al., 2017), where it promotes splicing to remove the 3′ terminal intron to create a stable mature *MAT2A* mRNA during low-SAM conditions. Based on this model, loss of *Mettl16* leads to reduced splicing of the terminal exon, resulting in intron-retained unstable transcripts and hence not detected in our sequencing experiments (Figures 4 and 5). Consistently, we detected higher levels of intronic reads in the *Mettl16* knockout embryos (Figure 4H). The C-terminal vertebrate conserved regions (VCRs) of METTL16 are proposed to mediate this activity (Pendleton et al., 2017). We speculate that some of the splicing factors we identified in endogenous METTL16 complexes from mouse tissues, and from transfected human cell lines, may participate in this role (Figures S5F and S5G). Future studies using a catalytic-dead METTL16 mutant mouse should help settle this issue of methylation-mediated decay versus splicing role. However, it is also possible that multiple pathways co-exist to control *Mat2a* levels.

HeLa cell extracts were originally shown to harbor an activity that adds an m^{6}A mark at position A43 in the splicing machinery component U6 snRNA (Shimba et al., 1995). This activity was

later identified to be METTL16 (Pendleton et al., 2017; Warda et al., 2017). Mutation of this methylation site in yeast U6 snRNA, which lies within a region that base pairs with the 5′ splice site of pre-mRNAs, causes lethality (Madhani et al., 1990). However, our analysis of the *Mettl16* knockout mouse embryos did not reveal any gross changes in splicing patterns across the transcriptome (Figures S5A and 5G). Thus, it is possible that methylation of U6 snRNA is not critical for splicing in higher organisms, perhaps due to the diversity of 5′ splice site sequences (Yang et al., 2013). Alternatively, complementation by another MTase activity in our mutant might account for lack of splicing defects. In conclusion, our studies place the m^{6}A writer METTL16 in a dominant position to influence early developmental decisions in the mouse embryo via regulation of SAM synthetase expression.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animal Work
  - *Mettl16* knockout mice
  - Genotyping
  - Mouse embryos
- METHOD DETAILS
  - Clones and constructs
  - Antibodies
  - Recombinant protein production
  - Purification of METTL3-METTL14 complex
  - Purification of METTL16
  - Limited proteolysis of hMETTL16-FL
  - Crystallization and data collection
  - Structure determination and refinement
  - *In vitro* transcription of RNA substrates for methylation assay
  - *In vitro* RNA methylation assay with METTL16
  - UV crosslinking assay
  - Cell culture and transfections
  - Isolation of human METTL16 complexes for mass spectrometry
  - Isolation of METTL16 complexes from mouse testes and spleen
  - Mass spectrometry
  - Preparation of RNA libraries
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - *In vitro* methylation with METTL16 and m^{6}A-IP-RNaseq
  - Transcriptome analysis of Mettl16 mutant mouse embryos
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and three tables and can be found with this article at https://doi.org/10.1016/j.molcel.2018.08.004.

ACKNOWLEDGMENTS

We thank Olivier Fazio and Leonardo Beccari for help with mouse embryo collection, Margot Riggi for scientific illustrations, Magdalena Wojtas for the METTL3/14 protein complex, and Johanna Mattay for critical reading of the manuscript. We thank the following facilities: iGE3 Genomics Platform, University of Geneva, and the EMBL Genomics core facility for sequencing; the Transgenic Mouse Facility, University of Geneva for mouse creation; the Functional Genomics Center, Zurich for proteomics analyses. M.M. is grateful for a PhD Fellowship from the Boehringer Ingelheim Fonds. This work was supported by grants to R.S.P. from the Swiss National Science Foundation: ERC Transfer Grant (GermMethylation), Project Grant (Origin-of-pi), and funding from the NCCR RNA & Disease. Work in the Pillai lab is supported by the Republic and Canton of Geneva.

AUTHOR CONTRIBUTIONS

M.M. performed all biochemical and mouse experiments with help from P.G.; K.-M.C produced recombinant proteins and performed structural analyses with A.A.M.; R.R.P. prepared the *Mettl16* knockout mouse and sequencing libraries; D.H. conducted all computational analysis; manuscript preparation and writing was by R.S.P. with input from everyone.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: May 7, 2018
Revised: June 22, 2018
Accepted: July 30, 2018
Published: September 6, 2018

REFERENCES

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. *2*, 28–36.

Bailey, A.S., Batista, P.J., Gold, R.S., Chen, Y.G., de Rooij, D.G., Chang, H.Y., and Fuller, M.T. (2017). The conserved RNA helicase YTHDC2 regulates the transition from proliferation to differentiation in the germline. eLife *6*, e26116.

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.

Bieniossek, C., Imasaki, T., Takagi, Y., and Berger, I. (2012). MultiBac: expanding the research toolbox for multiprotein complexes. Trends Biochem. Sci. *37*, 49–57.

Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Sharff, A., Smart, O.S., Vonrhein, C., et al. (2016). BUSTER (Global Phasing Ltd.).

Brown, J.A., Kinzig, C.G., DeGregorio, S.J., and Steitz, J.A. (2016). Methyltransferase-like protein 16 binds the 3′-terminal triple helix of MALAT1 long noncoding RNA. Proc. Natl. Acad. Sci. USA *113*, 14013–14018.

Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21.

Chen, E.Y., Tan, C.M., Kou, Y., Duan, Q., Wang, Z., Meirelles, G.V., Clark, N.R., and Ma'ayan, A. (2013). Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics *14*, 128.

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.

Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., and Baker, N.A. (2004). PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. Nucleic Acids Res. *32*, W665-7.

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.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501.

Flot, D., Mairs, T., Giraud, T., Guijarro, M., Lesourd, M., Rey, V., van Brussel, D., Morawe, C., Borel, C., Hignette, O., et al. (2010). The ID23-2 structural biology microfocus beamline at the ESRF. J. Synchrotron Radiat. *17*, 107–118.

Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m^{6}A RNA methylation. Nat. Rev. Genet. *15*, 293–306.

Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., et al. (2015). Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science *347*, 1002–1006.

Goto-Ito, S., Ito, T., Kuratani, M., Bessho, Y., and Yokoyama, S. (2009). Tertiary structure checkpoint at anticodon loop modification in tRNA functional maturation. Nat. Struct. Mol. Biol. *16*, 1109–1115.

Gupta, Y.K., Chan, S.H., Xu, S.Y., and Aggarwal, A.K. (2015). Structural basis of asymmetric DNA methylation and ATP-triggered long-range diffusion by EcoP15I. Nat. Commun. *6*, 7363.

Hahne, F., and Ivanek, R. (2016). Visualizing genomic data using Gviz and Bioconductor. Methods Mol. Biol. *1418*, 335–351.

Hartley, S.W., and Mullikin, J.C. (2016). Detection and visualization of differential splicing in RNA-Seq data with JunctionSeq. Nucleic Acids Res. *44*, e127.

Haussmann, I.U., Bodi, Z., Sanchez-Moran, E., Mongan, N.P., Archer, N., Fray, R.G., and Soller, M. (2016). m^{6}A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. Nature *540*, 301–304.

Hsu, P.J., Zhu, Y., Ma, H., Guo, Y., Shi, X., Liu, Y., Qi, M., Lu, Z., Shi, H., Wang, J., et al. (2017). Ythdc2 is an N^6-methyladenosine binding protein that regulates mammalian spermatogenesis. Cell Res. *27*, 1115–1127.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. Nat. Methods *12*, 115–121.

Ikeda, S., Kawahara-Miki, R., Iwata, H., Sugimoto, M., and Kume, S. (2017). Role of methionine adenosyltransferase 2A in bovine preimplantation development and its associated genomic regions. Sci. Rep. *7*, 3800.

Ivanova, I., Much, C., Di Giacomo, M., Azzi, C., Morgan, M., Moreira, P.N., Monahan, J., Carrieri, C., Enright, A.J., and O'Carroll, D. (2017). The RNA m(6)A reader YTHDF2 is essential for the post-transcriptional regulation of the maternal transcriptome and oocyte competence. Mol. Cell *67*, 1059–1067.

Jain, D., Puno, M.R., Meydan, C., Lailler, N., Mason, C.E., Lima, C.D., Anderson, K.V., and Keeney, S. (2018). *ketu* mutant mice uncover an essential meiotic function for the ancient RNA helicase YTHDC2. eLife *7*, e30919.

Kabsch, W. (2010). Xds. Acta Crystallogr. D Biol. Crystallogr. *66*, 125–132.

Kasowitz, S.D., Ma, J., Anderson, S.J., Leu, N.A., Xu, Y., Gregory, B.D., Schultz, R.M., and Wang, P.J. (2018). Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. PLoS Genet. *14*, e1007412.

Ke, S., Alemu, E.A., Mertens, C., Gantman, E.C., Fak, J.J., Mele, A., Haripal, B., Zucker-Scharff, I., Moore, M.J., Park, C.Y., et al. (2015). A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes Dev. *29*, 2037–2053.

Ke, S., Pandya-Jones, A., Saito, Y., Fak, J.J., Vågbø, C.B., Geula, S., Hanna, J.H., Black, D.L., Darnell, J.E., Jr., and Darnell, R.B. (2017). m^{6}A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. Genes Dev. *31*, 990–1006.

Knuckles, P., Carl, S.H., Musheev, M., Niehrs, C., Wenger, A., and Bühler, M. (2017). RNA fate determination through cotranscriptional adenosine methylation and microprocessor binding. Nat. Struct. Mol. Biol. *24*, 561–569.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. *44* (W1), W90–W97.

Lence, T., Akhtar, J., Bayer, M., Schmid, K., Spindler, L., Ho, C.H., Kreim, N., Andrade-Navarro, M.A., Poeck, B., Helm, M., and Roignant, J.Y. (2016). m^{6}A modulates neuronal functions and sex determination in Drosophila. Nature *540*, 242–247.

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.

Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature *518*, 560–564.

Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol. Biol. *6*, 26.

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.

Madhani, H.D., Bordonné, R., and Guthrie, C. (1990). Multiple roles for U6 snRNA in the splicing pathway. Genes Dev. *4* (12B), 2264–2277.

McCarthy, A.A., Barrett, R., Beteva, A., Caserotto, H., Dobias, F., Felisaz, F., Giraud, T., Guijarro, M., Janocha, R., Khadrouche, A., et al. (2018). ID30B—a versatile beamline for macromolecular crystallography experiments at the ESRF. J. Synchrotron Radiat. *25*, 1249–1260.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Cryst. *40*, 658–674.

Mohammed, H., Hernando-Herraez, I., Savino, A., Scialdone, A., Macaulay, I., Mulas, C., Chandra, T., Voet, T., Dean, W., Nichols, J., et al. (2017). Single-cell landscape of transcriptional heterogeneity and cell fate decisions during mouse early gastrulation. Cell Rep. *20*, 1215–1228.

Morin, A., Eisenbraun, B., Key, J., Sanschagrin, P.C., Timony, M.A., Ottaviano, M., and Sliz, P. (2013). Collaboration gets the most out of software. eLife *2*, e01456.

Patil, D.P., Pickering, B.F., and Jaffrey, S.R. (2018). Reading m^{6}A in the transcriptome: m^{6}A-binding proteins. Trends Cell Biol. *28*, 113–127.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods *14*, 417–419.

Pendleton, K.E., Chen, B., Liu, K., Hunter, O.V., Xie, Y., Tu, B.P., and Conrad, N.K. (2017). The U6 snRNA m(6)A methyltransferase METTL16 regulates SAM synthetase intron retention. Cell *169*, 824–835.

Picelli, S., Faridani, O.R., Björklund, A.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nat. Protoc. *9*, 171–181.

R Core Team (2017). R: A Language and Environment for Statistical Computing (Vienna, Austria: R Foundation for Statistical Computing).

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science *293*, 1089–1093.

Ruszkowska, A., Ruszkowski, M., Dauter, Z., and Brown, J.A. (2018). Structural insights into the RNA methyltransferase domain of METTL16. Sci. Rep. *8*, 5311.

Schwartz, S., Agarwala, S.D., Mumbach, M.R., Jovanovic, M., Mertins, P., Shishkin, A., Tabach, Y., Mikkelsen, T.S., Satija, R., Ruvkun, G., et al. (2013). High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell *155*, 1409–1421.

Sergiev, P.V., Serebryakova, M.V., Bogdanov, A.A., and Dontsova, O.A. (2008). The ybiN gene of Escherichia coli encodes adenine-N6 methyltransferase

specific for modification of A1618 of 23 S ribosomal RNA, a methylated residue located close to the ribosomal exit tunnel. J. Mol. Biol. *375*, 291–300.

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 N^6-adenosine methylation and mRNA degradation involving METTL16 and YTHDC1. Cell Rep. *21*, 3354–3363.

Shimba, S., Bokar, J.A., Rottman, F., and Reddy, R. (1995). Accurate and efficient N-6-adenosine methylation in spliceosomal U6 small nuclear RNA by HeLa cell extract in vitro. Nucleic Acids Res. *23*, 2421–2426.

Śledź, P., and Jinek, M. (2016). Structural insights into the molecular mechanism of the m(6)A writer complex. eLife *5*, e18434.

Slobodin, B., Han, R., Calderone, V., Vrielink, J.A., 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.

Soneson, C., Love, M.I., and Robinson, M.D. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Res. *4*, 1521.

Tickle, I.J., Flensburg, C., Keller, P., Paciorek, W., Sharff, A., Vonrhein, C., and Bricogne, G. (2017). STARANISO (Global Phasing Ltd.).

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., and Bricogne, G. (2011). Data processing and analysis with the autoPROC toolbox. Acta Crystallogr. D Biol. Crystallogr. *67*, 293–302.

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

Wang, P., Doxtader, K.A., and Nam, Y. (2016a). Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. Mol. Cell *63*, 306–317.

Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., Gong, Z., Wang, Q., Huang, J., Tang, C., et al. (2016b). Structural basis of N(6)-adenosine methylation by the METTL3-METTL14 complex. Nature *534*, 575–578.

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 *N*^{6}-methyladenosine (m^{6}A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep. *18*, 2004–2014.

Wojtas, M.N., Pandey, R.R., Mendel, M., Homolka, D., Sachidanandam, R., and Pillai, R.S. (2017). Regulation of m6A transcripts by the 3′→5′ RNA helicase YTHDC2 is essential for a successful meiotic program in the mammalian germline. Mol. Cell *68*, 374–387.

Yang, F., Wang, X.Y., Zhang, Z.M., Pu, J., Fan, Y.J., Zhou, J., Query, C.C., and Xu, Y.Z. (2013). Splicing proofreading at 5′ splice sites by ATPase Prp28p. Nucleic Acids Res. *41*, 4660–4670.

Zhao, B.S., Wang, X., Beadell, A.V., Lu, Z., Shi, H., Kuuspalu, A., Ho, R.K., and He, C. (2017). m^{6}A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. Nature *542*, 475–478.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Polyclonal rabbit anti-m^{6}A	Synaptic Systems	Cat. no. 202003; RRID:AB_2279214
Polyclonal rabbit anti-METT10D	Abcam	Cat. no. ab186012
Mouse IgG control antibody	Santa Cruz	Cat. no. sc-2025; RRID:AB_737182
Bacterial and Virus Strains		
DH10EMBacY bacterial strain	(Bieniossek et al., 2012)	N/A
Biological Samples		
PMSG	MSD Animal Health	Folligon
HCG	MSD Animal Health	Chorulon
Chemicals, Peptides, and Recombinant Proteins		
Sodium deoxycholate	Sigma	30968
Complete EDTA-free protease inhibitor	Roche	11 873 580 001
^{14}C-S-ADENOSYL-L-METHIONINE	Perkin Elmer	NEC363010UC
N^6-methyl adenosine	Sigma-Aldrich	M2780
Anti-HA Affinity Matrix	Roche	Cat. no. 11815016001; RRID:AB_390914
Critical Commercial Assays		
NEBNext Multiplex Small RNA Library Prep Set for Illumina	NEB	E7300
MinElute Gel Extraction Kit	QIAGEN	28604
MEGAshortscript T7 Transcription Kit	Life technologies	Cat. no. AM1354
Dynabeads Protein A	Life Technologies	10002D
Deposited Data		
Deep sequencing datasets	This study	GEO accession: GSE116329
All raw gel data are deposited at Mendeley Data.	This study	https://doi.org/10.17632/ny82j2ngt5.1
Structure: METTL16-core, crystal form 1	This study	PDB ID: 6GFN
Diffraction images: METTL16-core, form 1	This study	DOI:10.15785/SBGRID/578
Structure: METTL16-core, crystal form 2	This study	PDB ID: 6GT5
Diffraction images: METTL16-core, form 2	This study	DOI:10.15785/SBGRID/579
Structure: METTL16-DN	This study	PDB ID: 6GFK
Diffraction images: METTL16-DN	This study	DOI:10.15785/SBGRID/577
Experimental Models: Cell Lines		
Sf21 insect cells for protein production	Eukaryotic Expression Facility, EMBL Grenoble, France	N/A
High Five insect cells for protein production	Eukaryotic Expression Facility, EMBL Grenoble, France	N/A
Experimental Models: Organisms/Strains		
Mouse: *Mettl16 knock-out*	This study	Available from Lead Contact
Oligonucleotides		
DNA and RNA oligos		See Table S1
Recombinant DNA		
pACEBac2	Bieniossek et al., 2012	N/A
Human *Mettl16* cDNA	This study	NP_076991; NM_024086

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Cutadapt		http://journal.embnet.org/index.php/embnetjournal/article/view/200
ENRICHR	Chen et al., 2013; Kuleshov et al., 2016	http://amp.pharm.mssm.edu/Enrichr/
MEME - Motif discovery tool	Bailey and Elkan, 1994	
R	R Core Team, 2017	https://www.r-project.org
DESeq2	Love et al., 2014	N/A
Bioconductor	Huber et al., 2015	https://www.bioconductor.org/
Gviz	Hahne and Ivanek, 2016	N/A
STAR	Dobin et al., 2013	N/A
Salmon	Patro et al., 2017	N/A
tximport	Soneson et al., 2015	N/A
featureCounts	Liao et al., 2014	N/A
JunctionSeq	Hartley and Mullikin, 2016	N/A
RNAfold	Lorenz et al., 2011	N/A
Phaser	McCoy et al., 2007	http://www.phaser.cimr.cam.ac.uk/index.php/Phaser_Crystallographic_Software
XDS suite	Kabsch, 2010	http://xds.mpimf-heidelberg.mpg.de
autoPROC	Vonrhein et al., 2011	http://www.globalphasing.com/autoproc/
STARANISO	Tickle et al., 2017	http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi
Coot	Emsley et al., 2010	http://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot
BUSTER	Bricogne et al., 2016	http://www.globalphasing.com/buster
MOLPROBITY	Chen et al., 2010	http://molprobity.biochem.duke.edu
PyMOL	Molecular Graphics System, Version 1.8.6 Schrodinger, LLC	https://pymol.org/2/
PDB2PQR	Dolinsky et al., 2004	http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/
SBgrid	Morin et al., 2013	https://sbgrid.org/
Other		
Chelating Sepharose Fast Flow beads	GE Healthcare	17-0575-01
StrepTrap HP	GE Healthcare	28-9075-46
Superdex S75 10/300 GL	GE Healthcare	17-5174-01
Superdex 200 10/300 GL	GE Healthcare	17-5175-01
MethaPhor agarose	Lonza	50180

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ramesh S. Pillai (ramesh.pillai@unige.ch).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal Work

Mutant mice were generated at the Transgenic Mouse Facility of University of Geneva. The mice were bred in the Animal Facility of Sciences III, University of Geneva. The use of animals in research at the University of Geneva is regulated by the Animal Welfare Federal Law (LPA 2005), the Animal Welfare Ordinance (OPAn 2008) and the Animal Experimentation Ordinance (OEXA 2010). The Swiss legislation respects the Directive 2010/63/EU of the European Union. Any project involving animals has to be approved by the Direction Générale de la Santé and the official ethics committee of the Canton of Geneva, performing a harm-benefit analysis of the project. Animals are treated with respect based on the 3Rs principle in the animal care facility of the University of Geneva. We use the lowest number of animals needed to conduct our specific research project. Discomfort, distress, pain and injury is limited to what is

indispensable and anesthesia and analgesia is provided when necessary. Daily care and maintenance are ensured by fully trained and certified staff. This particular work was approved by the Canton of Geneva (GE/6/18).

Mettl16 knockout mice

The *Mettl16* gene locus is located on mouse chromosome 11 and consists of 10 exons (Figure S4A). We targeted the endogenous *Mettl16* locus in mouse embryos of the B6D2F1/J hybrid line (also called B6D2; The Jackson Laboratory, stock no. 100006). It is a cross between C57BL/6J (B6) and DBA/2J (D2), and heterozygous for all B6 and D2 alleles. Single-cell mouse embryos were injected with a guide RNA (gRNA) that directs the DNA endonuclease Cas9, and a 170 nt single-stranded DNA (ssDNA) repair template (IDT). The ssDNA carries a triple-stop codon flanked by a 81 nt 5′ homology arm and a 75 nt 3′ homology arm. Founder mice were identified by genotyping PCR (Figure S4B) and crossed with wild-type C57BL/6JRj (Janvier) partners to obtain germline transmission. We obtained two lines: line #2112 where the homology recombination template was inserted, resulting in a triple-stop codon cassette (sequence: ATGTAAATAGATGA) in exon 3, and line #2175 where a 7 bp deletion led to removal of a splicing donor site in intron 4. It is expected that creation of premature termination codons in both lines should result in removal of the transcripts via nonsense-mediated decay (NMD). Heterozygous *Mettl16*$^{+/-}$ mice of both sexes were viable and fertile, while homozygous mutants were not recovered in born litters (Figure S4C). Indeed, our analysis indicates that homozygous *Mettl16*$^{-/-}$ mutation results in embryonic lethality around implantation (Figures 4, 5, and S4D–S4F). Both the generated lines showed the embryonic lethality phenotype. We used the line #2112 (with the triple-stop codon cassette) for sequence analysis of early embryos.

Preparation of gRNAs: A cloning-free method was used to prepare DNA template for *in vitro* transcription of the chimeric crRNA-tracrRNA, termed single guide RNA (sgRNA or gRNA). Briefly, a common reverse primer (CRISPR sgR primer) and a gene specific forward primer (CRISPR F primer) with T7 promoter sequence was used to PCR amplify the single-stranded sgDNA template. Primer sequences are provided in Table S1.

Forward (F) primer design template:

5′-GAAATTAATACGACTCACTATAGGNNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGC-3′

N represent the gene-specific sequence.

The following components were mixed to prepare the PCR reaction: 20 μl 5X Phusion HF buffer, 67 μl ddH2O, 2 μl 10 mM dNTPs, 5 μl of 10 μM CRISPR F primer, 5 μl of 10 μM CRISPR sgR primer, and 1 μl Phusion DNA polymerase. The PCR reaction was set as follows: 98°C for 30 s, 35 cycles of [98°C for 10 s, 60°C for 30 s and 72°C for 15 s], 72°C for 10 min, and finally at 4°C to hold the reaction. The PCR product (~110bp) was agarose gel-purified using mini-elute gel extraction kit (QIAGEN, cat. no. 28604). The purified DNA was used to produce gRNA by *in vitro* transcription via the T7 promoter. *In vitro* transcription was carried out with the MEGAshortscript T7 Transcription Kit (Life technologies; cat no. AM1354) for 4 hours at 37°C. Reactions were treated with DNase I to remove template DNA, phenol-chloroform extracted and precipitated with ethanol. Quality of the generated gRNA was verified by 1.2% agarose gel electrophoresis.

Denaturing formaldehyde-agarose gel electrophoresis: Quality of generated gRNAs were verified by 1.2% agarose-formaldehyde gel electrophoresis. Agarose gel was prepared by mixing 0.6 g agarose, 36.5 mL H_2O, 5 mL of 10x MOPS buffer (0.2 M MOPS, 80 mM sodium acetate, 10 mM EDTA) and 8.5 mL of 37% formaldehyde. Approximately, 4 μg of RNA was dissolved in the 4xRNA loading buffer (50% formamide, 6.5% formaldehyde, MOPS buffer 1x, bromophenol blue 0.2%, ethidium bromide 50 μg/ml) and heated to 65°C for 10 min. RNA was loaded into the gel and run at 70V for approximately 90 minutes. Gel was imaged in the E-Box VX5 (Vilber Lourmat, France) imaging station.

Preparation of injection mix: We mixed 12.5 ng/μl of the gRNA with 12.5 ng/μl of the 170 nt ssDNA repair template (IDT), and 25 ng/μl of Cas9 mRNA (ThermoFischer Scientifique; A29378), in injection buffer (10 mM Tris pH 7.5, 1 mM EDTA, pH 8.0). Prepare aliquots of 20 μL and store at −80°C.

Sequence of ssDNA repair template used: The triple-stop codon sequence is highlighted (bold, italic).

ssDNA (negative-strand sequence)

AGTTGAGAATGCAAAACCTATGGAAGTAAGAACCACCTACCTATGTCAATTCCTCTTCGGAGAGTAGTTTTGTCCGAATCC***TCAT CTATTTACAT***TGGTGACCAATCAAATCTTCTACCCAGTGAATATAGTTAAGTCTCAAGGGGACTGTGGGAATTAGTCTCTCCAAA

Injection of mouse embryos of the hybrid background B6D2F1/J (black coat color) was carried out at the Transgenic Mouse Core Facility, University Medical Centre (CMU), University of Geneva. The B6D2F1/J hybrid line (also called B6D2; The Jackson Laboratory, stock no. 100006) is a cross between C57BL/6J (B6) and DBA/2J (D2), and heterozygous for all B6 and D2 alleles. The NMRI (Naval Medical Research Institute) mice, which have a white coat color were used as foster mothers.

Genotyping

Ear punches of the weaned animals (21 days-old) were digested in 100 μl of buffer containing 10 mM NaOH, 0.1 mM EDTA for 120 min at 95°C. After centrifugation at 3000 rpm for 10 min, 50 μl of supernatant was transferred to a new tube containing 50 μl of TE buffer (20mM Tris-HCl, pH 8.0 and 0.1 mM EDTA). An aliquot of 2 μl of the digestion mix was used for PCR.

Primers to detect bands (Figure S4B) corresponding to the wild-type (344 bp, WT), the triple-stop codon knock-in (358 bp, 2112) and 7 bp deletion (337 bp, 2175) alleles were MMoligo109 and MMoligo110 (Table S1). Identity of the bands were confirmed by Sanger sequencing.

Reaction mix for 25 μl PCR reactions: 1 × Taq buffer (without $MgCl_2$, ThermoFisher cat. no. B38), 2 mM $MgCl_2$, 0.5 μl dNTPs mix (stock 10 mM), 0.5 μl primer mix (stock 10 nM each), 2.0 μl tail DNA (100-200 ng), 0.5 μl Taq Pol (EMBL Protein Expression Facility, Heidelberg), water to make 25 μl final volume. Reactions were run using the following conditions (94°C, 20 s; 60°C, 30 s; 72°C, 30 s) for 35 cycles. Reactions were examined by 2.5% agarose gel electrophoresis (Figure S4B).

Mouse embryos

Heterozygous *Mettl16*$^{+/-}$ adult (8 weeks-old) females were superovulated by hormone injections for E2.5 and E3.5 embryo collections. Briefly, one intraperitoneal (IP) injection of five International Units (IU) per mouse (volume, 0.1ml) of pregnant mare serum gonadotropin (PMSG; Folligon, MSD Animal Health) was given two days before crossing with males (at day −2). A second IP injection of 5 IU/mouse (volume, 0.1ml) of human chorionic gonadotropin (HCG; Chorulon, MSD Animal Health) at day 0 was administered to the females. The females were mated with *Mettl16*$^{+/-}$ males immediately after the injections and checked for plugs the day-after (E0.5). The females were sacrificed 2 or 3 days later (embryonic days E2.5 or E3.5) to collect embryos at 16-cell morula and ≤ 64-cell blastocyst stages, respectively.

For single-embryo transcriptome sequencing, the isolated E2.5 and E3.5 embryos were visually examined for viability and cell number, and transferred separately into single tubes of 0.2 mL thin-walled 8-tube PCR strips (Thermo, AB-0451). The tubes contained 2 μL of the following mix: 0.4% Triton X-100 (vol/vol) in H2O + 2U/μl SUPERase• In RNase Inhibitor (20 U/μL; Thermo, AM2694). Embryos were stored at −80°C prior to processing for Smart-seq2 library preparation (Picelli et al., 2014).

For genotyping E2.5, E3.5 embryos, these were collected as above from superovulated heterozygous *Mettl16*$^{+/-}$ females and placed individually into single tubes of 0.2 mL thin-walled 8-tube PCR strips with 10 μl of lysis buffer [GoTaq G2 DNA Polymerase buffer (Promega, M7841), 200 μg/ml Proteinase K]. Embryos were lysed for 1h at 55°C, and then Proteinase K was inactivated by heating to 96°C for 10 min. 5 μl of the mix was used for PCR. Reaction mix for 20 μl: 5 x GoTaq G2 DNA Polymerase Buffer, 200 μM dNTP mix, 250 μM primers, 0.25 μl GoTaq G2 DNA Polymerase, 5 μl DNA. Reactions for oligo pair MMoligo109 + MMoligo110 were run using the following conditions (94°C, 20 s; 60°C, 30 s; 72°C, 30 s) for 35 cycles. Reactions were examined by 2% agarose gel electrophoresis.

For genotyping E6.5, E8.5 and E12.5 embryos, these were collected from heterozygous *Mettl16*$^{+/-}$ females without superovulation. *Mettl16*$^{+/-}$ females were mated with *Mettl16*$^{+/-}$ males and plugs were checked on the day-after (E0.5). Plugged animals were separated. The females were sacrificed 6, 8 or 12 days later, in the late afternoon (between 4 pm to 7 pm). After dissection, embryos were placed in 50 μl of RNA*later* Stabilization Solution (ThermoFisher, AM7020) and kept at −80°C until isolation. RNA and DNA were extracted simultaneously using DNeasy Blood and Tissue Kit (QIAGEN, 69504) and RNeasy Plus Micro Kit (QIAGEN, 74034). RNA was stored at −80°C. 2 μl of DNA was used for genotyping. Reaction mix for 20 μl: 5 x GoTaq G2 DNA Polymerase Buffer, 200 μM dNTP mix, 250 μM primers, 0.25 μl GoTaq G2 DNA Polymerase, 2 μl DNA. Reactions for oligo pair MMoligo109 + MMoligo110 were run using the following conditions (94°C, 20 s; 60°C, 30 s; 72°C, 30 s) for 35 cycles. Reactions were examined by 2% agarose gel electrophoresis. For embryos for which agarose gel electrophoresis was not conclusive, the PCR was repeated and reaction products were cloned into pCR 2.1 vector using The Original TA Cloning Kit (ThermoFisher, 45-0046). Positive clones were selected and sequenced by Sanger sequencing.

METHOD DETAILS

Clones and constructs

Constructs for mammalian cell expression

Coding sequence for full-length (FL) human METTL16 (hMETTL16; 562 aa; Accession number NP_076991) was amplified from human HeLa cell total RNA by reverse transcription-PCR (RT-PCR). A mammalian expression vector (pCI-neo vector backbone) allowing production of 3xFLAG-HA tagged proteins from a cytomegalovirus (CMV) promoter was used. Sequence of the tag: ATG GACTACAAAGACCATGACGGTGATTATAAAGATCATGATATCGATTACAAGGATGACGATGACAAGggcggcagcggcTACCCATATG ATGTTCCAGATTACGCT.

Constructs for insect cell expression

For the production of FL proteins, we used Baculovirus-mediated expression in insect cells. The full-length (1-562 aa) human METTL16 (hMETTL16) was cloned into the pACEBac2-SUMO acceptor vector (Bieniossek et al., 2012) for expression as an N-terminal 6xHis-Strep-SUMO-TEV fusion in the insect cells. For co-expression of human METTL3 and METTL4, full-length coding sequence for human METTL3 (1-580 aa) was cloned into the NheI and SphI restriction sites of the modified acceptor vector pACEBac2 to express the recombinant proteins with N-terminal 6xHis-SUMO-StrepIII-TEV fusions. The full-length coding sequence for untagged hMETTL14 (1-456) was cloned into the donor vector pIDK between KpnI/XhoI restriction sites. The proteins were co-expressed by taking advantage of MultiBac system (Bieniossek et al., 2012) which allows the generation of multi-gene constructs via Cre-lox recombination. The acceptor and donor vectors were combined in Cre-mediated reaction in total volume of 20 μl where 2 μg of each vector was mixed with 2 μl of 10x Cre buffer and 1 μl of Cre recombinase (NEB, cat no. M0298S). The reaction was incubated at 37°C for 1 h. After that, 5 μl of Cre reaction was transformed to 100 μl of competent TOP10 cells and plated on LB agar with appropriate antibiotics. The clones were verified by restriction digestion of the isolated plasmid, as well as by PCR.

Constructs for bacterial expression

Constructs covering only the core methyltransferase domain of hMETTL16 (1-291 aa) or its point mutant/deletion versions were cloned into the bacterial expression vector (pETM-11-SUMO vector; EMBL Protein Expression and Purification Core Facility) as 6xHis-Strep-SUMO-TEV fusions. The following constructs were prepared:

METTL16-ΔN: 40-291 aa, N-terminal deletion version similar to that used in PDB ID: 2H00 [Structural Genomics Consortium (SGC)].

METTL16-core: 1-291 aa.

METTL16-core mutants

1. Single amino acid changes: K5A, R10A, R12A, K14A, K16A, K5E, K10E, K10D, K47E, R74E, R82E, F187G, R279E, R282E.
2. MUT1: five residues (K5, R10, R12, K14, and K16) mutated to As.
3. MUT2: two residues (K26 and K31) mutated to As.
4. MUT3: combination of MUT1 and MUT2 sites mutated to As.
5. PP185-186AA: two residues (P185 and P186) mutated to As.
6. Loop-4P-A: four residues (P202, P205, P206, P207) mutated to As.
7. Loop-3R-E: three residues (R200, R203, R204) mutated to Es.
8. Loop-del: deletion of disordered loop 190-218 aa and replaced with a linker GGGSGGGS.
9. Double mutations in the binding groove: two residues (K47 and R279) mutated to Es.

Antibodies

The polyclonal rabbit anti-m^{6}A (Synaptic Systems; 202003), antibody for detecting mouse METTL16- polyclonal rabbit anti-METT10D (abcam, ab186012) and normal mouse IgG (Santa Cruz, sc-2025) antibodies were purchased. Anti-HA affinity matrix (Roche; cat. no. 11815016001) and Pierce HA Epitope Tag Antibody (ThermoFisher, cat.no. #26181) were used for immunoprecipitations.

Recombinant protein production

Production of full-length recombinant proteins was carried out in insect cell lines using the baculovirus expression system. The ovary-derived cell lines used are: High Five (Hi5) insect cell line originating from the cabbage looper (*Trichoplusia ni*) and the Sf9 cells derived from the fall army worm *Spodoptera frugiperda*. Briefly, recombinant full-length hMETTL16 coding sequence was cloned into pACEBac2-Sumo acceptor vector (His-Strep-Sumo tag) (Bieniossek et al., 2012). Plasmids were transformed into DH10EMBacY competent cells for recombination with the baculovirus genomic DNA (bacmid). The bacmid DNA was extracted and transfected with FuGENE HD (Promega, cat. no. E231A) into the *Sf*9 insect cells for virus production. The supernatant (V_0) containing the recombinant baculovirus was collected after 72 to 96 hours post-transfection. To expand the virus pool, 6.0 mL of the V_0 virus stock was added into 25 mL of *Sf*9 (0.5×10^6/mL) cells. The resulting cell culture supernatant (V_1) was collected 24 h post-proliferation arrest. For large-scale expression of the protein, Hi5 cells were infected with virus (V_1) and cells were harvested 72 h post-proliferation arrest.

For bacterial expression, plasmids were transformed into the *E. coli* BL21(DE3) strain and expression was initiated by addition of 0.7 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) when the culture density reached 0.6 (OD_{600}). The proteins were then expressed overnight at 20°C following induction.

Purification of METTL3-METTL14 complex

Insect cells co-expressing hMETTL3 and hMETTL14 were resuspended in the lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 40 mM Imidazole, 5% glycerol, 0.1% Triton X-100, 5 mM 2-mercaptoethanol, proteinase inhibitor (Roche, Complete EDTA-free) and Benzonase (Millipore), sonicated with MISONIX Sonicator S-4000 and the lysate was centrifuged at 20,000 rpm for 30 min at 4°C. The clarified supernatant was incubated at 4°C for 2h with the Ni^{2+} chelating Sepharose FF beads (GE Health; cat. no. 17057501). The beads were washed with buffer W300 (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 50 mM Imidazole, 0.1% Triton X-100, 5 mM 2-mercaptoethanol) and W500 (50 mM Tris-HCl pH 8.0, 500 mM NaCl, 40 mM Imidazole, 0.2% Triton X-100, 5 mM 2-mercaptoethanol). Finally, His-tag proteins bound to the beads were eluted with the elution buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 300 mM Imidazole, 0.1% Triton X-100, 5 mM 2-mercaptoethanol). The tag (His-Strep-Sumo) was cleaved overnight with TEV in the dialysis buffer (50 mM Tris-HCl pH 8.0, 250 mM NaCl, 5 mM 2-mercaptoethanol). After cleavage, second Ni-column purification was performed and supernatant containing the cleaved protein was collected. Proteins were further purified over the ion exchange column (HiTrap™ Q Sepharose HP, 1ml, GE healthcare, cat. no. 17-1153-01). Fractions containing the recombinant proteins were further purified by gel filtration chromatography using Superdex S200 10/300GL equilibrated with gel-filtration buffer containing: 50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM 2-mercaptoethanol (GE Healthcare, cat. no. 17-5175-01). The fractions eluting at 11 mL of elution volume were checked by SDS-PAGE analysis (Figure S1C) and pure hMETTL3-hMETTL14 protein complexes were concentrated and flash frozen in liquid nitrogen after addition of 10% glycerol.

Purification of METTL16

The insect cells or bacterial cells were collected by centrifugation and lysed by sonication [25 mM Tris-HCl, pH 8.0, 400 mM NaCl, 5% Glycerol, 0.5% Tween-20, 5 mM 2-mercaptoethanol, 20 mM Imidazole and protease inhibitor (Roche complete EDTA-free)]. After incubation for two hours with Ni-NTA beads, the fusion protein was eluted with Imidazole (250 mM), and the His-SUMO tag was

cut by the TEV protease (10 μg of protease per 1 mg of fusion protein; EMBL Protein expression and purification facility). The cleaved tag was removed by a second purification on Ni-NTA beads. The protein was further purified by gel filtration chromatography (Superdex S75 or Superdex 200, GE Healthcare) in the buffer (25 mM HEPES, pH 7.2, 150 mM NaCl, 2 mM DTT). The elution volumes of both full-length METTL16 and METTL16-core and METTL16-ΔN during gel-filtration chromatography are consistent with the proteins being a monomer (Figure S1B). The pure fractions were verified by SDS-PAGE (Figure S1E), and used for crystallization and biochemical assays. One of the METTL16-core mutants (Loop-3R-E) showed aberrant migration in the denaturing gel, but its identity was confirmed by mass spectrometry and shows normal elution profiles during gel-filtration chromatography (Figure S2F).

Limited proteolysis of hMETTL16-FL

For limited proteolysis, we used a 1:1000 ratio of protease:protein (if the protease is freshly prepared, use 1:500 ratio). Take 100 μL of METTL16-FL (concentration 1 μg/μl) protein solution and mix with 2 μL the protease Trypsin (concentration is 50 ng/μl). This makes a total of 102 μL reaction mix. Incubate at 25°C and remove aliquots of 25 μL at time-points 0, 5, 30 and 60 minutes. Aliquots are immediately mixed with gel loading dye, boiled at 95°C, and stored at −20°C. Reactions are then resolved via SDS-PAGE (Figure S1E). Peptide boundaries of proteolysis fragments were identified by mass spectrometry at the Proteomics Core Facility, EMBL, Heidelberg.

Crystallization and data collection

Optimal crystallization conditions for full-length human METTL16 (1-562 aa) and the human METTL16-core (1-291 aa) proteins were sought by robot screening at the High Throughput Crystallization Facility at EMBL Grenoble, France. Only the METTL16-core gave crystals in this screen. Once conditions were identified, crystals were manually produced: 2 μL protein solution at 13 mg/ml was manually mixed with 2 μL reservoir solution using the hanging drop method at room temperature. The reservoir conditions used were either 0.2 M di-sodium tartrate, 20% (w/v) PEG 3350 or 0.1 M Bis-Tris propane, pH 6.5, 0.2 M potassium-sodium tartrate, 20% (w/v) PEG 3350. We additionally crystallized the human METTL16-ΔN (40-291 aa) version using conditions previously described in PDB ID: 2H00 [Structural Genomics Consortium (SGC)]. The crystals were then flash-frozen at 100K after transferring them to identical crystallization conditions containing 20% glycerol. Diffraction data were collected on ID23-2 (Flot et al., 2010) and ID30B (McCarthy et al., 2018) at the European Synchrotron Radiation Facility (Grenoble, France), and integrated using the XDS suite (Kabsch, 2010). The diffraction data from hMETTL16-core (1-291 aa) crystals were highly anisotropic, with diffraction limits of ~2.8 Å and 2.4 Å along the best direction for crystal form 1 and 2 respectively, but only ~3.6 Å in the weakly diffracting directions. Therefore, data were processed using STARANISO (Tickle et al., 2017), as implemented in autoPROC (Vonrhein et al., 2011), which applies non-elliptical anisotropic limits based on a locally averaged mean $I/\sigma(I)$ cut-off, performs a Bayesian estimation of structure amplitudes, and applies an anisotropic correction to the data. Detailed crystallographic statistics are provided in Table 1.

Structure determination and refinement

The hMETTL16-core (1-291 aa) structure was solved by molecular replacement using the METTL16-ΔN, N-terminal deletion structure (PDB ID: 2h00) as a search model with Phaser (McCoy et al., 2007). Several rounds of manual building with Coot (Emsley et al., 2010), and structure refinement with BUSTER (Bricogne et al., 2016) were carried out for all structures. MOLPROBITY (Chen et al., 2010) was used for model validation and all the crystallographic information is summarized in Table 1. The atomic coordinates and structure factors have been deposited in the Protein Data Bank with the accession codes: 6GFN (METTL16-core, crystal form 1), 6GT5 (METTL16-core, crystal form 2) and 6GFK (METTL16-ΔN). For modeling a bound RNA into the METTL16-core structure we used a tRNA from PDB ID: 2ZZM (Goto-Ito et al., 2009). Structural figures were prepared with PyMOL (Schrödinger, LLC). The electrostatic potential was calculated using PDB2PQR (Dolinsky et al., 2004) and displayed in PyMOL using the APBS plugin. For modeling of adenosine into the METTL16-core structure we used a 20-mer DNA from the complex structure of the MTase EcoP15I (PDB: 4ZCF, chain B)(Gupta et al., 2015).

While this study was in preparation, Ruszkowska et al. reported the crystal structure (PDB ID: 6B92) of METTL16 core MTase domain (Ruszkowska et al., 2018). A comparison with our structure (PBD ID: 6GFN) reveals a very high degree of overlap (rmsd = ~0.38 Å for superimposition of 187 C_α atoms) (Figure S1F). Nevertheless, there are some differences. First, the N terminus in our structure is longer by four amino acids. Second, there are differences in labile loops between α4-β3 (96-99 disordered in ours); α7-β6 (poor density in ours); and β6-α8 (the long catalytic loop). The catalytic NPPF residues (in both our METTL16-core structures, Table 1) are more similar to their apo form (PDB ID: 6B91) than their SAH-bound form (PDB ID: 6B92) (Ruszkowska et al., 2018). They also have eight additional residues on α8, but these are not helical as in our METTL16-ΔN structure (PBDB ID: 6GFK) (Figure 1D).

In vitro transcription of RNA substrates for methylation assay

Templates for *in vitro* transcription (of full-length human *MAT2A* mRNA hairpin 1 and human U6 snRNA RNA) (Pendleton et al., 2017) were amplified in a PCR reaction to prepare a single-stranded DNA template with T7 promoter sequence. The T7 promoter sequence 5′- TAATACGACTCACTATAGGG −3′ was introduced at the 5′ end of forward primer followed by a specific sequence. The reverse primer had a 20 nt overlap with the forward primer allowing for efficient base pairing. The primers used for template preparation are given in Table S1. The following components were mixed to prepare the PCR reaction: 20 μl 5X Phusion HF buffer, 67 μl ddH2O, 2 μl 10 mM dNTPs, 5 μl of 10 μM Forward primer, 5 μl of 10 μM Reverse primer, and 1 μl Phusion DNA polymerase. The PCR reaction

Table 1. Data Collection and Refinement Statistics

Protein PDB Code	ΔN MTase 6GFK	MTase (form 1, SAH) 6GFN	MTase (form 2, apo) 6GT5
Wavelength (Å)	0.9763	0.8731	0.8731
Resolution range (Å)	46–2.3 (2.38–2.3)	82–2.86 (3.2–2.86)	80–2.5 (2.8–2.5)
Space group	$P3_121$	$I4_122$	$P4_12_12$
Unit cell (Å)	133.8, 133.8, 78.7 90, 90, 120	93.4, 93.4, 180.7 90, 90, 90	89.6, 89.6, 179.1 90, 90, 90
Unique reflections	36,074 (3,538)	5,724 (286)	12,391 (619)
Completeness (%)			
Spherical	99.4 (99.6)	60.9 (10.5)	45.9 (6.8)
Ellipsoidal	N/A	93.1 (78.7)	93.3 (79.9)
Mean < I/σI >	10.3 (1.4)	8.0 (1.8)	4.5 (1.6)
R_{pim} (%)	3.7 (55.0)	7.3 (50.3)	,12.4 (50.2)
CC*	0.994 (0.996)	0.996 (0.59)	0.967 (0.647)
R_{work} (%)	19.4 (21.2)	18.1 (22.1)	18.2 (24.0)
R_{free} (%)	22.8 (23.5)	21.6 (30.5)	23.4 (30.5)
Number of non-H Atoms			
Macromolecules	5,307	2,024	4,011
Water	100	11	25
SAH/ion	103	26	–
R_{msd} (bonds, Å)	0.009	0.01	0.009
R_{msd} (angles, °)	1.06	1.11	1.10
Ramachandran Plot (%)			
Favored	97.3	94.4	93.2
Allowed	2.7	4.4	6.0

Statistics for the highest resolution shell are shown in parentheses.

conditions were set as follows: 98°C for 30 s, 35 cycles of [98°C for 10 s, 60°C for 30 s and 72°C for 15 s], 72°C for 10 min, and finally at 4°C to hold the reaction. The PCR product (~110bp) was agarose gel-purified using mini-elute gel extraction kit (QIAGEN, cat. no. 28604). The purified DNA was used to produce RNA by *in vitro* transcription reaction via the T7 promoter. *In vitro* transcription was carried out with the MEGAshortscript T7 Transcription Kit (Life technologies; cat no. AM1354) for 4 hours at 37°C. Reactions were treated with DNase I to remove template DNA, phenol-chloroform extracted and precipitated with ethanol. *MAT2A* hairpin RNA was 82 nt long, while the U6 snRNA was 83 nt long. Quality of the generated RNA was verified by 1.2% agarose gel electrophoresis.

In vitro RNA methylation assay with METTL16

Some methylation assays were carried out with *in vitro* transcribed RNAs (*MAT2A* mRNA hairpin 1 or U6 snRNA), while the majority were with chemically synthesized RNA oligos (Microsynth, Switzerland) (Table S1). Recombinant human METTL16 proteins (FL, core, ΔN and mutant versions) or a heterodimer of human METTL3/METTL14 were used.

Prior to the experiment, the RNAs were refolded by heating 100 μM RNA solution in 10 mM NaCl in a thermoblock to 70°C for 5 min. and slowly cooling down to room temperature, while keeping the tubes in a heat block. All methylation reactions were performed in a 50 mM Tris-HCl, pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 2 mM DTT buffer with 10 μM of refolded single-stranded RNA, 5 μg of recombinant protein, 1 μl of RiboLock RNase Inhibitor (ThermoFisher, cat. no. EO0381) and 0.1 μCi of ^{14}C-SAM (Perkin Elmer, NEC363010UC) in a total volume of 50 μl. Unless otherwise indicated, all reactions were performed overnight at 37°C. For reactions with RNA oligos designed based on m^{6}A-IP-RNaseq experiment (Figure 3J), these were performed overnight at 22°C. RNA was subsequently isolated using phenol/chloroform extraction protocol. RNA pellets were resuspended in 2x RNA loading buffer (90% formamide, 0.02% SDS, 1 mM EDTA, 0.02% bromophenol blue, 0.02% xylene cyanol), heated for 5 min. at 70°C, cooled down to the room temperature and resolved in a 15% Urea-PAGE gel.

The 15% Urea-PAGE gel was prepared by mixing 12.6 g of urea, 3 mL of 10x TBE (1 M Tris base, 1 M boric acid, 0.02 M EDTA), 11.25 mL of 40% acrylamide (19:1) and 6.75 mL of H_2O. To catalyze gel polymerization, 240 μl of APS and 24 μl of TEMED were added. Gel was left for 40 min. at room temperature to polymerize. Wells were flushed with 1XTBE to remove urea deposits and gel was pre-run in 1X TBE at 20 W for 25 min to warm the gel. After the pre-run, ssRNA marker labeled with ^{32}P-γ-ATP and composed of four single-stranded RNA oligos (RP_RNA_19: 40 nt, RP_RNA_1: 30 nt, RP_RNA_3: 28 nt, RP_RNA_18: 16 nt; Table S1) was loaded into the gel, together with RNA samples from the *in vitro* methylation assay. Gel was run at 12 W for 1 h 30 min. Then, dried

in a gel dryer (Bio-Rad, model 583) with a gradual heating and cooling program, 80°C for 3 h. Dried gel was exposed to a phosphor screen BAS (GE Healthcare) for 24 h. The phosphor screen was scanned in a Typhoon FLA 9500 laser scanner (GE Healthcare) at 700V and 100 μm pixel size using control software (1.1 version) for Typhoon FLA 9500. Scans were analyzed using ImageQuant TL 8.1 software (GE Healthcare).

The quality of RNAs used for methylation assays were verified by Methylene Blue staining. In some experiments, after the methylation reaction products were resolved by urea-PAGE, the gel was stained with Methylene Blue, imaged to verify integrity of RNAs present in the reaction (Figure S3A) and then dried for exposure to the phosphor storage screen to detect radioactivity signals (Figure 3A).

UV crosslinking assay

Preparation of labeled RNA: RNA6 (100 pmol) was 5′-end labeled with [γ-^{32}P]ATP and T4 Polynucleotide Kinase (NEB, M0201) for 1 h at 37°C. Labeled RNA was resolved on 15% Urea-PAGE gel and exposed with phosphor screen BAS (GE Healthcare) for 5 minutes. RNA band corresponding to the size of 29 nt was cut from the gel. The RNA was eluted from gel by overnight incubation in 300 mM NaCl at room temperature and with shaking (750 rpm). RNA was extracted by phenol-chloroform and resuspended in 20 μl of H_2O. See Table S1 for RNA sequence.

METTL16 proteins (final concentration 0.4 μM and 2 μM) were mixed with 1 μl of labeled RNA6 in RNA binding buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl, 1 mM DTT, 1 mM SAH) in a final volume of 20 μl and incubated for 2 h on ice. After incubation, reaction mix was deposited inside the cap of the Eppendorf tube, and placed on ice such that the cap touches the ice. Tubes were placed around 4 cm from UV lamp (254 nm) and irradiated for 5 min. (UV Stratalinker 2400, Stratagene). After UV irradiation, samples were boiled for 5 min in SDS loading buffer and resolved by 12% SDS-polyacrylamide gel electrophoresis.

Gel was dried in a gel dryer (Bio-Rad, model 583) with a gradual heating and cooling program, 80°C for 3 h. Dried gel was transferred to the cassette and exposed with a phosphor screen BAS (GE Healthcare) for 24 h. After exposure, phosphor screen was scanned in a Typhoon FLA 9500 laser scanner (GE Healthcare) at 700V and 50 μm pixel size. Control software for Typhoon FLA 9500 was at 1.1 version. Scans were analyzed using ImageQuant TL 8.1 software (GE Healthcare).

Cell culture and transfections

Human embryonic kidney cell line 293 (HEK293) transformed with the SV40 large T antigen (HEK293T) were grown in Dulbecco's modified Eagle Medium (DMEM; Invitrogen, cat. No. 21969-035) supplemented with 10% fetal bovine serum (ThermoFisher; cat. no. 10270106), 1% Penicilline/Streptomycin (ThermoFisher; cat. No. 15140122), 1% 200 mM Glutamine (ThermoFisher; cat. no. 15140122), later referred to as DMEM complete medium (DMEM CM), and maintained in an environment with 5% CO_2 at 37°C. For transfection, cells growing in a 75 cm^2 flask were washed with warm (37°C) 1X PBS and incubated with 1 mL of Trypsin-EDTA 0.05% (ThermoFisher; cat. no. 25300-054) for 1-2 min to promote removal of cells from the growth surface. Subsequently, 10 mL warm DMEM media was added and cells were resuspended by pipetting. Cells were counted using Bürker-Türk and appropriate cell numbers were seeded in cell culture vessels.

Approximately, 4 mL of HEK293T cells were seeded in the 10 cm dish (Falcon, cat. no. 353003) and cultured as described above. When 40 – 50% confluence was reached, cells were transfected with FLAG-HA-METTL16 plasmid: 10 μg of plasmids was diluted in 500 μL of 150 mM NaCl. Simultaneously, 26 μg of linear polyethylenimine, MW 25000 (PEI, Polysciences Inc., cat. no. 23966) was diluted in 500 μL of 150 mM NaCl. Solutions were mixed together, vortexed vigorously for 15 s and incubated for 15 min. at room temperature. Then mix was added to the HEK293T cells in the DMEM CM. After 24 h, medium was changed for the fresh DMEM CM. Cells were grown for 72h in total.

Isolation of human METTL16 complexes for mass spectrometry

Cells in 10 cm dishes were washed 3x with ice cold PBS and 1 mL of lysis buffer [20 mM Tris pH 7.4, 150 mM NaCl, 0.5% Triton X-100, 0.1% sodium deoxycholate, 1 mM EDTA, 0.5 mM DTT, protease inhibitor (Complete Protease Inhibitor Cocktail Tablet, Roche, cat. no. 5056489001)] was added to the cells. Cells were removed from their growth surface using a cell scraper (Costar; cat. no. 3010) and transferred to 1.5 mL eppendorf tubes. Cell lysate was passed 5-times through a 26 G needle (B. Braun Medical Inc., #466-5457) and kept on ice for 15 min. The total cell lysate was spun at 12,000 x g for 10 min at 4°C. After centrifugation, supernatant was transferred to a fresh tube and spun again (12,000 x g, 10 min., 4°C). The cleared lysate was transferred to a fresh tube. While 50 μl of lysate was transferred to a fresh tube and flash-frozen in liquid nitrogen to use as an input, 950 μl was incubated for 4 h at 4°C with 20 μL of Anti-HA Affinity Matrix (Roche, cat. no. 11815016001). After, beads were collected by gentle centrifugation (500 x g for 1 min at 4°C) and eluate was discarded. Beads were washed 5 times with wash buffer (50 mM Tris pH 7.4, 150 mM NaCl, 0.1% Triton X-100, 1 mM EDTA) Then, beads were transferred to fresh 1.5 mL eppendorf tubes and 40 μL of 2x Laemmli buffer (4% SDS, 20% glycerol, 120 mM Tris-HCl pH 6.8, 10% β-mercaptoethanol, 0.02% bromophenol blue) was added. Beads were boiled at 95°C for 5 minutes and stored at −20°C. Proteins were identified by mass spectrometry at the Functional Genomics Center Zurich (ETH Zurich) (Figure S5F). Database searches were performed using the Mascot (SwissProt, human) search program. Applied settings: 1% protein false detection rate (FDR), min. 2 peptides per protein, 0.1% peptide FDR.

Isolation of METTL16 complexes from mouse testes and spleen

An aliquot of 80 μl of Dynabeads Protein A slurry (ThermoFisher, 10001D) was transferred to a fresh Eppendorf tube and washed three times with 1 mL of 20 mM sodium phosphate with 0.02% Tween20. Then, 20 μg of METTL16 antibody (abcam, ab186012) or 20 μg of mouse IgG control antibody (Santa Cruz, cat.no. sc-2025) in 500 μl of 20 mM sodium phosphate with 0.02% Tween20 was added to the beads and incubated overnight at 4°C with rotation.

Next day, two adult (P60) mouse testes and one spleen were cut into pieces using scalpel blade and placed into separate 1.5 mL eppendorf tubes. 500 μL of ice cold lysis buffer [20 mM Tris pH 7.4, 150 mM NaCl, 0.5% Triton X-100, 0.5% sodium deoxycholate, 1 mM DTT, 1 mM EDTA and protease inhibitor (Roche) was added to the tubes. Organs were dounced 15-times using a plastic pestle and left on ice for 10 min. Then, tubes were spun at 12000 x g at 4°C for 10 minutes. Supernatant was transferred to a fresh tube and centrifugation was repeated to clarify the lysate further. Supernatants were transferred to a fresh tube and diluted 2x with dilution buffer dilution buffer [20 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, protease inhibitor] to decrease sodium deoxycholate and Triton X-100 concentration to 0.25%. An aliquot of 50 μl of lysate was transferred to a fresh tube and flash-frozen in liquid nitrogen to use as an input, while rest was transferred to antibody-bound Dynabeads prepared above, and incubated at 4°C for 4 h with rotation.

After 4 h, the supernatant was removed and beads were washed with wash buffer (20 mM Tris pH 7.4, 150 mM NaCl, 0.2% Triton X-100, 1 mM EDTA, 0.5 mM DTT). Washing was repeated four more times, after which beads were transferred to the fresh 1.5 mL eppendorf tubes and 40 μL of 2x Laemmli buffer (4% SDS, 20% glycerol, 120 mM Tris-HCl pH 6.8, 10% β-mercaptoethanol, 0.02% bromophenol blue) was added. Beads were boiled at 95°C for 5 minutes and stored at −20°C. Proteins in the samples were identified at the Functional Genomics Center Zurich (ETH Zurich) with the shotgun mass spectrometry analysis (Figure S5G). Database searches were performed using the Mascot (SwissProt, human) search program. Applied settings: 1% protein false detection rate (FDR), min. 1 peptides per protein, 0.1% peptide FDR.

Mass spectrometry

Mass spectrometry to confirm purified recombinant proteins were carried out at the Proteomics Core Facility, EMBL, Heidelberg. Identification of components within an immunoprecipitated complex was carried out at the Functional Genomics Center Zurich (ETH Zurich) using the shotgun mass spectrometry analysis. Database searches were performed using the Mascot (SwissProt, all species) search program. Applied settings if not stated differently are 1% protein false detection rate (FDR), min. 2 peptides per protein, 0.1% peptide FDR.

Preparation of RNA libraries

In vitro *methylation with METTL16 and m^{6}A-IP-RNaseq*

Libraries of randomized 30 nt RNA sequences were chemically synthesized (Microsynth, CH). The sequences had a constant central 9-mer sequence flanked by randomized (represented by N) sequences (MM-RNA-14: N_{11}-UAC**A**GAGAA-N_{10}). The 9-mer sequence originates from the hairpin 1 of the human *MAT2A* mRNA and carries the m^{6}A methylation site for METTL16 (Pendleton et al., 2017). RNA solutions (100 μM) with 50 mM NaCl were denatured at 80°C for 1 min, and refolded by allowing to cool to room-temperature. *In vitro* methylation reactions containing 15 μL (100 μM) of the above RNA library were carried out in 100 mL reactions (50 mM Tris-HCl, pH 7.5, 100 mM KCl, 5 mM $MgCl_2$, 20U of Riboblock RNase inhibitor, 0.64 mM SAM) with 20 μg of hMETTL16. Reactions were carried out in duplicates and incubated at 37°C, overnight. Reactions were then removed and frozen at −20°C prior to further processing.

A small portion (10%) was left aside to be used as input sample, while the remainder was subjected to immunoprecipitation. The m^{6}A immunoprecipitation was performed as described (Ke et al., 2015). Briefly, 100 μL of Dynabeads Protein A (Life Technologies; 10002D) were washed once in PXL buffer (1 × PBS, 0.1% SDS, 0.5% sodium deoxycholate, 0.5% NP-40) followed by pre-treatment with BSA (final concentration 1μg/ μL) in 200 μL PXL buffer for 45 minutes at room-temperature (RT). BSA pre-treated beads was then conjugated with m^{6}A rabbit polyclonal antibody (5 μg; Synaptic Systems, catalog no. 202003) in 200 μL PXL buffer supplemented with 4 μL of RNasin RNase inhibitor (Promega; N2611) for one hour at RT on a rotating wheel. Dynabeads were further washed twice with PXL buffer and finally beads were resuspended in 400 μL of PXL buffer and 5 μL of RNasin. The *in vitro* methylation reaction prepared above was added to the beads and incubated at 4°C for 2 hours on a rotating wheel. After two hours incubation, the beads were washed twice by ice-cold Nelson low-salt buffer (15 mM Tris at pH 7.5, 5 mM EDTA), once by ice-cold Nelson high-salt buffer (15 mM Tris at pH 7.5, 5 mM EDTA, 2.5 mM EGTA, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1 M NaCl), once by ice-cold Nelson stringent wash buffer (15 mM Tris at pH 7.5, 5 mM EDTA, 2.5 mM EGTA, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 120 mM NaCl, 25 mM KCl), and last by ice-cold NT-2 buffer (50 mM Tris at pH 7.4, 150 mM NaCl, 1 mM MgCl2, 0.05% NP-40). Antibody-bound RNAs were eluted by incubating the beads with 0.5 mg/mL N^6-methyl adenosine (Sigma-Aldrich; M2780) in NT2 buffer for one hour at 4°C. The eluted RNAs were precipitated with ethanol and glycogen and dissolved in RNase-free water.

The input and IP RNAs were first 3′ end dephosphorylated with T4 PNK (NEB; M0201S, 10 U/μL) in the absence of ATP at 37°C for 45 minutes (40 μL reaction: 35.5 μL RNA, 4 μL 10X T4 PNK buffer, 0.5 μL of T4 PNK) followed by phosphorylation of 5′ end (50 μL reaction: 40 μL dephosphoryated RNA, 6.5 μL water, 1 μL RNasin, 0.5 μL 100 mM ATP, 1 μL 10X T4 PNK buffer 1 μL T4 PNK) at 37°C for 45 minutes. RNAs were phenol chloroform-extracted, ethanol precipitated and resupended in 6 μL of RNase-free water. The input RNA fragments and the immunopurified RNAs after the phosphorylation step were directly used for library preparation (barcoded at 3′ end) using NEBNext® Multiplex Small RNA Library Prep Set for Illumina® (NEB; catalog No. E7560S) following

manufacturer's instructions. The synthesized cDNA libraries were resolved on 3% high-resolution MethaPhor agarose (Lonza; cat. No. 50180) gels in 1X TAE buffer at 70 V. Fragments in the size-range of ~150-250 bp were gel-extracted with the use of MinElute Gel Extraction Kit (QIAGEN; cat No. 28604). Multiple libraries with different barcodes (at 3′ end) were mixed in equimolar ratios and sequenced with the HiSeq Illumina Platform (EMBL GeneCore facility, Heidelberg). The maximum sequencing length was 50 nt. The list of sequencing libraries generated are provided in Table S2.

Mouse single-embryo library preparation

Polyadenylated transcripts in single embryos (E2.5 morula or E3.5 blastocysts) were amplified using the Smart-seq2 protocol (Picelli et al., 2014). The protocol generates libraries that lack strand specificity. Multiple libraries with different barcodes (at the 3′ end) were mixed in equimolar ratios and paired-end sequencing reads were obtained with the HiSeq Illumina Platform (EMBL GeneCore facility, Heidelberg). The maximum sequencing length was 80 nt. The list of sequencing libraries generated are provided in Table S2.

QUANTIFICATION AND STATISTICAL ANALYSIS

In vitro methylation with METTL16 and m^{6}A-IP-RNaseq

Reads were sorted into individual libraries based on the barcodes and the 3′ adaptor sequences were removed using cutadapt 1.9.1 (http://journal.embnet.org/index.php/embnetjournal/article/view/200). Only reads of final length of 30 nucleotides with correctly sequenced TACAGAGAA consensus motif at position 12-20 and without any Ns were kept for further analysis using R 3.4.3 (R Core Team, 2017) and Bioconductor (Huber et al., 2015). To search for possible preference of human METTL16 for specific structured RNA features, we analyzed the predicted secondary structures of the sequenced oligos and compare their representation in between the m^{6}A-IP and input libraries. For each sequence we obtained the minimum free energy (MFE) secondary structure using RNAfold (Lorenz et al., 2011). We used DESeq2 1.18.1 bioconductor package (Love et al., 2014) to obtain the lists of structures significantly enriched or depleted in IP (immunoprecipitation) libraries when compared to input libraries (adjusted p value < 0.1). Top enriched structures were plotted in dot bracket notation (DBN) (Figures 3E and 3F). To search for preferred features in IP-enriched structures, for every structure and each position based on DBN, we checked whether it is part of the stem, is in a loop, is in a bulge or if it is in between two stems. Then we compared the proportion of the structures having nucleotide at specific position in a stem, loop, etc, in between IP-enriched structures, IP-depleted structures and structures with no difference in their abundance between m^{6}A-IP and input (Figure 3G). In IP-enriched structures we observed a clear preference of A at position 15 (in the motif UAC**A**GAGAA), which is methylated by METTL16, to be in a single nucleotide bulge or to lie in a region surrounded by stem structures (Figure 3G). To see the differences in IP-enriched and depleted structures, we also plotted the log2 difference of the frequencies for IP-enriched (or depleted) structures when related to the structures not differentially represented between m^{6}A IP and input (Figure 3H). While the IP-enriched structures had higher proportion of 15A in a single nucleotide bulge or lying between two stems, the IP-depleted structures showed the opposite trend, with less proportion of structures with 15A in a bulge or in between two stems. For calculations of these log2 ratios of the frequencies, the frequencies lower than 0.5% were considered to be 0.5. Proportion of individual structures in which 15A can be found is summarized in a barplot (Figure 3I). We also directly compared the frequencies of oligos with individual positions in a stem, loop etc. in between m^{6}A-IP libraries and input libraries (Figure S3C) and their log2 ratios, separately for both replicas (Figure S3D). In m^{6}A-IP libraries we observed increased proportion of oligos where 15A is in single nucleotide bulge or in between two stems.

To check whether there is specific sequence preference outside the TAC**A**GAGAA consensus motif, we compared the nucleotide frequencies at individual oligo positions between IP and input samples and plotted their log2 ratios (Figure S3E). In the IP-oligos we observed general higher frequencies of G and C. We used MEME - Motif discovery tool 4.11.2 (Bailey and Elkan, 1994) to search for any sequence motif in the IP enriched left 11-mers and right 10-mers surrounding the TAC**A**GAGAA (Figure S3F).

Transcriptome analysis of Mettl16 mutant mouse embryos

Paired-end reads were sorted into individual libraries based on the barcodes and aligned to NCBI RefSeq transcripts (build mm10) using Salmon v0.7.2 (Patro et al., 2017). The genotype of the mouse embryos giving rise to the individual samples was assessed based on the presence of the reads derived from WT *Mettl16* allele (containing CACCAGGATTCGGACAAAACTA or TAGTTTTGTCC GAATCCTGGTG sequence, since libraries are non-strand-specific) and from *Mettl16* KO allele (containing TCACCAATGTAAATA GATGAGG or CCTCATCTATTTACATTGGTGA sequence). For E2.5 there were 5 *Mettl16*$^{+/+}$ (WT), 14 *Mettl16*$^{+/-}$ (HET) and 12 *Mettl16*$^{-/-}$ (KO) samples. For E3.5 we got 18 *Mettl16*$^{+/+}$ (WT), 9 *Mettl16*$^{+/-}$ (HET) and 12 *Mettl16*$^{-/-}$ (KO) samples

The transcript estimates were imported into DESeq2 1.18.1 (Love et al., 2014) and summarized to gene levels using tximport 1.2.0 (Soneson et al., 2015). The DESeq2 was used to obtain lists of differentially expressed genes with statistical significance (adjusted p value $\leq$ 0.1). The MA plots were plotted using graphics::smoothScatter function and the individual genes with significantly different expression were highlighted (Figures 4D and 5C).

For the E2.5 dataset, twenty genes were found to be differentially expressed between some of the genotypes (Figure 4E; Table S3) and their expression was visualized by heatmap using the made4::heatplot function (Figure 4E). Only four of the genes had significantly different expression in *Mettl16*$^{-/-}$ versus *Mettl16*$^{+/-}$ and also in *Mettl16*$^{-/-}$ versus *Mettl16*$^{+/+}$ comparison (Figure S5B). Boxplots of normalized counts were plotted for those genes, with individual samples plotted as dots using graphics:stripchart function (Figures 4F and S5C). To visualize the coverage of individual genomic loci, the sequenced reads were aligned to reference mm10

genome using STAR (Dobin et al., 2013) and the normalized coverage was calculated using GenomicRanges::coverage function. Mean coverages were plotted for individual genotypes using Gviz 1.22.3 (Hahne and Ivanek, 2016) together with the transcript annotation obtained either from NCBI RefSeq track from UCSC or from GENCODE M17 (Figures 4G and S5E). Gviz was also used to plot the coverage of individual exons or introns. To compare the amount of reads coming from individual introns of *Mat2a*, featureCounts (Liao et al., 2014) was used to obtain the counts for individual genomic exons and introns which were then normalized by DESeq2. Boxplots of the counts were plotted for individual introns of *Mat2a* normalized to library sizes or to overall *Mat2a* counts (Figure S5D). The intron coordinates used were shortened by 10 nucleotides from both sides so that the intron counts were not affected by exonic reads partially protruding into the introns. The JunctionSeq 1.8.0 (Hartley and Mullikin, 2016) was used to search for differential usage of splice junctions among the genotypes. Only few splice junctions were significantly (adjusted p value $\leq$ 0.01) differentially used between *Mettl16*$^{-/-}$ versus *Mettl16*$^{+/-}$ and also in *Mettl16*$^{-/-}$ versus *Mettl16*$^{+/+}$ comparison (Table S3). Whereas the reads spanning the splice junction of last *Mat2a* (NM_145569 = ENSMUST00000059472.9) intron were depleted in *Mettl16*$^{-/-}$ when normalized to overall *Mat2a* transcript levels, alternative splice junction (common for ENSMUST00000206904.1 and ENSMUST00000206692.1) was elevated (Figure 4H) as shown by boxplots. This was accompanied by overall increase of last *Mat2a* intron counts in *Mettl16*$^{-/-}$ which is however not significant.

For the E3.5 dataset, 5166 genes were found to be differentially expressed between some of the genotypes (Figure 5D; Table S3) and their expression was visualized by heatmap using the made4::heatplot function (Figure 5D). Most of the genes which were found to be dysreguled in *Mettl16*$^{-/-}$ versus *Mettl16*$^{+/-}$ were also differentially expressed between *Mettl16*$^{-/-}$ and *Mettl16*$^{+/+}$ (Figure 5E). Enriched gene ontology biological processes for upregulated and downregulated genes were identified by ENRICHR (Chen et al., 2013; Kuleshov et al., 2016) and are summarized in Table S3. Boxplots of normalized counts were plotted for *Mettl16*, *Mat2a* and top dysregulated genes with individual samples plotted as dots using graphics:stripchart function (Figures 5B and S6D). FeatureCounts was used to obtain summarized counts for introns, exons and repeats which did not show any differences between the genotypes neither in E2.5 nor in E3.5 (Figures S5A and 5G) and also to obtain the counts for individual genomic exons and introns. Boxplot was used to compare number of reads arising from last intron of *Mat2a* and splice junction reads crossing the last intron (Figure S6A) whose counts were obtained from JunctionSeq analysis. All the splice junctions which were significantly (adjusted p value $\leq$ 0.01) differentially used between *Mettl16*$^{-/-}$ versus *Mettl16*$^{+/-}$ and also in *Mettl16*$^{-/-}$ versus *Mettl16*$^{+/+}$ comparison are summarized in Table S3. Overall counts of uniquely mapping reads crossing the splice junctions were obtained from SJ.out.tab files generated by STAR and their proportion was compared between individual samples (Figure 5F). To find out whether genes specific to any developmental stage are misregulated in the mutant, we checked the expression of the key transcription and chromatin factors characteristic for different stages (Table S1 of Mohammed et al., 2017). Heatmap of log2 (normalized counts +1) expression was plotted using gplots::heatmap.2 from individual samples (Figure S6B). Boxplots were used to compare the average expression change between *Mettl16*$^{-/-}$ versus *Mettl16*$^{+/-}$ and *Mettl16*$^{-/-}$ versus *Mettl16*$^{+/+}$, with individual genes plotted as dots (Figure S6C).

DATA AND SOFTWARE AVAILABILITY

Deep sequencing data generated in this study are deposited with Gene Expression Omnibus under the accession number GSE116329. Crystallographic data are deposited with Protein Data Bank under PDB accessions: 6GFN, 6GT5 and 6GFK. Other raw data associated with this study are deposited with Mendeley Data under the accession https://doi.org/10.17632/ny82j2ngt5.1. The *Mettl16* knockout mutant mouse generated in this study will be available from the Lead Contact.

Molecular Cell, Volume *71*

Supplemental Information

Methylation of Structured RNA by the m^{6}A Writer METTL16 Is Essential for Mouse Embryonic Development

Mateusz Mendel, Kuan-Ming Chen, David Homolka, Pascal Gos, Radha Raman Pandey, Andrew A. McCarthy, and Ramesh S. Pillai

INVENTORY OF SUPPLEMENTAL INFORMATION

One PDF with

1) Supplemental Figures (S1-6)

2) Supplemental figure legends (S1-6)

3) Supplemental Tables (S1-2) and legends

Supplemental Figures

Figure S1. Full-length human METTL16 exists as monomers. Related to Figure 1.

Figure S2. Mutational analysis of the human METTL16-core methyltransferase domain to define the RNA-binding groove. Related to Figures 1 and 2.

Figure S3. In vitro methylation with human METTL16-FL and a randomized RNA library reveals structural and sequence requirements for m^6A RNA methylation. Related to Figure 3.

Figure S4. Embryonic lethality in *Mettl16* knockout mice around implantation stage. Related to Figure 4.

Figure S5. Lack of METTL16 has very specific and limited effect on the transcriptome of E2.5 embryos. Related to Figure 4.

Figure S6. Dramatically altered transcriptome of E3.5 *Mettl16* knockout embryos. Related to Figure 5.

Supplemental Tables

Table S1. DNA primers and RNA oligonucleotides used in this study. Related to STAR Methods and Figure 1,2 and 3.

Table S2. List of all deep-sequencing libraries created in this study. Related to STAR Methods and Figure 4 and 5.

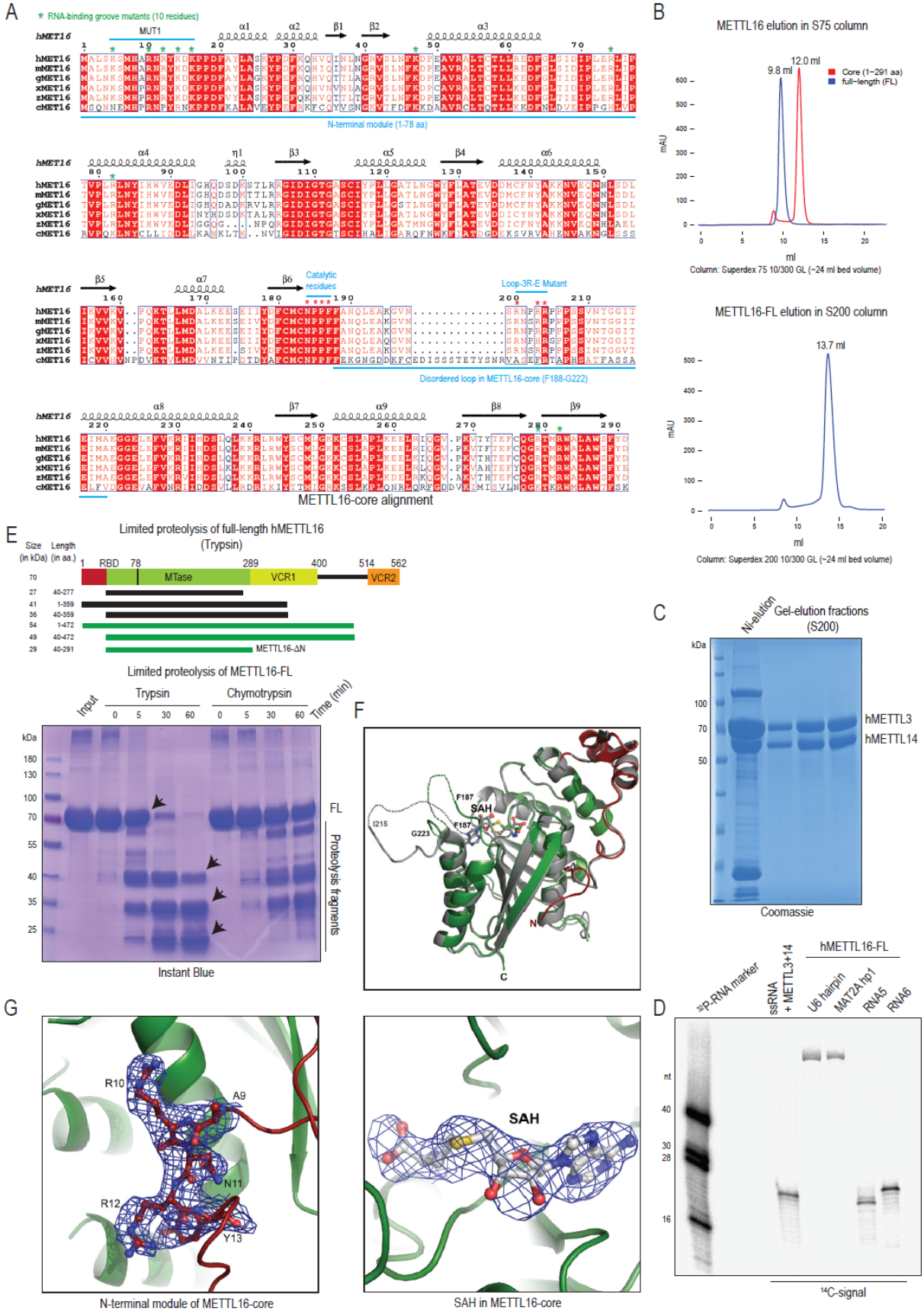

A
RNA-binding groove mutants (10 residues)
MUT1
N-terminal module (1-78 aa)
Catalytic residues
Loop-3R-E Mutant
Disordered loop in METTL16-core (F188-G222)
METTL16-core alignment
B
METTL16 elution in S75 column
12.0 ml
9.8 ml
Core (1-291 aa)
full-length (FL)
mAU
ml
Column: Superdex 75 10/300 GL (~24 ml bed volume)
METTL16-FL elution in S200 column
13.7 ml
Column: Superdex 200 10/300 GL (~24 ml bed volume)
C
Ni-elution
Gel-elution fractions (S200)
kDa
hMETTL3
hMETTL14
Coomassie
D
hMETTL16-FL
32P-RNA marker
ssRNA
+ METTL3+14
U6 hairpin
MAT2A hp1
RNA5
RNA6
nt
14C-signal
E
Limited proteolysis of full-length hMETTL16 (Trypsin)
Size (in kDa)
Length (in aa.)
RBD
MTase
VCR1
VCR2
METTL16-ΔN
Limited proteolysis of METTL16-FL
Input
Trypsin
Chymotrypsin
Time (min)
FL
Proteolysis fragments
Instant Blue
F
SAH
F187
I215
G223
N
C
G
R10
A9
N11
R12
Y13
N-terminal module of METTL16-core
SAH
SAH in METTL16-core

Figure-S1

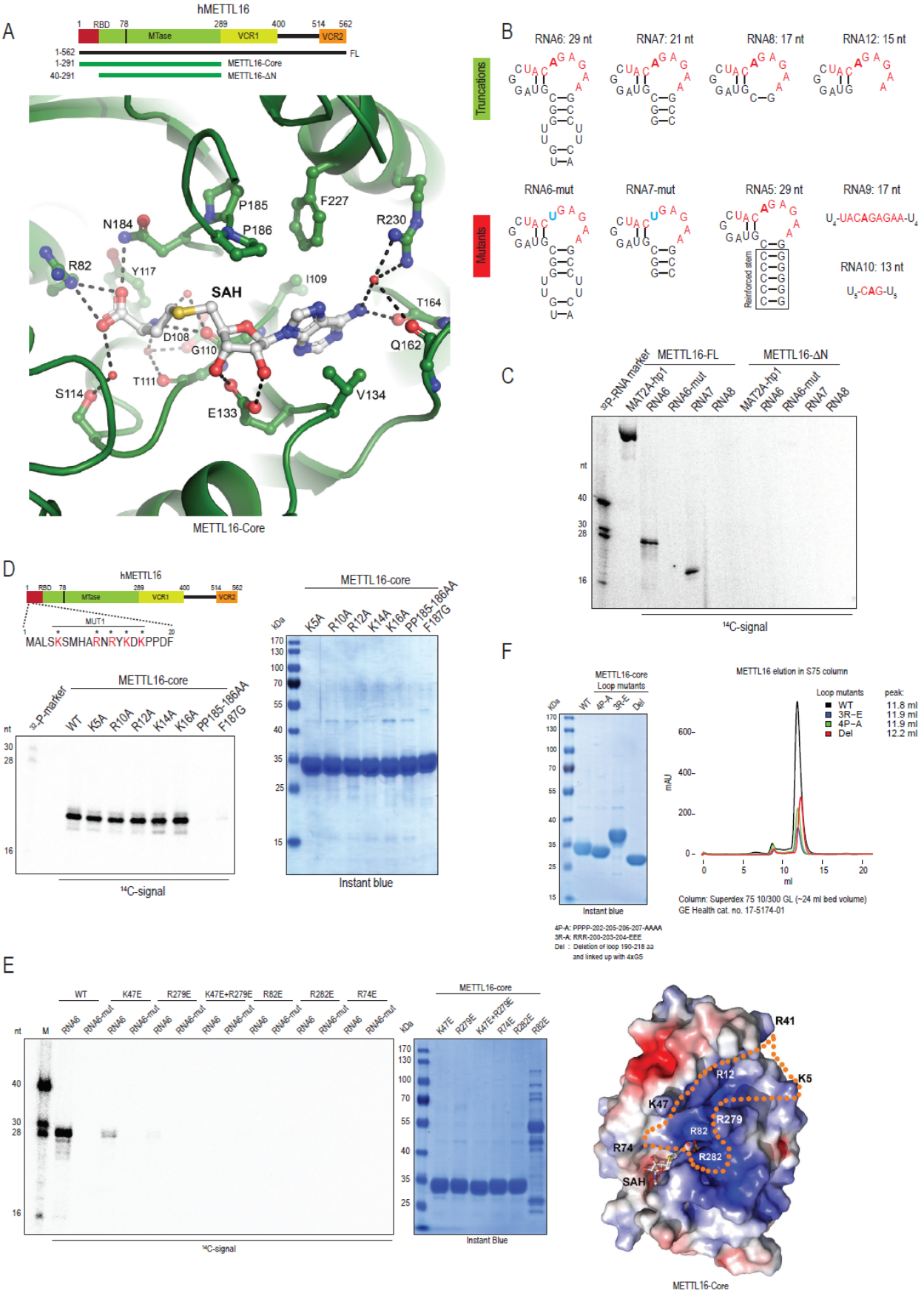

A
hMETTL16
1 RBD 78 MTase 289 VCR1 400 514 562 VCR2
1-562 FL
1-291 METTL16-Core
40-291 METTL16-ΔN
P185
F227
N184
P186
R230
R82
Y117
I109
SAH
T164
D108
G110
Q162
T111
S114
V134
E133
METTL16-Core
B
Truncations
RNA6: 29 nt
RNA7: 21 nt
RNA8: 17 nt
RNA12: 15 nt
Mutants
RNA6-mut
RNA7-mut
RNA5: 29 nt
RNA9: 17 nt
U4-UACAGAGAA-U4
Reinforced stem
RNA10: 13 nt
U5-CAG-U5
C
32P-RNA marker
METTL16-FL
METTL16-ΔN
MAT2A-hp1
RNA6
RNA6-mut
RNA7
RNA8
nt
40
30
28
16
14C-signal
D
hMETTL16
MUT1
MALSKSMHARNRYKDKPPDF
METTL16-core
32P-marker
WT
K5A
R10A
R12A
K14A
K16A
PP185-186AA
F187G
kDa
170
130
100
70
55
40
35
25
15
14C-signal
Instant blue
F
METTL16-core
Loop mutants
WT
4P-A
3R-E
Del
Instant blue
4P-A: PPPP-202-205-206-207-AAAA
3R-A: RRR-200-203-204-EEE
Del : Deletion of loop 190-218 aa
and linked up with 4xGS
METTL16 elution in S75 column
Loop mutants
peak:
WT 11.8 ml
3R-E 11.9 ml
4P-A 11.9 ml
Del 12.2 ml
mAU
600
400
200
0
0
5
10
15
20
ml
Column: Superdex 75 10/300 GL (~24 ml bed volume)
GE Health cat. no. 17-5174-01
E
WT
K47E
R279E
K47E+R279E
R82E
R282E
R74E
RNA6
RNA6-mut
nt
M
40
30
28
16
14C-signal
METTL16-core
K47E
R279E
K47E+R279E
R74E
R282E
R82E
kDa
170
130
100
70
55
40
35
25
Instant Blue
R41
R12
K5
K47
R279
R82
R74
R282
SAH
METTL16-Core

Figure-S2

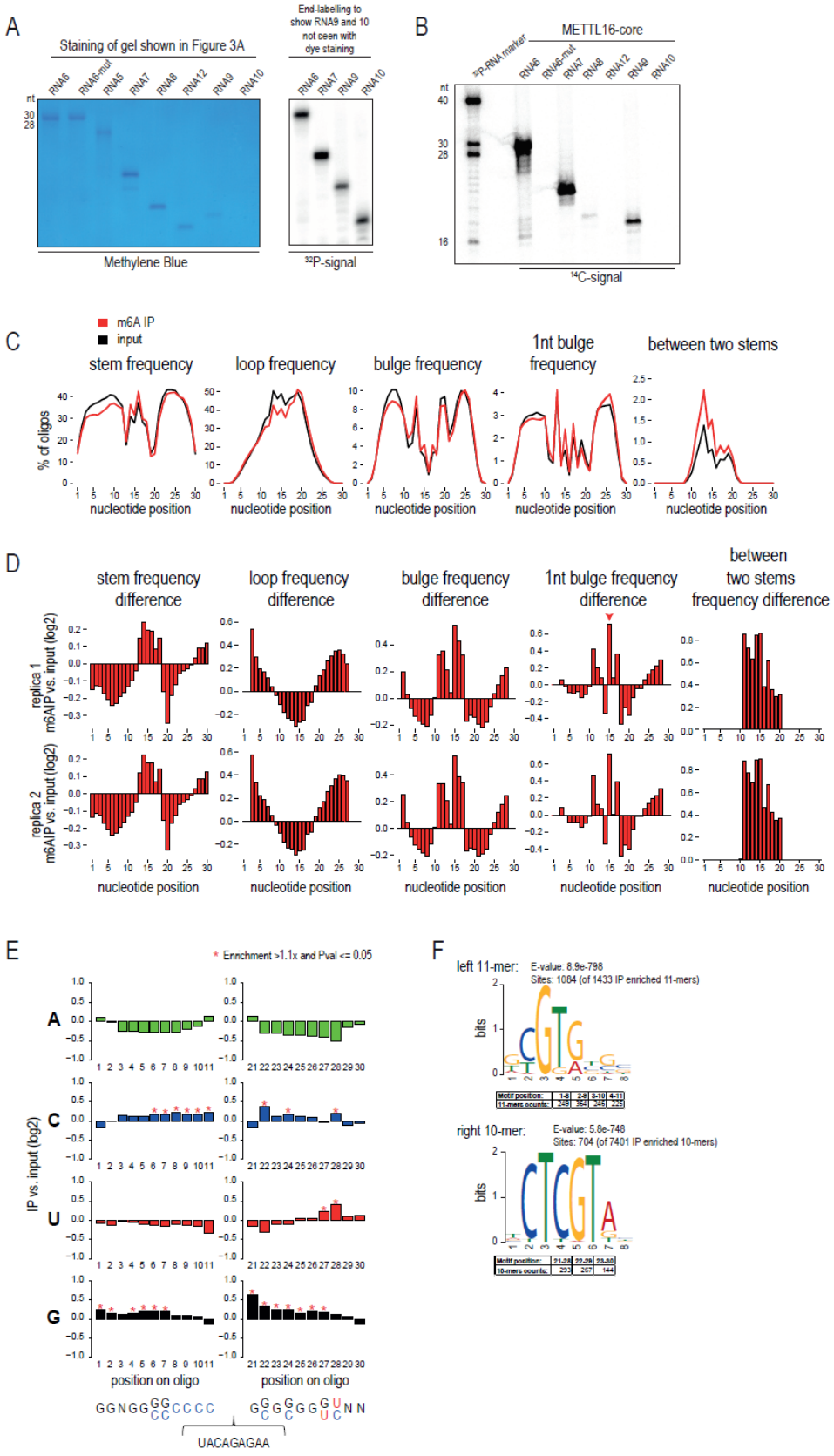
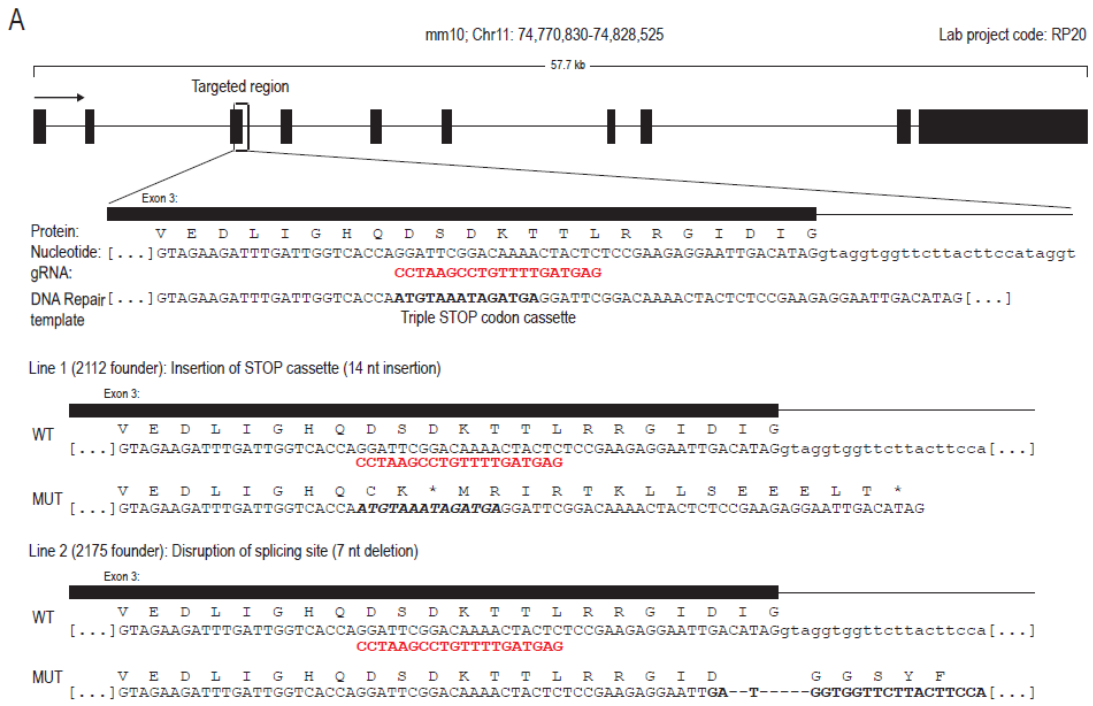
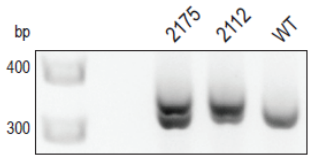

A
Staining of gel shown in Figure 3A
End-labelling to show RNA9 and 10 not seen with dye staining
RNA6
RNA6-mut
RNA5
RNA7
RNA8
RNA12
RNA9
RNA10
nt
30
28
Methylene Blue
32P-signal
B
METTL16-core
32P-RNA marker
nt
40
30
28
16
14C-signal
C
m6A IP
input
stem frequency
loop frequency
bulge frequency
1nt bulge frequency
between two stems
% of oligos
nucleotide position
D
stem frequency difference
loop frequency difference
bulge frequency difference
1nt bulge frequency difference
between two stems frequency difference
replica 1 m6AIP vs. input (log2)
replica 2 m6AIP vs. input (log2)
nucleotide position
E
* Enrichment >1.1x and Pval <= 0.05
A
C
U
G
IP vs. input (log2)
position on oligo
GGNGGGCCCCC
GGGGGGUCNN
UACAGAGAA
F
left 11-mer:
E-value: 8.9e-798
Sites: 1084 (of 1433 IP enriched 11-mers)
bits
Motif position:
11-mers counts:
right 10-mer:
E-value: 5.8e-748
Sites: 704 (of 7401 IP enriched 10-mers)
10-mers counts:

Figure-S3

A



B



C

Genotype of animals in litters at weaning age (P21) from *Mettl16+/-* x *Mettl16+/-* crosses

	Total	KO	HET	WT	Male	Female
Number	156	0	114	42	73	83
Percentage	-	0.0%	73.1%	26.9%	46.8%	53.2%

D

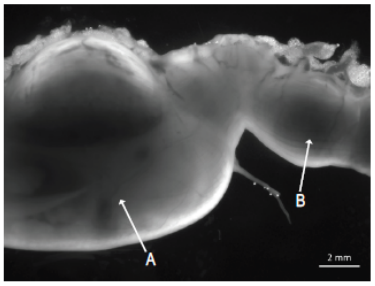
Genotype of E6.5 embryos from *Mettl16+/-* x *Mettl16+/-* crosses

	Placentas	Embryos	WT	HET	KO	Unknown
Number	68	52	16	32	1	3
Percentage	-	76.5%	30.8%	61.5%	1.9%	5.8%

E Genotype of E8.5 embryos from *Mettl16+/-* x *Mettl16+/-* crosses

	Placentas	Embryos	WT	HET	KO
Number	24	14	1	13	0
Percentage	-	58%	7%	93%	

F



E12.5

Heterozygote (A) Empty amniotic sac (B)

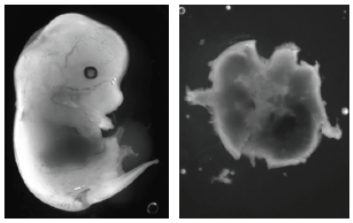
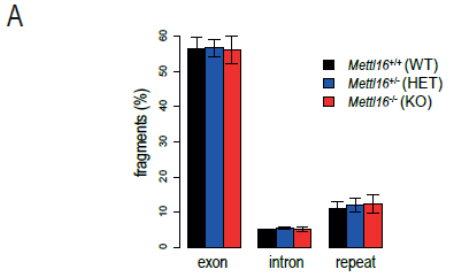
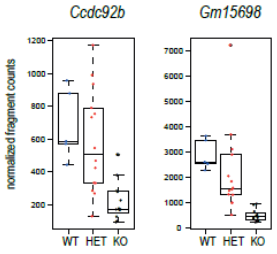


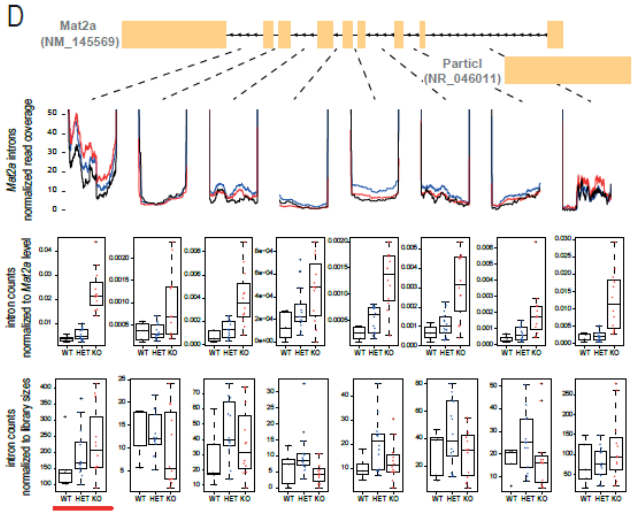
Figure-S4

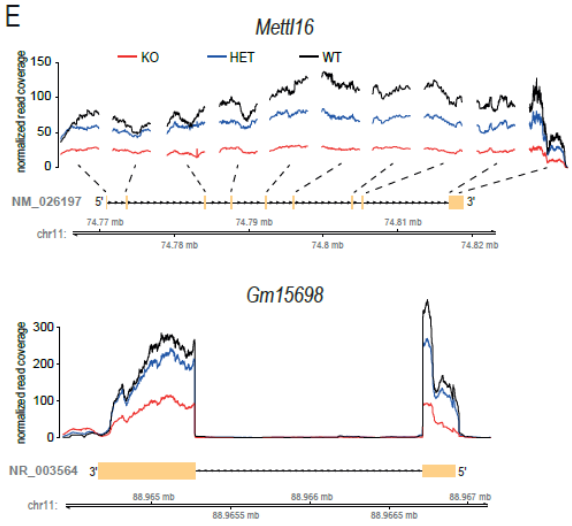


B

	log2 fold change			adjusted p-value		
gene	KO vs WT	KO vs HET	HET vs WT	KO vs WT	KO vs HET	HET vs WT
Ccdc92b	-1.63	-1.39	-0.24	4.79E-02	6.43E-03	1.00E+00
Mettl16	-1.83	-1.32	-0.52	4.31E-11	3.13E-10	1.00E+00
Mat2a	-2.41	-2.00	-0.41	2.69E-31	5.03E-40	1.00E+00
Gm15698	-2.58	-2.16	-0.41	6.40E-07	1.65E-09	1.00E+00







F

METTL16 Complex from HEK293 cells

				Exclusive Unique Peptide Count				
				METTL16			GFP	
No.	Name	ID	MW [kDa]	Rep. 1	Rep. 2	Rep. 3	Rep. 1	Rep. 2
1	Kanadaptin	SLC4A1AP	89	26	24	18	4	0
2	U6 small nuclear RNA (adenine-(43)-N(6))-methyltransferase	METTL16	64	74	84	77	13	5
3	Lupus La protein	SSB	47	9	7	4	0	2
4	DNA repair protein RAD50	RAD50	154	5	3	6	0	2
5	Splicing factor 3B subunit 2	SF3B2	100	19	17	15	5	2
6	Deoxynucleoside triphosphate triphosphohydrolase SAMHD1	SAMHD1	72	4	9	7	0	3
7	Splicing factor 3B subunit 1	SF3B1	146	34	21	33	9	6
8	Kinesin-like protein KIF18B	KIF18B	94	4	6	6	1	2
9	Structural maintenance of chromosomes protein 4	SMC4	147	4	3	4	0	0
10	Zinc finger CCCH domain-containing protein 8	ZC3H8	34	2	5	4	0	2
11	WD repeat-containing protein 82	WDR82	35	4	2	5	2	0
12	Splicing factor 3B subunit 3	SF3B3	136	23	8	22	6	6
13	Furin	FURIN	87	3	5	5	0	3
14	Probable 28S rRNA (cytosine(4447)-C(5))-methyltransferase	NOP2	89	2	7	3	2	0
15	Galectin-1	LGALS1	15	2	3	6	2	0

G

METTL16 Complex from mouse tissues

				Exclusive Unique Peptide Count				
No.	Name	ID	MW [kDa]	Testis	Testis	Spleen	IgG CTRL	IgG CTRL
1	Splicing factor 3B subunit 1	SF3B1	146	28	33	51	0	0
2	Splicing factor 3B subunit 3	SF3B3	136	21	25	33	0	0
3	Protein PRRC2A	PRRC2A	229	24	32	13	0	0
4	Kinesin-like protein KIF13B	KIF13B	203	24	23	7	0	0
5	Probable ATP-dependent RNA helicase DDX5	DDX5	69	22	20	33	0	0
6	Cleavage and polyadenylation specificity factor subunit 1	CPSF1	161	28	19	22	0	0
7	pre-mRNA 3' end processing protein WDR33	WDR33	145	24	19	22	0	0
8	Splicing factor 3B subunit 2	SF3B2	100	20	17	28	0	0
9	Splicing factor 3A subunit 1	SF3A1	89	10	19	30	0	0
10	Protein scribble homolog	SCRIB	174	19	21	5	0	0
11	Gem-associated protein 5	GEMIN5	167	21	13	18	0	0
12	U6 small nuclear RNA (adenine-(43)-N(6))-methyltransferas	METTL16	62	17	17	14	0	0
13	Serine/arginine repetitive matrix protein 2	SRRM2	295	12	28	16	0	0
14	Cleavage and polyadenylation specificity factor subunit 2	CPSF2	88	21	14	16	0	0
15	Heterogeneous nuclear ribonucleoprotein M	HNRNPM	78	19	14	15	0	0

Figure-S5

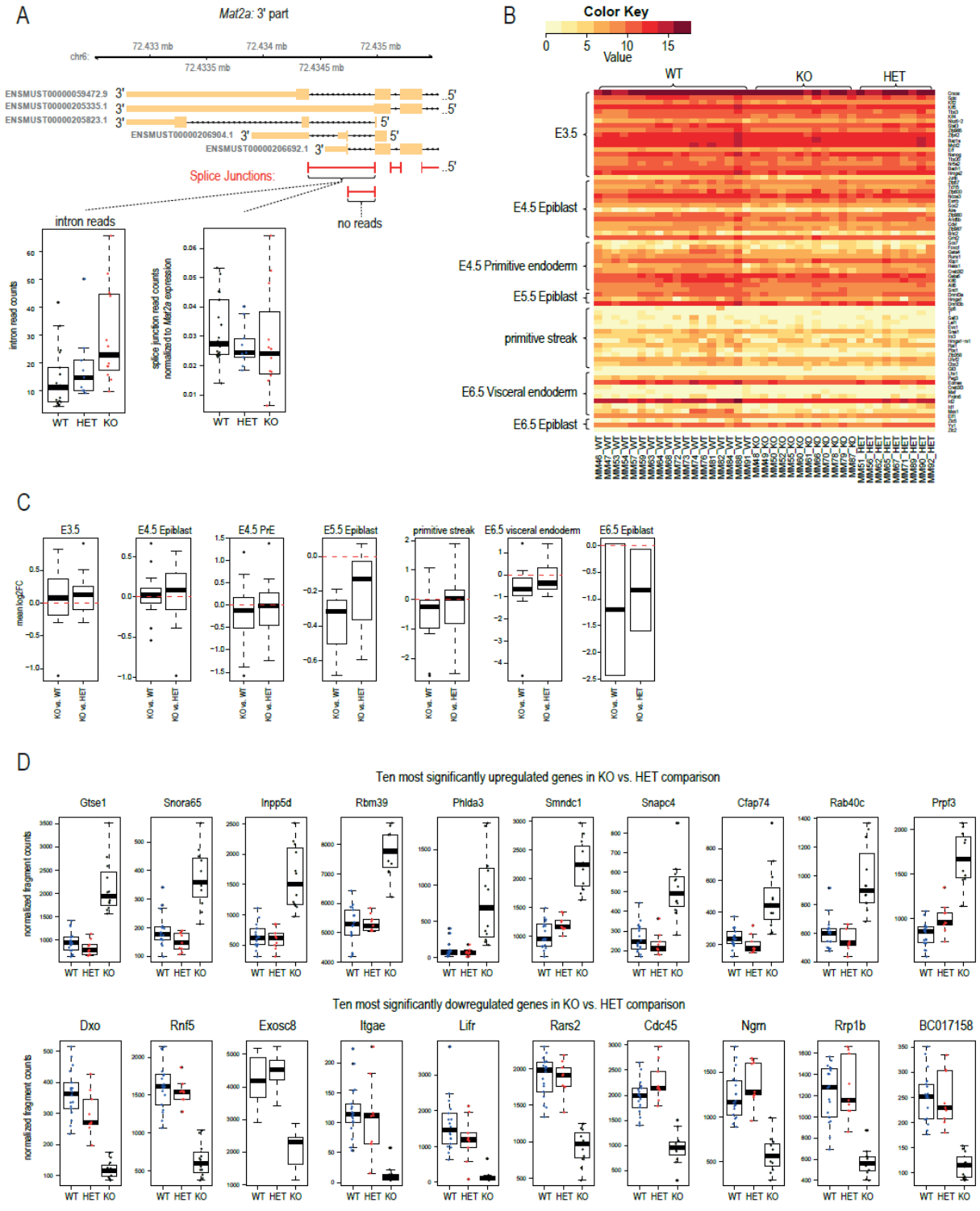

A
Mat2a: 3' part
chr6:
72.433 mb
72.434 mb
72.435 mb
72.4335 mb
72.4345 mb
ENSMUST00000059472.9
ENSMUST00000205335.1
ENSMUST00000205823.1
ENSMUST00000206904.1
ENSMUST00000206692.1
Splice Junctions:
no reads
intron reads
intron read counts
splice junction read counts normalized to Mat2a expression
WT HET KO
B
Color Key
Value
WT
KO
HET
E3.5
E4.5 Epiblast
E4.5 Primitive endoderm
E5.5 Epiblast
primitive streak
E6.5 Visceral endoderm
E6.5 Epiblast
C
E3.5
E4.5 Epiblast
E4.5 PrE
E5.5 Epiblast
primitive streak
E6.5 visceral endoderm
E6.5 Epiblast
mean logFC
KO vs. WT
KO vs. HET
D
Ten most significantly upregulated genes in KO vs. HET comparison
Gtse1
Snora65
Inpp5d
Rbm39
Phlda3
Smndc1
Snapc4
Cfap74
Rab40c
Prpf3
normalized fragment counts
WT HET KO
Ten most significantly dowregulated genes in KO vs. HET comparison
Dxo
Rnf5
Exosc8
Itgae
Lifr
Rars2
Cdc45
Ngrn
Rrp1b
BC017158

Figure-S6

Table S1. DNA and RNA oligonucleotides used in this study.
Related to STAR Methods and Figure 1,2 and 3.

DNA oligos		
Name	**Sequence**	**Comments**
CRISPR F primer	GAAATTAATACGACTCACTATAGGGAGTAGTTTTGTCCGAATCCG TTTTAGAGCTAGAAATAGC	RRoligo770
CRISPR sgR primer	CAAAATCTCGATCTTTATCGTTCAATTTTATTCCGATCAGGCAAT AGTTGAACTTTTTCACCGTGGCTCAGCCACGAAAA	
MMoligo109	ATTCCCACAGTCCCCTTGAG	Genotyping PCR
MMoligo110	CAATGCCCAACCAACACAGA	Genotyping PCR
XXXXX	AGTTGAGAATGCAAAACCTATGGAAGTAAGAACCACCTACCTATG TCAATTCCTCTTCGGAGAGTAGTTTTGTCCGAATCC***TCATCTATT TACAT***TGGTGACCAATCAAATCTTCTACCCAGTGAATATAGTTAA GTCTCAAGGGGACTGTGGGAATTAGTCTCTCCAAA	ssDNA repair template
MMoligo81	TAATACGACTCACTATAGGGTGTTAGCCTTTTTTCCCCAGACTTG TTGGCGTAGGCTACA	MAT2A hp1 template forward
MMoligo82	GGAAGGAGGGCCCTTTCCCTCAGAGCTTGAAGGCTTCTCTGTAGC CTACGCCAACAAGTC	MAT2A hp1 template reverse
MMoligo83	TAATACGACTCACTATAGGGAAAATTGGAACGATACAGAGAAGAT TAGCATGGCCCCTGC	U6 snRNA template forward
MMoligo84	AAAATATGGAACGCTTCACGAATTTGCGTGTCATCCTTGCGCAGG GGCCATGCTAATCTT	U6 snRNA template reverse
RNA oligos		
Lab Name	**Sequence**	**Name in this study**
MMRNA6	UGUUGGCGUAGGCUACAGAGAAGCCUUCA	RNA6 ; 29nt
MMRNA11	UGUUGGCGUAGGCUACUGAGAAGCCUUCA	RNA6-mut ; 29nt
MMRNA7	GGCGUAGGCUACAGAGAAGCC	RNA7 ; 21nt
MMRNA8	CGUAGGCUACAGAGAAG	RNA8 ; 17nt
MMRNA5	CCCCCCCGUAGGCUACAGAGAAGGGGGGG	RNA5 ; 29nt
MMRNA12	GUAGGCUACAGAGAA	RNA12 ; 15nt
MMRNA9	UUUUUACAGAGAAUUUU	RNA9 ; 17nt
MMRNA10	UUUUUCAGUUUUU	RNA10 ;13nt
MMRNA16	GGGGUACAGAGAAGGGG	RNA16 ; 17nt
MMRNA14	N11-UACAGAGAA-N10	Library;30nt; N= randomized
MMRNA21	CAGUAGCGACGUACAGAGAAACAUCUUCUC	RNA21 ; 30nt
MMRNA22	CAGUAGCGACGUACUGAGAAACAUCUUCUC	RNA22 ; 30nt
MMRNA23	CCGGCUUAGCCUACAGAGAACCUUUCUCGU	RNA23 ; 30nt
MMRNA24	CCGGCUUAGCCUACUGAGAACCUUUCUCGU	RNA24 ; 30nt
MMRNA28	GGGCGUAGGCNACANNNNNGCCC	RNA28 ; 23 nt ; N= randomized
MMRNA29	GGGCNNAGGCNNNANNNNNGCCC	RNA29 ; 23 nt ; N= randomized
MMRNA30	GGCGUAGGCUACAGACUUGCC	RNA30 ; 21 nt
MET1 RNA	UACACUCGAUCUGGACUAAAGCUGCUC	METTL3/14 substrate
MET2 RNA	UACACUCGAUCUGGAUUAAAGCUGCUC	METTL3/14 substrate(mut)
RP_RNA_1	UGACAUGAACACAGGUGCUCAGAUAGCUUU	RNA marker- 30nt
RP_RNA_3	UGACAUGAACACAGGUGCUCAGAUAGCU	RNA marker-28nt
RP_RNA_18	AGCACCGUAAAGACGC	RNA marker-16nt
RP_RNA_19	GCGUCUUUACGGUGCUUAAAACAAAACAAAACAAAACAAA	RNA marker-40nt

Supplemental Table S2. List of all deep-sequencing libraries created in this study. Related to STAR Methods

experiment	sample	*Mettl16* genotype	reads
RNASeq of E2.5 embryo	MM14	HET	34679836
RNASeq of E2.5 embryo	MM15	HET	28734242
RNASeq of E2.5 embryo	MM16	KO	34160424
RNASeq of E2.5 embryo	MM17	KO	36085991
RNASeq of E2.5 embryo	MM18	HET	33601058
RNASeq of E2.5 embryo	MM19	HET	32744958
RNASeq of E2.5 embryo	MM20	KO	36987647
RNASeq of E2.5 embryo	MM21	HET	33169466
RNASeq of E2.5 embryo	MM22	HET	32156604
RNASeq of E2.5 embryo	MM23	KO	32000383
RNASeq of E2.5 embryo	MM24	KO	30556292
RNASeq of E2.5 embryo	MM25	WT	23023349
RNASeq of E2.5 embryo	MM26	HET	33038807
RNASeq of E2.5 embryo	MM27	WT	31126991
RNASeq of E2.5 embryo	MM28	HET	32684662
RNASeq of E2.5 embryo	MM29	KO	29378767
RNASeq of E2.5 embryo	MM31	HET	35120157
RNASeq of E2.5 embryo	MM32	WT	32219943
RNASeq of E2.5 embryo	MM33	HET	32009135
RNASeq of E2.5 embryo	MM34	KO	32613795
RNASeq of E2.5 embryo	MM35	KO	29911795
RNASeq of E2.5 embryo	MM36	KO	35371390
RNASeq of E2.5 embryo	MM37	HET	34461243
RNASeq of E2.5 embryo	MM38	KO	34035525
RNASeq of E2.5 embryo	MM39	HET	33553354
RNASeq of E2.5 embryo	MM40	KO	31275672
RNASeq of E2.5 embryo	MM41	KO	32321617
RNASeq of E2.5 embryo	MM42	WT	34131480
RNASeq of E2.5 embryo	MM43	HET	33347148
RNASeq of E2.5 embryo	MM44	HET	35295970
RNASeq of E2.5 embryo	MM45	WT	25935853

experiment	sample	*Mettl16* genotype	reads
RNASeq of E3.5 embryo	MM46	WT	27418677
RNASeq of E3.5 embryo	MM47	WT	27497901
RNASeq of E3.5 embryo	MM48	KO	30804681
RNASeq of E3.5 embryo	MM49	KO	31884501
RNASeq of E3.5 embryo	MM50	KO	34965952
RNASeq of E3.5 embryo	MM51	HET	33312807
RNASeq of E3.5 embryo	MM52	KO	28476667
RNASeq of E3.5 embryo	MM53	WT	31001004
RNASeq of E3.5 embryo	MM54	WT	26071380
RNASeq of E3.5 embryo	MM55	KO	31392715
RNASeq of E3.5 embryo	MM56	HET	26518367
RNASeq of E3.5 embryo	MM57	WT	36455849
RNASeq of E3.5 embryo	MM59	WT	35482923
RNASeq of E3.5 embryo	MM60	KO	31141823
RNASeq of E3.5 embryo	MM61	KO	30971406
RNASeq of E3.5 embryo	MM62	HET	31630583
RNASeq of E3.5 embryo	MM63	WT	42925667
RNASeq of E3.5 embryo	MM64	WT	46540640
RNASeq of E3.5 embryo	MM65	HET	47186347
RNASeq of E3.5 embryo	MM66	KO	46106111
RNASeq of E3.5 embryo	MM67	HET	42360160
RNASeq of E3.5 embryo	MM68	WT	44228364
RNASeq of E3.5 embryo	MM70	KO	34314476
RNASeq of E3.5 embryo	MM71	HET	38949449
RNASeq of E3.5 embryo	MM72	WT	38861981
RNASeq of E3.5 embryo	MM73	WT	51611166
RNASeq of E3.5 embryo	MM74	WT	50932521
RNASeq of E3.5 embryo	MM76	WT	45598521
RNASeq of E3.5 embryo	MM78	KO	27609275
RNASeq of E3.5 embryo	MM79	KO	38122632
RNASeq of E3.5 embryo	MM81	WT	43648518
RNASeq of E3.5 embryo	MM82	WT	37246976
RNASeq of E3.5 embryo	MM83	NA	39136258
RNASeq of E3.5 embryo	MM84	WT	36941697
RNASeq of E3.5 embryo	MM87	KO	24036687
RNASeq of E3.5 embryo	MM88	WT	103815156
RNASeq of E3.5 embryo	MM89	HET	34466973
RNASeq of E3.5 embryo	MM90	HET	62009799
RNASeq of E3.5 embryo	MM91	WT	39750111
RNASeq of E3.5 embryo	MM92	HET	39921773

experiment	sample	description	reads	filtered reads
METTL16 in vitro IP and RNASeq	RR582	Input sample 11	42755991	21312366
METTL16 in vitro IP and RNASeq	RR583	Input sample 12	37993827	18869613
METTL16 in vitro IP and RNASeq	RR586	m6A IP sample 11	42519953	23104554
METTL16 in vitro IP and RNASeq	RR587	m6A IP sample 12	36825565	1203361

SUPPLEMENTAL FIGURE LEGENDS

Figure S1. Full-length human METTL16 exists as monomers. Related to Figure 1.

(A) Protein sequence alignment of the methyltransferase domain of METTL16 proteins. MET16, METTL16; h, human (mammal); m, mouse (mammal); g, *Gallus* (bird); x, *Xenopus* (amphibian); z, zebrafish (fish); c, *C.elegans* (nematode). The secondary structure features present in the human METTL16-core (PDB ID: 6GFN) are indicated above: α helices, β-strands and η-3_{10} helix. Residues (marked with green asterisks) in the putative RNA-binding groove where mutated and shown to affect methylation activity and/or RNA-binding (tested for MUT1 only, in Figure 2F). Deletion of the disordered loop or other mutated residues (red asterisks) also abolish activity. (B) Gel-filtration chromatography profile for indicated proteins on two different Superdex columns (S75 and S200). The elution profile of the proteins are consistent with them being a monomer. The full-length (FL) METTL16 was produced in insect cells, while the METTL16-core was expressed in *E.coli*. (C) Purification of full-length human METTL3/METTL14 complex. See STAR Methods. (D) In vitro methylation assay with ^{14}C-SAM, METTL16-FL or the METTL3+METTL14 complex, and indicated RNAs. The single-stranded RNA (ssRNA) used for METTL3+14 reaction is MET1RNA, while full-length human U6 snRNA and full-length human *MAT2A* hairpin (hp) 1 were in vitro transcribed (STAR methods), and others were purchased (Table S1). Single-stranded RNA markers (size in nucleotides, nt) are ^{32}P-end-labelled. (E) Limited proteolysis of METTL16-FL with indicated proteases. An Instant Blue-stained SDS-PAGE with aliquots of the time-course reaction, with incubation times in minutes (min) and protein markers (in kDa) is shown. Protein boundaries of the fragments were identified by mass spectrometry from bands (indicated by arrows) in the gel. (F) Comparison of our crystal structure of the human METTL16 core (Green/Magenta; PDB ID: 6GFN) with the one recently published by Ruszkowska et al. (Grey; PDB ID: 6B92). A comparison reveals a very high degree of overlap between two structures with only minor differences (see STAR Methods). (G) Electron density for key regions identified in the METTL16-core structure. Shown on left, a 2Fo-Fc map contoured at 1.2 σ and coloured in blue for the METTL16 N-terminal region, A9 to Y13 (PDB: 6GTS). Shown on right, a 2Fo-Fc omit map contoured at 2 σ and coloured in blue for the METTL16 core SAH-binding site (PDB: 6GFN).

Figure S2. Mutational analysis of the human METTL16-core methyltransferase domain to define the RNA-binding groove. Related to Figures 1 and 2.

(A) Cartoon indicating protein domains of human METTL16. Boundaries of the two constructs crystallized in this study are shown (in green). A zoom of the catalytic domain in human METTL16-core domain showing the bound S-adenosyl-homocysteine (SAH) (PDB ID: 6GFK). Residues coordinating the SAH and the catalytic residues N184, P185, P186 are highlighted. (B) Predicted structures of various synthetic RNAs used in in vitro methylation assays. They are all derived from the human *MAT2A* hairpin (hp) 1 and the consensus methylation motif of METTL16 is highlighted (red). The adenosine (A) that is methylated is indicated in bold and this is mutated to a uridine (U) in some of the RNAs. RNA5 has a reinforced stem with artificial G:C pairs. (C) In vitro methylation assay with ^{14}C-SAM, METTL16 proteins and indicated RNAs. *MAT2A* hp1 was in vitro transcribed, while the others were purchased (Table S1). Single-stranded RNA markers (size in nucleotides, nt) are ^{32}P-end-labelled. Note that the METTL16-ΔN protein is inactive. (D) Cartoon showing the N-terminal 20 amino acids of human METTL16, with the positively charged residues that were individually mutated being highlighted (red with asterisks). In vitro methylation assay with wildtype (WT) or mutant METTL16-core versions indicated. Individual mutations of N-terminal positively charged residues to neutral alanine do not affect activity. See also Figure 2D. Catalytic-dead

mutations PP185-186AA and F187G result in absence of any activity. Quality of recombinant proteins used is shown (on the right). (E) In vitro methylation assay with human METTL16-core MTase proteins carrying mutations within the putative RNA-binding groove. The groove is outlined on the surface charge representation of the METTL16-core, on the right (this is same as shown in Figure 2C). Quality of the proteins used is shown. Note that almost all mutations abolish methylation activity. The RNA used are indicated (Table S1). (F) SDS-PAGE gel of the METTL16-core with mutations in the disordered loop. One of the proteins (Loop-3R-E) displays a retarded migration. Gel-filtration profile of the protein shows that its elution profile is same as the one seen for the wildtype protein or other mutants. See also Figure 2G.

Figure S3. In vitro methylation with human METTL16-FL and a randomized RNA library reveals structural and sequence requirements for m^{6}A RNA methylation. Related to Figure 3.

(A) Methylene blue dye staining of the same gel as shown in Figure 3A to reveal RNAs present in the reactions after in vitro methylation assays. After staining with the dye, gels were dried and exposed for detection of radioactivity, and is presented in Figure 3A. This shows the presence of RNAs of the expected sizes in all the lanes. RNA9 and 10 (uridine-rich sequences) are poorly stained with the dye. On the right, a few of the RNAs used were ^{32}P-end-labelled to reveal the integrity of the RNAs used (including RNA9 and RNA10). (B) A repetition of similar experiment as shown in Figure 3A, showing that dramatic truncations to the stem region of the *MAT2A* hairpin RNA abolishes activity. Note that RNA5 was not included in this experiment. (C) A randomized RNA library carrying the nonamer consensus sequence was incubated with METTL16-FL and m^{6}A-containing RNAs were enriched by immunoprecipitation (IP) and sequenced. See Figure 3D. Frequency of RNA oligos forming stem, loop and other selected features at individual positions is compared between m^{6}A-IP and input oligos. Methylated oligos have higher frequency of 15A (which is the adenosine predicted to be methylated in the nonamer consensus motif) in a bulge and surrounded by stems. See also Figure 3G. (D) Log_2 differences of oligo frequencies between m^{6}A-IP and input oligos indicate the 15A bulge and its positioning in between stem structures is important for being recognized and methylated by METTL16. Two replicate experiments show the same pattern. (E) Comparison of the nucleotide frequencies at individual oligo positions surrounding the consensus motif reveals higher G and C occurrence in methylated oligos. (F) Sequence motifs that were identified in the m^{6}A-IP-enriched 11-mers or 10-mers surrounding the consensus motif.

Figure S4. Embryonic lethality in *Mettl16* knockout mice around implantation stage. Related to Figure 4.

(A) Strategy for insertion of a triple-stop codon cassette into exon 3 of mouse *Mettl16* genomic locus using a guide RNA (gRNA) that targets the Cas9 DNA endonuclease. Homologous recombination (HR) introduces the DNA repair template with the triple-stop codon cassette into exon 3, disrupting the coding sequence, creating a knockout (KO) allele. Two independent lines were obtained and both showed identical embryonic lethality phenotype. Line#1 was used in this study for sequence analysis of embryos. (B) Genotyping PCR with mouse tail DNA from indicated mutant lines and wildtype. See STAR methods for PCR conditions. (C) Genotype of animals recovered in litters at the weaning stage (P21, post-natal day 21) from a cross of *Mettl16*$^{+/-}$ heterozygous (HET) parents. No homozygous *Mettl16*$^{-/-}$ knockout (KO) animals were present, indicating early lethality. (D) Genotyping of embryos collected at E6.5 from *Mettl16*$^{+/-}$ females crossed with *Mettl16*$^{+/-}$ males. Note that only one KO embryo was recovered. (E) Genotyping of embryos collected at E8.5 from *Mettl16*$^{+/-}$ females crossed with *Mettl16*$^{+/-}$ males. These studies reveal that loss of mouse *Mettl16* results in embryonic lethality around implantation stage. See also Figure 4B. (F) Examination of E12.5 embryos by genotyping. No KO embryos were detected, indicating early lethality. Note that only one litter was examined at E12.5. Scale bar in millimetre (mm) is shown.

Figure S5. Lack of METTL16 has very specific and limited effect on the transcriptome of E2.5 embryos. Related to Figure 4.

(A) Lack of METTL16 does not affect the global transcription from exons, introns and repeats. Error bars refer to standard deviation. (B) Only the transcripts of four genes have significantly differential abundance between in *Mettl16*$^{-/-}$ (KO) and *Mettl16*$^{+/-}$ (HET) and also between *Mettl16*$^{-/-}$ (KO) and *Mettl16*$^{+/+}$ (WT). (C) Boxplot of two genes significantly downregulated in *Mettl16*$^{-/-}$ (KO). Both of the genes lie on chromosome 11, same as the targeted *Mettl16*. (D) Normalized read coverage of *Mat2a* introns is shown together with the overall read counts for each intron. The intron read counts were normalized either just to library sizes (bottom row) or to *Mat2a* levels (middle row). One of the boxplots (marked with a red line) is reproduced in Figure 4H. (E) Normalized read coverage of the loci with three of the genes (except *Mat2a*) found to be significantly differentially expressed in *Mettl16*$^{-/-}$(KO). Only exon coverage is shown for *Mettl16*. Note that both *Gm15698* and *Ccdc92b* lie on chromosome 11 as well as *Mettl16* and their differential expression might be just the consequence of different chromosome 11 DNA sequence resulting from the mixed genetic background of parental strain which was crossed to pure B6 mouse (see STAR Methods). (F) Mass spectrometry analysis of 3xFLAG-HA-hMETTL16 immunopurified from transfected HEK293T cells. 3xFLAG-HA-GFP was used as a negative control. Presented protein hits are the top 15 most enriched proteins in the hMETTL16 IP when compared to the negative control (3xFLAG-HA-GFP). HA-tag purification was performed (STAR METHODS). Exclusive Unique Peptide Count shown. Protein threshold: 1% FDR; min. peptides: 2; Peptide threshold: 0.1% FDR. (G) Mass spectrometry analysis of endogenous mMETTL16 immunopurified from adult mouse testes and spleen. Presented protein hits are the top 15 most enriched proteins in the mMETTL16 IP when compared to the negative control. Multiple splicing factor 3B and 3A subunits are among the top candidates. Purification was done using anti-METTL16 antibody (abcam, ab186012) without RNase-treatment, negative control was beads bound with normal mouse IgG (Santa Cruz, sc-2025). Exclusive Unique Peptide Count shown. Protein threshold: 1% FDR; min. peptides: 2; Peptide threshold: 0.1% FDR.

Figure S6. Dramatically altered transcriptome of E3.5 *Mettl16* knockout embryos. Related to Figure 5.

(A) Counts of the reads spanning the splice junction (SJ) of last *Mat2a* (ENSMUST00000059472.9) intron are plotted together with read counts originating in the intron. Boxplots are shown where counts of individual samples are plotted as dots. (B) The heatmap shows the expression of transcription factors and chromatin modifiers enriched for distinct embryonic lineages [as defined by (Mohammed et al., 2017)] in individual samples. Expression of these key factors is unaltered in the *Mettl16* KO embryos. (C) Boxplots display the average log2 fold-changes (FC) between *Mettl16* mutant (KO) and control (WT or HET) mice for the groups of lineage-specific genes. Dots depict the individual genes. None of the group of lineage-specific genes seems to be differentially expressed between *Mettl16*$^{-/-}$ and *Mettl16*$^{+/-}$ or *Mettl16*$^{-/-}$ and *Mettl16*$^{+/+}$. (D) Boxplots compare the expression of the top 10 downregulated or upregulated genes in the *Mettl16*$^{-/-}$ KO blastocysts. Transcript levels of individual samples are shown as dots.

SUPPLEMENTAL TABLE LEGENDS

Table S1. DNA primers and RNA oligonucleotides used in this study. Related to STAR Methods and Figures 1-5.

Table S2. List of all deep-sequencing libraries created in this study. Related to STAR Methods and Figures 3-5.

Data is available from GEO under accession no. GSE116329.

7.2. Chapter II – Regulation of splicing by m^{6}A methylation

This chapter consists of a peer-reviewed article entitled "Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing", published in Cell journal in June 2021. In this study, we show that *C. elegans* homologue of METTL16, METT-10, is an m^{6}A mRNA methyltransferase that methylates U6 snRNA and *sams-3/-4/-5* transcripts encoding for SAM synthetases. The *sams* transcripts are m^{6}A methylated at the 3′SS of intron 2, with methylation correlated with impaired splicing. We show that splicing is impaired because m^{6}A blocks the binding of the U2AF35 splicing factor, inhibiting splice site recognition. Methylation of this site is regulated by diet and is essential for the maintenance of stable SAM levels. Next, we show that although in mammals, SAM synthetase transcript splicing is regulated differently, the mechanism of U2AF35 inhibition by m^{6}A is conserved. We identify approximately 1000 potentially regulated 3′SS in mice and show that two of them might be regulated during mouse embryonic development.

I contributed to this project by being the lead author, who performed most of the experiments, discussed the project and particular experiments with Ramesh Pillai, David Homolka and other co-authors as well as coordinated all the experiments. I generated and maintained *Mettl16* catalytic-dead mouse mutants, performed all biochemical assays and cell culture experiments, analysed worm experiments (RNA isolation and RT-PCR). Kamila Delaney generated, maintained, and collected all the *C. elegans* lines and conducted all worm experiments with help from Joanna Wenda and Florian Steiner. Raman Radha Pandey prepared all m^{6}A-IP experiments as well as RNA for MS analysis, Kuan-Ming Chen produced all recombinant proteins and performed ITC measurements, Cathrine Broberg Vagbo did RNA mass spectrometry analysis, David Homolka conducted all computational analyses. The manuscript was written by Ramesh Pillai, with my and other authors input. I was involved in editing the manuscript at every stage of the publication process.

The manuscript was not modified for the purpose of this thesis and thus figure numeration and bibliography are separate from the rest of the thesis.

The graphic on the next page is our cover design proposal originally submitted to the journal. The cover was created by Marzia Munafò (https://www.munafomarzia.com, Twitter: @munafomarzia).

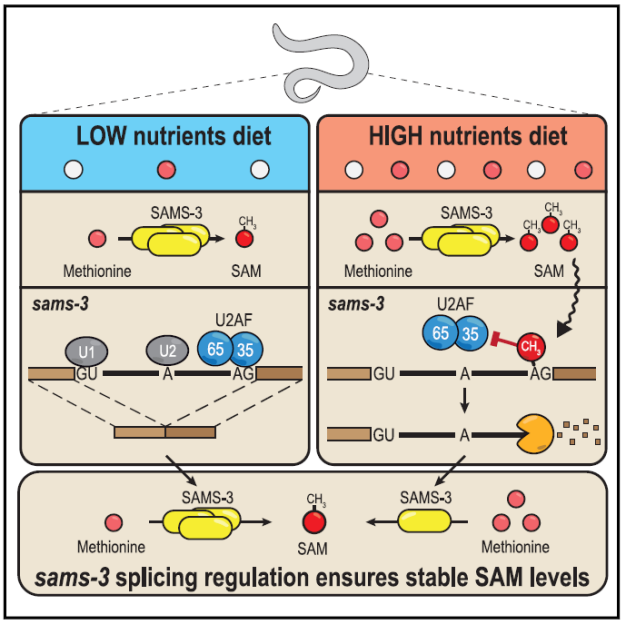

Cell
Cell
Volume XX
Number XX
Month XX, 2020
Volume XX Number XX Pages 000–000 Month XX, 2020
HN
CH3
CellPress

Cell

Article

Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing

Graphical abstract



Highlights

- m^{6}A deposited at 3′ splice site by worm METT-10 inhibits splicing
- Methylation blocks 3′ splice site recognition by splicing factor U2AF35
- Methylation and splicing inhibition is a response to change in worm diet
- Splicing inhibition by 3′ splice site m^{6}A is conserved in mammals

Authors

Mateusz Mendel, Kamila Delaney, Radha Raman Pandey, ..., Florian A. Steiner, David Homolka, Ramesh S. Pillai

Correspondence

david.homolka@unige.ch (D.H.), ramesh.pillai@unige.ch (R.S.P.)

In brief

m^{6}A methylation of a 3′ splice site blocks its recognition by splicing factors to inhibit pre-mRNA splicing in nematodes and mammals. In worms, this mechanism is used to modulate splicing in response to a change in diet.

Mendel et al., 2021, Cell *184*, 1–18
June 10, 2021
https://doi.org/10.1016/j.cell.2021.03.062

CellPress

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

Cell

CellPress
OPEN ACCESS

Article

Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing

Mateusz Mendel,[1] Kamila Delaney,[1] Radha Raman Pandey,[1] Kuan-Ming Chen,[1] Joanna M. Wenda,[1] Cathrine Broberg Vågbø,[2] Florian A. Steiner,[1] David Homolka,[1,*] and Ramesh S. Pillai[1,3,*]

[1]Department of Molecular Biology, Science III, University of Geneva, 30 Quai Ernest-Ansermet, 1211 Geneva 4, Switzerland

[2]Proteomics and Modomics Experimental Core (PROMEC), Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology (NTNU) and St. Olavs Hospital Central Staff, Trondheim, Norway

[3]Lead contact

*Correspondence: david.homolka@unige.ch (D.H.), ramesh.pillai@unige.ch (R.S.P.)

https://doi.org/10.1016/j.cell.2021.03.062

SUMMARY

The N^6-methyladenosine (m^{6}A) RNA modification is used widely to alter the fate of mRNAs. Here we demonstrate that the *C. elegans* writer METT-10 (the ortholog of mouse METTL16) deposits an m^{6}A mark on the 3′ splice site (AG) of the S-adenosylmethionine (SAM) synthetase pre-mRNA, which inhibits its proper splicing and protein production. The mechanism is triggered by a rich diet and acts as an m^{6}A-mediated switch to stop SAM production and regulate its homeostasis. Although the mammalian SAM synthetase pre-mRNA is not regulated via this mechanism, we show that splicing inhibition by 3′ splice site m^{6}A is conserved in mammals. The modification functions by physically preventing the essential splicing factor U2AF35 from recognizing the 3′ splice site. We propose that use of splice-site m^{6}A is an ancient mechanism for splicing regulation.

INTRODUCTION

It has been known since the early 1970s that RNAs can be modified with N^6-methyladenosine (m^{6}A) (Desrosiers et al., 1974, 1975; Schibler et al., 1977; Wei and Moss, 1977; Wei et al., 1975a, 1975b). It is the most abundant internal modification on eukaryotic mRNA (Fu et al., 2014; Patil et al., 2018; Roignant and Soller, 2017), with ~4 m^{6}A/10^4 nucleotides (nt) detected in poly(A)$^+$ RNA from adult mouse testes (Pandey et al., 2020). The mammalian heterodimeric METTL3/METTL14 RNA methyltransferase complex is the dominant m^{6}A "writer," with orthologs in organisms such as yeast, flies, and plants (Liu et al., 2014; Śledź and Jinek, 2016; Wang et al., 2016). The complex installs the m^{6}A mark within a loosely defined RR**m^{6}A**CH motif at thousands of sites in the transcriptome, with a bias towards the 3′ end of the RNA, where it is enriched near the stop codon (Dominissini et al., 2012; Kan et al., 2017; Meyer et al., 2012; Schwartz et al., 2013). m^{6}A marks are recognized by various "reader" proteins, like those belonging to the YTH family (Patil et al., 2018), to modulate RNA splicing, stability, and translation (Li et al., 2014; Theler et al., 2014; Zhang et al., 2010). Gene regulation by this writer-reader system is essential for embryonic development in plants and mice (Batista et al., 2014; Geula et al., 2015; Kasowitz et al., 2018; Lasman et al., 2020; Zhong et al., 2008), mammalian fertility (Hsu et al., 2017; Ivanova et al., 2017; Jain et al., 2018; Wojtas et al., 2017), sex determination in flies (Haussmann et al., 2016; Lence et al., 2016), and many other developmental processes. Notably, this m^{6}A writer-reader system is absent in nematodes.

The second mRNA m^{6}A writer, METTL16, is highly conserved, with current knowledge of the enzyme coming from investigation of the protein in mammals. METTL16 (Brown et al., 2016) has a very strict requirement for target methylation because it methylates an adenosine within a nonamer consensus motif (UAC**m^6**AGAGAA) only when it is present in a structured RNA context (Doxtader et al., 2018; Mendel et al., 2018; Pendleton et al., 2017). S-adenosylmethionine (SAM) synthetase *MAT2A* mRNA and the spliceosomal U6 small nuclear RNA (snRNA) are the two known targets of mammalian METTL16 (Pendleton et al., 2017; Warda et al., 2017). SAM synthetase is the enzyme responsible for production of the methyl donor SAM, which is required for methylation reactions in the cell. In the case of human *MAT2A* mRNA, there are six methylation sites in the 3′ UTR, each with the motif occupying the single-stranded region of a stem-loop structure (Pendleton et al., 2017; Shima et al., 2017; Warda et al., 2017). Methylation of these sites has been proposed to recruit the nuclear reader protein YTHDC1, which promotes decay of the *MAT2A* mRNA (Shima et al., 2017). However, the central gene-regulatory role of METTL16 appears to be non-catalytic because it has been shown to bind the stem-loop structure to promote splicing of a frequently retained terminal intron (Pendleton et al., 2017). Efficient splicing is critical to produce the MAT2A enzyme and maintain cellular SAM levels. Mammalian METTL16 has a highly conserved N-terminal RNA methyltransferase domain and a C-terminal region that is present only in vertebrates. Importantly, this non-catalytic C-terminal vertebrate-conserved region (VCR) of METTL16 is critical for

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

splicing regulation of the human SAM synthetase *MAT2A* mRNA (Pendleton et al., 2017). Supporting such a non-catalytic splicing stimulation role, loss of mouse *Mettl16* leads to reduced levels of mature *Mat2a* mRNA, causing pre-implantation embryonic lethality (Mendel et al., 2018). This raises the question of the relevance of METTL16's catalytic activity, which is conserved from bacteria to humans. By investigating the invertebrate and vertebrate orthologs of the enzyme, our study identifies 3′ splice-site m^{6}A methylation as a conserved mechanism to regulate splicing.

RESULTS

The m^{6}A transcriptome of *C. elegans*

To study the conserved role of the catalytic activity of METTL16, we chose the nematode *Caenorhabditis elegans* (hereafter referred to as worm). The worm ortholog METT-10 (Dorsett et al., 2009) contains the highly conserved RNA methyltransferase domain (Figure S1A) but lacks the VCRs found in mammalian METTL16 (Figure 1A). We began the study by detecting various ribose and base modifications in total and poly(A)$^+$ RNAs from adult worms (Figure 1B; STAR Methods). RNA from adult mouse testes and an insect cell line (*Bombyx mori* BmN4 cells) were used for comparison. The m^{6}A modification is detected in poly(A)$^+$ RNA from all three biological sources, including *C. elegans* (Figure 1B), which is important for this study.

To identify worm transcripts that carry the m^{6}A methylation, we carried out m^{6}A-IP-seq (Ke et al., 2015) with a mixture of poly(A)$^+$ RNAs from adult *C. elegans* and mouse testicular RNA (Figure 1C; Table S1; STAR Methods). The mouse RNA serves as an internal technical control because m^{6}A sites are already mapped in this system (Wojtas et al., 2017). Compared with over 20,000 mouse peaks, we identified only 176 m^{6}A peaks in the worm poly(A)$^+$ transcriptome (Figures S1B–S1D), which likely reflects the absence of the METTL3/METTL14 writer complex in worms (Sendinc et al., 2020; van Delft et al., 2017). Indeed, a motif analysis of the mouse peaks reveals the presence of the expected RRACH context (R = A and G; H = A, C, and U) used by the dominant mammalian METTL3/METTL14 writer (Figure 1D; Dominissini et al., 2012; Ke et al., 2015; Meyer et al., 2012), and this is absent in worms. Meta-analysis of m^{6}A-IP reads mapping to all mouse transcripts produces the typical profile, characterized by high levels of methylation over the coding sequences with peaks at the 5′ end and over the stop codon (Figure 1E). In contrast, such a pattern of m^{6}A distribution is clearly absent over worm sequences (Figure 1E), and a motif search did not recover any particular sequence context for the worm m^{6}A-enriched reads (Figure S1E).

Worm METT-10 is an m^{6}A writer for U6 snRNA and SAM synthetase RNA

Having confirmed the presence of m^{6}A on worm poly(A)$^+$ RNA, we wished to determine the contribution of METT-10 (Dorsett et al., 2009). To search for its methylation targets, we used a comparative analysis of m^{6}A-IP-seq datasets to identify poly(A)$^+$ transcripts that show reduced m^{6}A methylation (m^{6}A-IP reads/input reads) in the *mett-10* knockout (KO) mutant (Figure 1F and S1F; Table S2). Of these, the top 20 encode the U6 snRNA sequences (Figures 1G and S1G). U6 snRNA is a non-polyadenylated transcript, so its presence in the poly(A)$^+$ dataset is likely due to remnants left after poly(A)$^+$ enrichment from total RNA. Consistent with this, a separate m^{6}A-IP-seq experiment conducted with total RNA samples shows a higher enrichment of the U6 snRNA reads (Figure 1H). Human U6 snRNA is methylated within a nonamer motif (UAC**m^{6}A**GAGAA) by human METTL16 (Pendleton et al., 2017; Warda et al., 2017), and mapping of m^{6}A-IP reads shows that the worm U6 snRNA is also methylated within an identical site (Figures 1I and S1G). Importantly, this m^{6}A signal is lost in the *mett-10* KO (Figures 1H and 1I). As an independent validation, we used the SCARLET method, which allows examination of the methylation status in a nucleotide-specific manner (Liu et al., 2013). Analysis of total RNA from adult worms confirms methylation of this specific adenosine within the methylation consensus motif, and this is completely lost in the *mett-10* KO (Figures 1J and S1H). Loss of methylation in the *mett-10* KO has a slightly positive influence on the overall U6 snRNA levels (Figures S1I and S1J).

Other transcripts that display a significant drop in m^{6}A levels in the *mett-10* KO are *sams-3*, *sams-4*, and *sams-5* (Figure 1G). These duplicated genes encode the SAM synthetase, the enzyme responsible for production of the methyl donor SAM, which is required for methylation reactions in the cell. We identify worm METT-10 as an m^{6}A RNA methyltransferase, and, like its mammalian ortholog METTL16, it has U6 snRNA and SAM synthetase RNA as conserved targets.

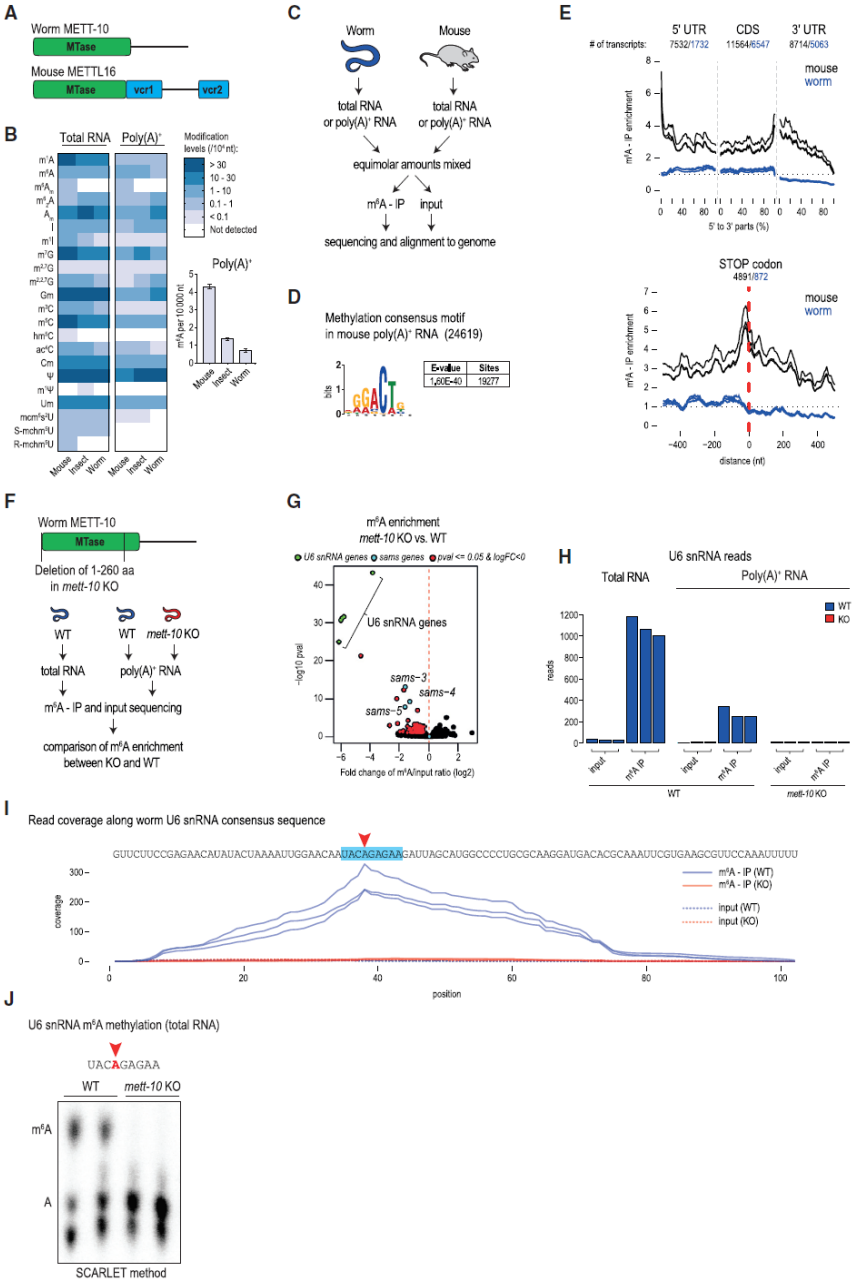
3′ splice site m^{6}A methylation of SAM synthetase pre-mRNA inhibits its splicing

Although mammalian METTL16 and worm METT-10 methylate SAM synthetase RNA, mapping of the m^{6}A-IP reads reveals very different locations for the modification. There are six methylation sites within the 3′ UTR of mammalian *MAT2A* SAM synthetase mRNA (Pendleton et al., 2017; Warda et al., 2017). In contrast, mapping of reads from three independent m^{6}A-IP datasets to the worm genome reveals a single discrete peak over the intron 2/exon 3 junction of the *sams* pre-mRNAs, and this signal is not detectable in the *mett-10* KO (Figure 2A). This peak is seen when reads were mapped over the duplicated *sams-3*, *sams-4*, and *sams-5* genes (Figure S2A). Because *sams-3* and *sams-4* are identical in sequence at this junction region and, hence, indistinguishable, we refer to these genes together in some of the analyses (Figure 2A). Compared with the methylation motif in worm U6 snRNA and in mammalian targets (UAC**m^{6}A**GAGAA), a variant motif is identified at the *sams* m^{6}A peak (UAC**m^{6}A**GAAAC; identical sequences are underlined). Importantly, the methylated adenosine within this motif is at the 3′ splice site (AG) of intron 2 (Figures 2A and S2B).

The significance of this finding became clear when we examined sequence databases. In addition to the mature spliced protein-coding (PC isoform) version of the *sams-3/4* transcript, two noncoding versions that fail to use the 3′ splice site within intron 2 are detected (Figures 2C and S2A). One noncoding version is an alternative splice (AS isoform) variant that uses an upstream cryptic 3′ splice site, and the other is an intron-retained version (IR isoform) that retains the complete intron 2. Compared with wild-type (WT) worms, the overall intron 2 read counts are lower in the *mett-10* KO (Figure 2B), indicating its efficient splicing in

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS



(legend on next page)

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

the absence of 3′ splice site methylation. Consistent with this, quantification of splice junction reads in the RNA sequencing (RNA-seq) datasets shows that this particular 3′ splice site (producing the PC isoform) is used preferentially (~8-fold higher) in the *mett-10* KO, whereas use of the upstream AS site (producing the AS isoform) and intron 2 retention (IR isoform) is higher in the WT (Figure 2C). This suggests that m^{6}A methylation at the 3′ splice site prevents its use and, instead, promotes use of an alternative upstream 3′ splice site or intron retention. The consequence of this m^{6}A-mediated splicing inhibition is a general increase in *sams* mRNA levels in the *mett-10* KO (Figures 2D and S2C). We show that worms use METT-10-mediated 3′ splice site m^{6}A methylation to inhibit splicing and production of SAM synthetase mRNA.

An RNA secondary structure is required for m^{6}A methylation at the 3′ splice site

Methylation by mammalian METTL16 in the *MAT2A* 3′ UTR requires the presence of the methylation consensus motif in the context of a stem-loop structure (Doxtader et al., 2018; Mendel et al., 2018; Pendleton et al., 2017). Similarly, secondary structure prediction shows that a 30-nt RNA fragment of the *sams-3* pre-mRNA that spans the 3′ splice site folds into a stem-loop structure, with the consensus motif (UAC**m^{6}A**GAAAC) occupying part of the loop region (Figure 3A). To confirm that this sequence can be methylated by worm METT-10, we incubated the 30-nt RNA with recombinant full-length worm METT-10 and radioactive ^{14}C-SAM as a methyl donor (Figure 3B). The RNA is methylated specifically at the 3′ splice site (AG) because mutation (A→U) of the adenosine abolishes this activity (Figure 3B, compare RNA-1 with RNA-2). Single or triple (CUU) mutations within the consensus motif (Figure 3B, RNA-4 and RNA-5) also abolish *in vitro* methylation activity of METT-10.

The stem region is also critical for methylation because placement of the motif within a single-stranded context (poly-C flanks) kills all activity, and this cannot be rescued by placing the motif alone within an artificial C:G stem (Figure 3B, RNA-8 and RNA-9). Similarly, large-scale mutations that disrupt the stem cannot be rescued (Figure 3C, RNA-10 and RNA-11). Interestingly, 2-nt mutations (RNA-12 and RNA-14) that disrupt pairing within the stem abolish activity, while compensatory mutations (RNA-13 and RNA-15) that restore pairing at these sites can rescue the activity (Figure 3C). Furthermore, a limited 6-bp artificial C:G stem (RNA-16) in the context of the original sequence supports activity (Figure 3C). These results show that the methylation consensus motif and stem-loop formation at the 3′ splice site of the *sams* pre-mRNA are prerequisites for its recognition by METT-10.

Mutations that abolish 3′ splice site methylation alter splicing *in vivo*

To directly analyze the effect of 3′ splice site m^{6}A methylation *in vivo*, we prepared a wild-type (WT) transgene splicing reporter construct based on *sams-3* (STAR Methods), where the 3′ splice site of intron 2 is methylated by METT-10. We also created a mutant (MUT) version where the methylation consensus motif has the mutations AAC→CUU, which, as we demonstrated, abolish methylation *in vitro* (Figure 3B). These mutations are in exon 3 and do not alter the 3′ splice site. The constructs were injected into the gonads of WT worms, and multiple independent progeny lines showing stable expression of the transgene were established (Figure 3D). Using adult transgenic worms, splicing of these constructs was investigated by reverse-transcriptase polymerase chain reaction (RT-PCR) analysis with transgene-specific primers (Table S3). Each experiment consisted of analysis of three independent progeny lines per construct and was repeated at least three times. We observed three distinct RT-PCR products: the unspliced or IR isoform, the AS isoform, and the correctly spliced mature PC isoform (Figure 3D). The MUT construct, which has mutations preventing m^{6}A methylation (Figure 3B), shows increased use of the 3′ splice site and efficient splicing *in vivo*, as evidenced by higher PC isoform levels and a decrease in the AS isoform (Figure 3D). This demonstrates the direct role of m^{6}A in preventing 3′ splice site recognition and inhibition of splicing. Consistent with the requirement of m^{6}A methylation, there is no difference in splicing between WT and MUT transgenes when expressed in the *mett-10* KO worms (Figure S3A). We show that it is the presence of an m^{6}A at the 3′ splice site and not binding of METT-10 per se that regulates

Figure 1. Worm METT-10 is an m^{6}A writer for U6 snRNA and SAM synthetase mRNA

(A) Domain organization of the m^{6}A writers: mammalian METTL16 and *Caenorhabditis elegans* METT-10. MTase, methyltransferase domain; VCR, vertebrate-conserved region. See also Figure S1A.

(B) Quantification of RNA modifications in total and poly(A)$^+$ RNA from mouse (*Mus musculus*), insect (silkworm, *Bombyx mori*), and worm (*C. elegans*) using liquid chromatography-tandem mass spectrometry (LC-MS/MS). The barplot shows the level of m^{6}A in poly(A)$^+$ RNA.

(C) Scheme for mapping m^{6}A sites catalyzed by worm METT-10 with m^{6}A-IP-seq. Mouse testes RNA is used as an internal control. See also Figure S1B.

(D) The METTL3/METTL14 methylation consensus motif (RRACH) is found on the majority of the mouse m^{6}A peaks (total number of peaks in brackets).

(E) Meta-analysis of the distribution of m^{6}A reads over mouse and worm transcripts.

(F) Scheme for identification of m^{6}A targets of *C. elegans* METT-10 by m^{6}A-IP-seq. See also Figure S1F.

(G) Based on decreased m^{6}A enrichment in *mett-10* KO worms compared with the control wild type (WT), we identified the indicated transcripts to be targets of METT-10. See also Figure S1G.

(H) Worm U6 snRNA is enriched in m^{6}A-IP with total and poly(A)$^+$ RNA, and this enrichment is lost in the *mett-10* KO . The normalized counts (reads per million [rpm]) are plotted separately for biological replicates (n = 3).

(I) Coverage of m^{6}A-enriched reads along the worm U6 snRNA sequence identifies the adenosine (red arrowhead), which is part of the conserved UACm^{6}AGAGAA motif, that is methylated. Methylation is lost in *mett-10* KO worms. The normalized coverages (rpm) from three biological replicates are plotted separately.

(J) Detection of U6 snRNA m^{6}A (red arrowhead) in total RNA from WT control or *mett-10* KO worms (in biological duplicates). The thin-layer chromatography (TLC) analysis used in the SCARLET method (STAR Methods) is shown.

See also Figure S1H.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS

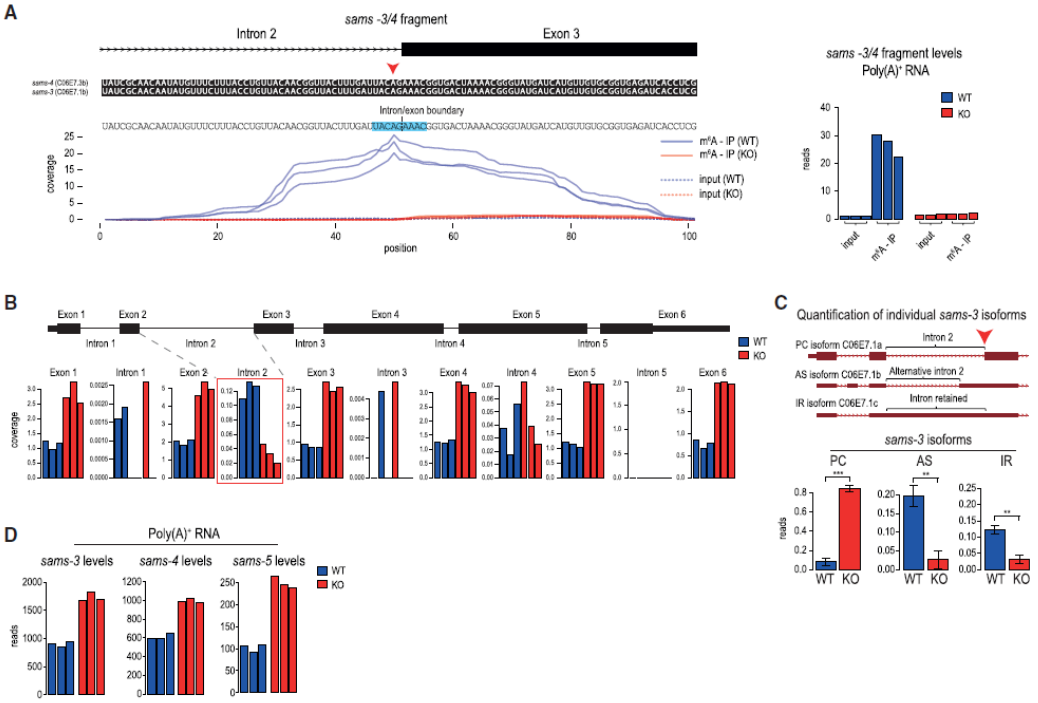


Figure 2. A 3′ splice site m^{6}A inhibits splicing of SAM synthetase pre-mRNA

(A) Mapping of m^{6}A reads identifies the 3′ splice site adenosine (red arrowhead) of intron 2 in the *sams-3/4* pre-mRNA as being methylated, and this methylation is lost in *mett-10* KO worms. The METT-10 methylation consensus motif is highlighted. The normalized coverages (rpm) from three biological replicates are plotted separately. See also Figures S2A and S2B. The barplot shows quantification (rpm) of the reads mapping to the *sams3/4* genomic window .

(B) Normalized read coverage (rpm) along the *sams-3* genomic locus shows uniformly increased exonic coverage and lower intron 2 coverage in the *mett-10* KO, suggesting more efficient splicing. Three biological replicates are plotted separately.

(C) Three *sams-3* isoforms that differ in utilization of the methylated 3′ splice site are annotated in ENSEMBL. Quantification of the different *sams-3* splice isoforms (rpm; STAR Methods) in WT and *mett-10* KO worms shows an increase in the mature, fully spliced PC isoform in the KO. PC, protein-coding; AS, alternative splice; IR, intron-retained. Mean values ± SD are plotted (n = 3). The p values were calculated using t tests. **p ≤ 0.01, ***p ≤ 0.001.

(D) Read counts (DESeq2 normalized) for the three different *sams* genes in the poly(A)$^+$ transcriptome from WT and *mett-10* KO worms show an overall increase in the KO. The three biological replicates are plotted separately.

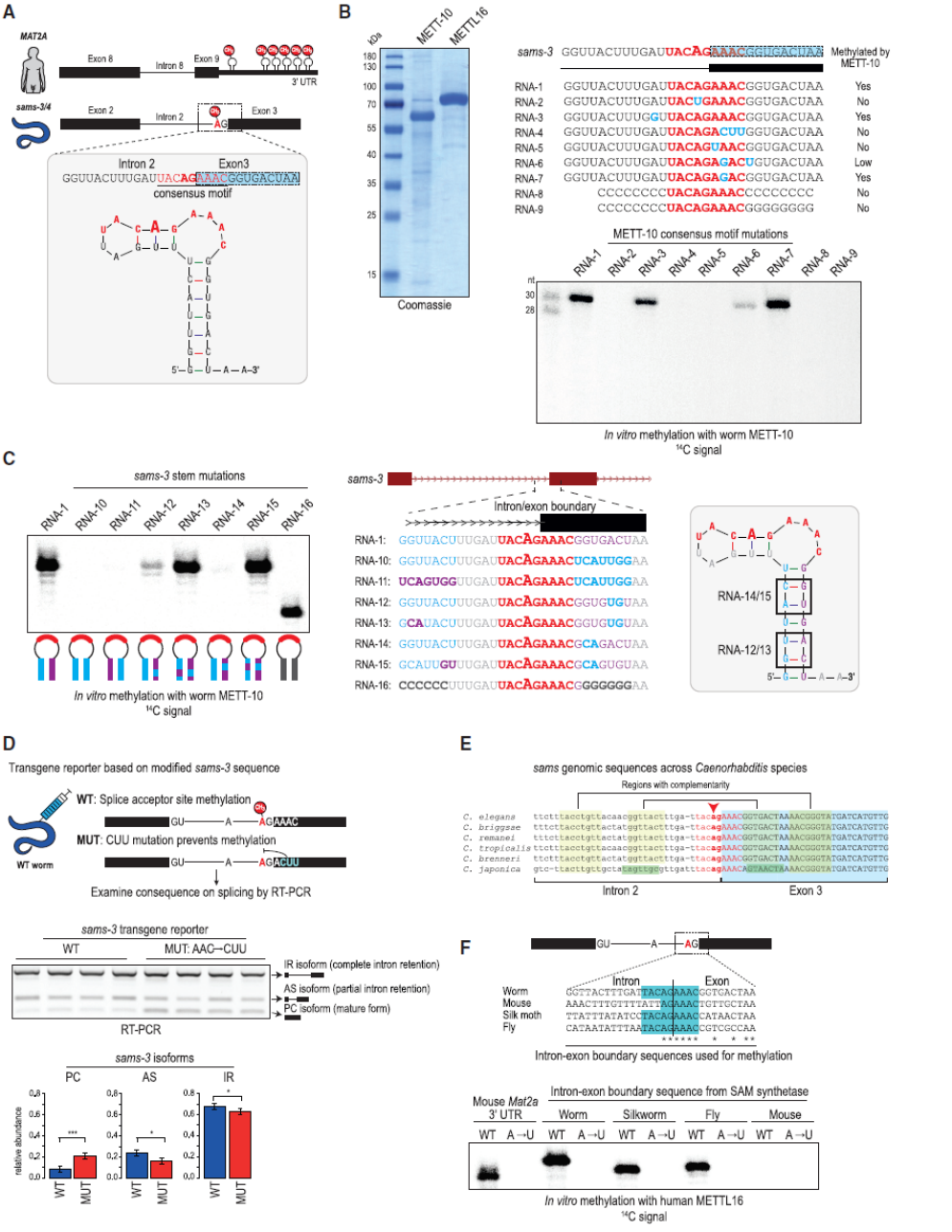
See also Figure S2C.

splicing and expression of the worm SAM synthetase transcript. This is in stark contrast to the mechanism used by METTL16 to regulate mammalian SAM synthetase pre-mRNA (Pendleton et al., 2017).

Interestingly, *sams* gene sequences surrounding the 3′ splice site from various *Caenorhabditis* species show strong conservation of the capacity to form the stem-loop structure, with the METT-10 methylation motif in the loop region (Figure 3E). Indeed, mutations found in the flanking regions in *Caenorhabditis japonica* are compensatory, allowing continued maintenance of pairing. Moreover, the motif can also be found at *sams* splice sites of other invertebrates, like the fruit fly *Drosophila melanogaster* and the silk moth *Bombyx mori* (Figure 3F), indicating potential evolutionary conservation of this type of splicing regulation among invertebrates.

To functionally validate these insect 3′ splice sites as targets for m^{6}A methylation, we carried out *in vitro* methylation assays with a 30-nt RNA spanning the region. We used human METTL16 (Figure 3F) or worm METT-10 (Figure S3B) as enzymes. The insect sequences are methylated specifically at the 3′ splice site (AG) adenosine within the consensus motif because mutation (A→U) of the splice site adenosine abolishes methylation of the RNA (Figure 3F). The homologous junction sequence from mouse *Mat2a* lacks the motif and is not methylated in this experiment, whereas the validated methylation site from the 3′ UTR of mouse *Mat2a* is methylated (Figure 3F). The presence of a conserved methylation motif within a structured RNA at the intron-exon boundary of invertebrate SAM synthetase pre-mRNA transcripts is required for 3′ splice site m^{6}A methylation and splicing regulation.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062



(legend on next page)

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS

Methylation of the *sams* 3′ splice site is triggered by a nutrient-rich diet

C. elegans is a bacterium-eating soil nematode that proliferates on rotting vegetal substrates (Félix and Duveau, 2012; Shtonda and Avery, 2006), but it is maintained in the laboratory on food that consists of different strains (OP50 or NA22) of *Escherichia coli*. For all experiments described above, where we noted 3′ splice site methylation-mediated splicing inhibition, the worms were grown on nutrient-high agar plates (peptone-rich medium + NA22 strain; Table S4). Changing the diet to nutrient-low agar plates (peptone-poor medium + OP50 strain; Table S4) led to the surprising loss of this splicing regulation and a similar isoform expression pattern among WT and *mett-10* KO worms (Figure 4A). RT-PCR analysis shows that intron 2 of the *sams-3* transcript is spliced efficiently in WT worms grown on nutrient-low agar plates, as evidenced by reduced levels of the AS isoform (Figure 4A, lanes 1 and 2 versus lanes 3 and 4). In fact, the splicing pattern in WT worms grown on nutrient-low plates very much resembled the pattern seen in the *mett-10* KO (Figure 4A), as if m^{6}A methylation on the 3′ splice site was absent in WT worms. This diet-dependent change in splicing pattern of endogenous *sams-3* was confirmed by RNA-seq analysis (Figure S4B) and also validated with our transgene reporter constructs based on *sams-3* (Figure S4C).

To directly establish that splice site m^{6}A methylation responds to a change in diet, we carried out m^{6}A-IP-seq with poly(A)$^+$ RNA from WT and *mett-10* KO worms grown on the two different diets. Strikingly, WT worms grown on nutrient-high plates display strong m^{6}A methylation of the 3′ splice site within intron 2 of the *sams-3* pre-mRNA, whereas this is reduced dramatically when WT worms are grown on nutrient-low plates (Figures 4B and S4A). The *mett-10* KO lacked this methylation under all conditions (Figure 4B), and, consequently, the splicing patterns were not altered when worms were grown on the different media (Figures 4A and S4B). This allows us to conclude that 3′ splice site m^{6}A methylation takes place in response to a nutrient-high diet to inhibit proper splicing and expression of SAM synthetase pre-mRNA.

m^{6}A-mediated inhibition of splicing represents negative feedback regulation of SAM levels

Because RNA methylation depends on SAM as a methyl donor, we examined whether the pathway serves to regulate cellular SAM levels by feedback inhibition. To investigate this further, we asked which constituents in the diet are responsible for triggering splice site methylation. Keeping the bacterial strain constant (NA22 or OP50), we prepared plates with nutrient-low medium or peptone-rich nutrient-high medium (Figure 4C). Worms were grown on such plates, and RT-PCR analysis was conducted to examine splicing of intron 2 in the endogenous *sams-3* pre-mRNA transcript. Irrespective of the bacterial strain used, the nutrient-high medium is responsible for strong splicing inhibition, as determined by quantification of the AS isoform (Figure 4C). Nevertheless, the level of splicing inhibition in nutrient-low plates is slightly higher when the NA22 bacterial strain is used, but the major driving factor was still the peptone-rich nutrient-high medium (Figure 4C).

Production of SAM requires enzymatic activities represented in the inter-linked methionine and folate cycles (Figure 4D). Briefly, SAM is produced from ATP and methionine by the SAM synthetase (*sams* in worms) within the methionine cycle, whereas the downstream by-product homocysteine is regenerated to methionine via methionine synthase, which requires folate (5-methyl tetrahydrofolate) and the co-factor vitamin B12. Importantly, the key metabolites, like the essential amino acid methionine, folic acid, and vitamin B12, are all acquired through the diet. Consistent with this, supplementing the nutrient-low medium with additional free methionine or vitamin B12, which directly enhances SAM production via the methionine cycle, triggered splicing inhibition of *sams-3* (as indicated by AS-isoform levels) similar to that seen with the nutrient-high medium (Figure 4C). However, supplementation with amino acids not involved in the methionine cycle (leucine and cysteine) or folic acid, which feeds into the folate cycle, did not lead to splicing inhibition. These results support a model of regulation by feedback inhibition, where constituents in the diet that directly increase cellular SAM levels via the methionine cycle trigger 3′

Figure 3. A conserved stem-loop structure containing the 3′ splice site identifies it for methylation by METT-10

(A) Position of m^{6}A marks introduced by human METTL16 on the 3′ UTR of human *MAT2A* SAM synthetase as well as *C. elegans* METT-10 on the 3′ splice site of worm *sams-3* pre-mRNA. A 30-nt RNA fragment (RNA-1; Table S3) spanning the intron 2-exon 3 boundary of the worm *sams-3* gene is predicted to fold into a stem-loop structure, with the METT-10 methylation motif UACm^{6}AGAAAC (red) present in the loop region. This is very similar to the substrate requirement of mammalian METTL16.

(B) Purification of recombinant worm METT-10 and human METTL16 proteins for *in vitro* methylation assays. Shown are *in vitro* methylation assays with METT-10 and the indicated RNA substrates, based on the *sams-3* intron 2/exon 3 junction sequence, using radioactive ^{14}C-SAM as a methyl donor. The UACAGAAAC motif (red) and residues that were mutated (blue) are highlighted. The reaction products were resolved by PAGE and exposed to detect the radioactivity (^{14}C) signal.

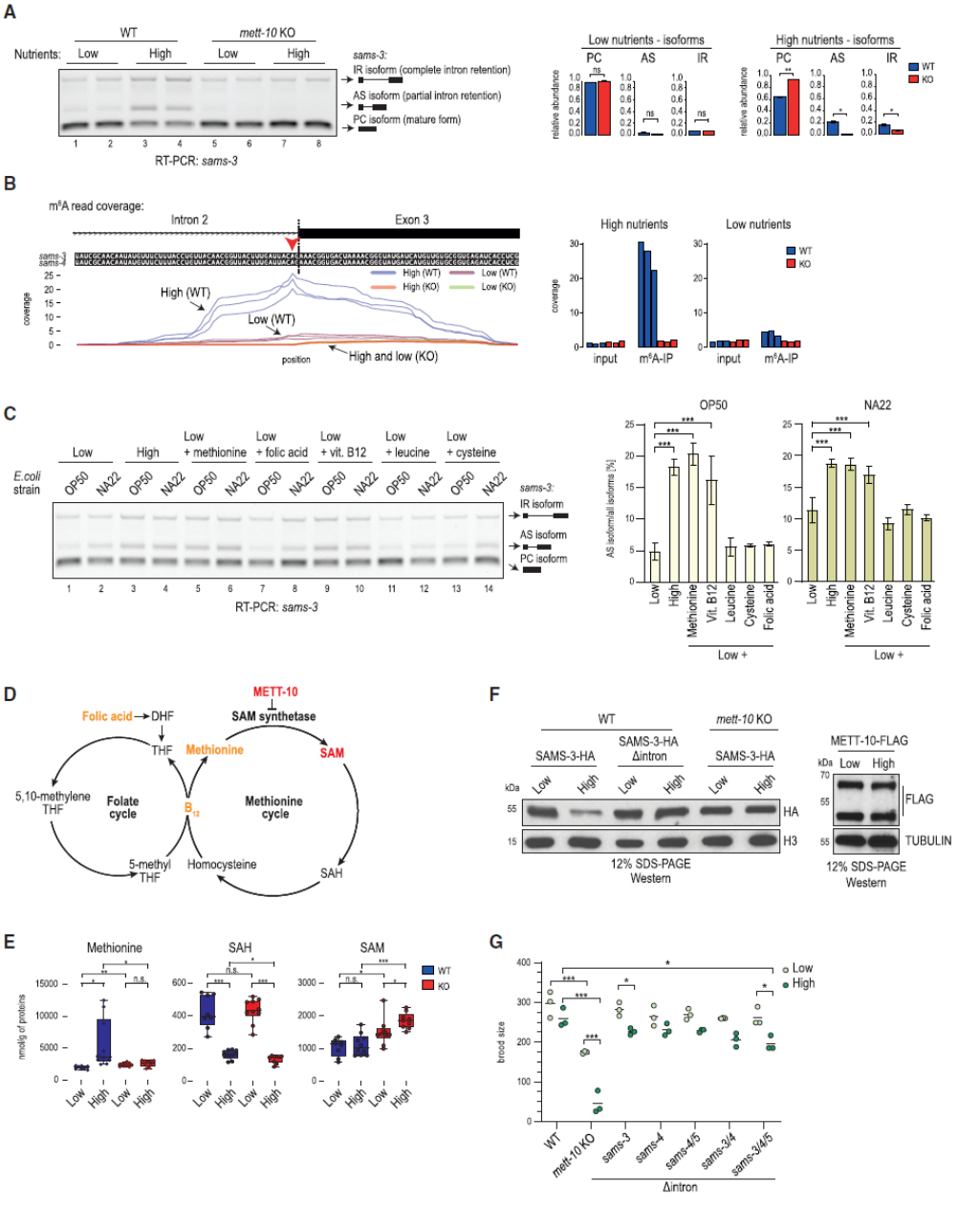
(C) *In vitro* methylation with recombinant METT-10 and the RNA substrates, based on the *sams-3* intron 2/exon 3 junction sequence, carrying mutations in the stem region.

(D) Splicing of WT and mutant (MUT) transgene reporter constructs injected into worm gonads. A MUT construct with triple mutations (AAC→CUU) within the methylation consensus motif (in the exon 3 part) increases 3′ splice site use, producing higher amounts of the PC isoform. Barplots depict the mean relative proportion of individual isoforms ± SD (n = 4). The p values were calculated using t tests. *p ≤ 0.05, ***p ≤ 0.001. See also Figure S3A for transgene analysis in the *mett-10* KO background.

(E) METT-10 consensus motif (red) and regions allowing secondary structure formation (yellow) are conserved in various worm species. Changes (green) in *C. japonica* are compensatory.

(F) Sequence alignment of the genomic region at the intron-exon boundary of the SAM synthetase gene from different organisms. The METT-10/METTL16 methylation consensus motif is highlighted (blue). Shown are *in vitro* methylations with ~30-nt RNAs corresponding to the intron-exon boundary sequence, carried out with recombinant human METTL16. The reaction products were resolved by PAGE and exposed to detect the radioactivity (^{14}C) signal. See also Figure S3B for the same reactions carried out with worm METT-10.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062



(legend on next page)

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

splice site m^{6}A methylation and splicing inhibition/alternative splicing of SAM synthetase pre-mRNA. This ensures optimal cellular SAM levels. Interestingly, it is known that a diet of the OP50 *E. coli* strain causes vitamin B12 deficiency in worms (Revtovich et al., 2019), probably explaining the reduced *sams-3* splicing inhibition compared with the NA22 strain (Figure 4C), and it also explains the reduced recycling of the by-product S-adenosylhomocysteine (SAH) via the methionine cycle under the nutrient-low diet (with the OP50 strain) condition (Figure 4E).

Validating the above model, metabolomics analyses (Figure 4E) show that, although WT worms are able to control SAM levels, the *mett-10* KO fails to do so. When grown on a nutrient-high diet that supplies an abundance of methionine, WT worms are able to maintain similar levels of SAM as under nutrient-low diet conditions (Figure 4E). Loss of *mett-10* upsets this homeostasis, resulting in elevated SAM concentrations under both diet conditions, with the levels being higher under the nutrient-high condition (Figure 4E). Thus, conditions that favor increased SAM production (such as nutrient-high diet) trigger m^{6}A methylation of the splice site in intron 2 of the *sams-3/4* pre-mRNA to inhibit production of the PC isoform version of SAM synthetase mRNA, regulating SAM biosynthesis. To directly verify protein levels of the enzyme during this regulation, we created a worm strain with SAMS-3 hemagglutinin (HA)-tagged at the endogenous locus and then derived a strain with intron 2 removed from the gene (STAR Methods; Table S5). Consistent with the RNA analyses, we observe a reduction in SAMS-3-HA protein levels under the nutrient-high diet condition (Figures 4F and S4D), and this depends on METT-10 and the presence of intron 2 in the *sams-3-HA* genomic locus (Figures 4F, S4D, and S4E). The level of the RNA methyltransferase (METT-10-FLAG; STAR Methods) does not change under the two dietary conditions (Figures 4F and S4F). Thus, m^{6}A-mediated reduction in protein levels of a key enzyme within the methionine cycle explains how WT worms cope with a diet that fuels this biosynthetic pathway to ensure SAM homeostasis.

Loss of *mett-10* results in a fertility defect phenotype (Dorsett et al., 2009), and here we examined the effect of diet. Compared with WT control animals, the *mett-10* KO has a reduced brood size with a nutrient-low diet, but this becomes worse with a nutrient-high diet, with very few progenies (Figure 4G). Interestingly, a triple-mutant worm strain (Table S5) lacking intron 2 in the three *sams* genes (*sams-3*$^{\Delta intron2}$, *sams-4*$^{\Delta intron2}$, and *sams-5*$^{\Delta intron2}$) also shows a small but significant reduction in brood size (Figure 4G). This shows that the ability to tune down SAM levels in response to a rich diet, using the m^{6}A-mediated splicing inhibition pathway we describe here, contributes to ensuring normal fertility in worms. Finally, the difference in the severity of the phenotypes of the *mett-10* KO and the triple mutant points to the existence of additional METT-10 targets that are required for fertility.

3′ splice site m^{6}A methylation inhibits splicing in mammalian cells

The above experiments show that worm METT-10 regulates splicing of *sams* pre-mRNA through m^{6}A methylation of a specific 3′ splice site. Because the basic mechanism of splicing is highly conserved from yeast to human (Fica and Nagai, 2017; Galej, 2018; Kastner et al., 2019), we wanted to find out whether the m^{6}A-mediated inhibitory pathway can be active in the mammalian system. To investigate this, we transfected the transgene reporter constructs based on worm *sams-3* into human HeLa cell cultures (Figure 5A). We already know that the 3′ splice site within worm *sams-3* RNA can be methylated by human METTL16 (Figures 3F and S5A). Strikingly, RT-PCR analysis of this reporter with the WT sequences revealed a splicing pattern similar to that seen when the same construct was expressed in worms, with 3′ splice site methylation reducing its use and promoting alternative splicing (AS isoform) via use of a cryptic upstream 3′ splice site (Figures 5A and S5B). As seen in worms, the transgene reporter with mutations (MUT, AAC→CUU) in the methylation consensus motif allows increased 3′ splice site use, reducing levels of the AS isoform (Figures 5A

Figure 4. Worms methylate the 3′ splice site of the SAM synthetase transcripts to downregulate their expression in response to a nutrient-high diet

(A) WT or *mett-10* KO worms were grown on plates that were high or low in nutrients. Splicing of intron 2 in the *sams-3* gene was monitored by RT-PCR analysis (biological duplicates are shown). Splicing of intron 2 in *sams-3* is different between WT and KO worms only under nutrient-high diet conditions. Barplots depict the mean relative proportion of individual isoforms ± SD (done in biological duplicates). The p values were calculated using t tests. *p ≤ 0.05, **p ≤ 0.01. See also Figure S4B for RNA-seq data.

(B) Mapping of m^{6}A-IP-seq reads (n = 3) from WT and *mett-10* KO worms fed on nutrient-high or nutrient-low plates. The m^{6}A coverage on the intron 2-exon 3 boundary of the *sams-3* gene is shown. The normalized coverages (rpm) from three biological replicates are plotted separately. See also Figures S4A–S4C. The barplot shows quantification (rpm) of the reads mapping to the *sams-3/4* genomic window shown. The read counts from three biological replicates are plotted separately.

(C) A nutrient-high diet inhibits splicing of *sams-3* intron 2 (RT-PCR analysis) in WT worms, as shown by an increased level of the AS isoform. Supplementing a nutrient-low diet with free methionine or vitamin B12 increases splicing inhibition. The barplots (mean ± SD) show quantification of the AS isoform band from three independent biological replicates. The nutrient-low and peptone-rich, nutrient-high media contained OP50 or the NA22 strain of *E. coli*. The p values were obtained by Tukey's HSD after ANOVA. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001.

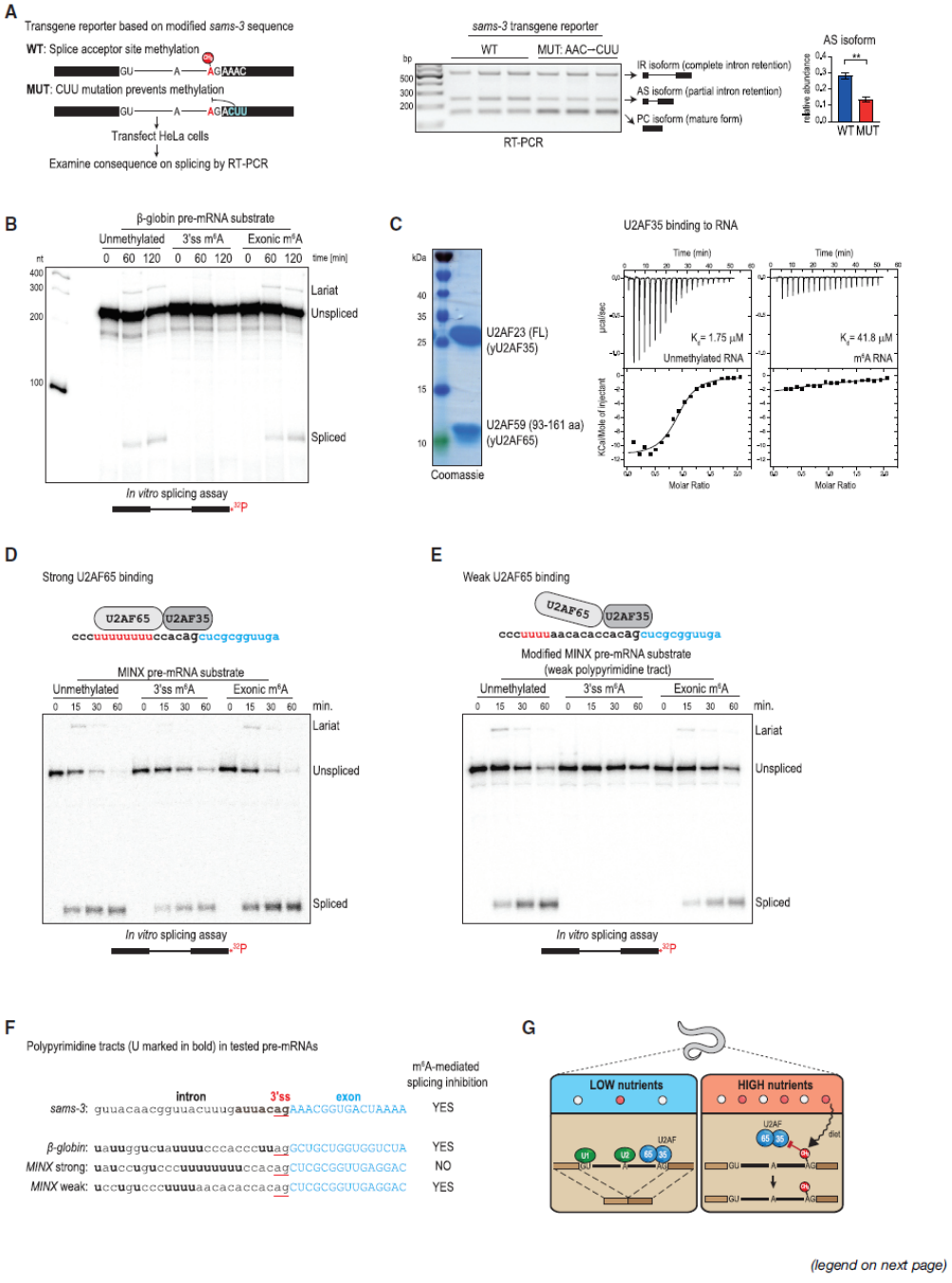
(D) A simplified scheme showing the methionine and folate cycles.

(E) Metabolomics analysis detecting the indicated metabolites. The p values were calculated using t tests and adjusted using Benjamini-Hochberg correction. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001.

(F) Western blot analysis of knockin worms expressing SAMS-3-HA or METT-10-FLAG proteins under different diet conditions. One of the worm lines has intron 2 deleted (Δintron) in the *sams-3-HA* gene locus. See also Figures S4D–S4F.

(G) Analysis of brood size in worms of the indicated genotypes (Table S5) when grown on nutrient-low or nutrient-high plates. Δintron, deletion of intron having the METT-10 methylated 3′ splice site. n = 3 independent experiments, each done in 2–5 technical replicates. The p values were calculated by two-way ANOVA followed by Tukey's HSD. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062



(legend on next page)

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS

and S5B). Thus, using this ectopic reporter system, we demonstrate that human METTL16 can catalyze 3′ splice site m^{6}A methylation, which leads to splicing modulation in human cells.

Next we wanted to know whether the observed splicing modulation is a direct consequence of the m^{6}A mark or whether the stem-loop structure that is required for recruitment of METTL16 plays any role. To demonstrate that the inhibitory effect is directly due to the presence of m^{6}A, we artificially introduced an m^{6}A at the 3′ splice site (by splint ligation; STAR Methods) of the unrelated human β-globin pre-mRNA and carried out *in vitro* splicing assays (Krainer et al., 1984). To this end, we prepared ^{32}P-labeled splicing substrates and incubated them with human HeLa S3 extracts (Figure 5B). Splicing takes place via two transesterification reactions (Fica and Nagai, 2017; Shi, 2017; Will and Lührmann, 2011). In step 1, the free 5′ exon and the intron lariat-3′ exon intermediate are produced. In step 2, exon ligation joins the 5′ exon with the 3′ exon, releasing the branched lariat. Splicing of the unmethylated substrate proceeded normally, as expected (Padgett et al., 1984; Ruskin et al., 1984), with production of the lariat intermediate and the mature spliced product observed (Figures 5B and S5C). However, splicing of the substrate with the m^{6}A modification at the 3′ splice site was blocked completely because the lariat and the mature product were absent (Figures 5B and S5C). Placing the m^{6}A mark in the exonic part of the substrate did not hinder splicing (Figures 5B and S5C), demonstrating the specificity of the 3′ splice site inhibitory mechanism. Thus, we conclude that the human splicing machinery is also sensitive to the presence of m^{6}A at the 3′ splice site, and this directly inhibits the first step of the splicing reaction.

m^{6}A methylation prevents splice site recognition by the essential splicing factor U2AF35

Recognition by splicing factors of the key *cis* elements within the pre-mRNA is critical for initiation of splicing in metazoans. The 5′ splice site is recognized by the U1 snRNP, the branchpoint sequence (BPS) by the mammalian branchpoint binding protein (mBBP)/SF1, and the 3′ splice site is bound by the U2 auxiliary factor (U2AF). mBBP/SF1 and U2AF then promote recruitment of the U2 snRNP, which pairs with the branch-site sequence. U2AF is a heterodimer composed of the U2AF35 and U2AF65 subunits (Zamore and Green, 1989). While U2AF65 recognizes the polypyrimidine tract that precedes the AG dinucleotide at the intron-exon junction (Sickmier et al., 2006; Zamore et al., 1992), U2AF35 has been shown to directly contact the 3′ splice site AG dinucleotide (Merendino et al., 1999; Soares et al., 2006; Wu et al., 1999; Zorio and Blumenthal, 1999a; Zuo and Maniatis, 1996). U2AF35 is highly conserved from fission yeast to human and essential for splicing *in vivo* in worms (Zorio and Blumenthal, 1999b) and flies (Rudner et al., 1996).

This prompted us to examine whether 3′ splice site methylation can hinder U2AF35 binding. Our attempts to express full-length human or worm U2AF35 alone in a recombinant form were unsuccessful. However, we could stabilize fission yeast (*Schizosaccharomyces pombe*) full-length U2AF35 by expressing it in complex (Yoshida et al., 2015) with a minimal fragment of U2AF65 (the U2AF35-interacting region) lacking the RNA binding domains (Zamore et al., 1992; Figure 5C; STAR Methods). U2AF35 with its two zinc fingers (Figure S6A) is the only component in this complex with the ability to bind RNA, hence, hereafter, this preparation will be referred to as U2AF35. We used a short RNA fragment mimicking the 3′ splice site (AG) to test interactions with U2AF35. Isothermal calorimetry (ITC) experiments revealed that, although U2AF35 strongly (K_D = 1.75 μM) interacts with the unmethylated RNA, the presence of 3′ splice site m^{6}A decreases the affinity by an order of magnitude (K_D = 41.8 μM) (Figures 5C and S6B–S6D). Thus, the 3′ splice site m^{6}A inhibits splicing by physically hindering its recognition by the essential splicing factor U2AF35.

Splice site m^{6}A methylation inhibits splicing of AG-dependent introns

Of the different splicing signals within the intron, the polypyrimidine tract is the most variable. Its composition, measured by the number of uridines in the tract (Singh et al., 1995; Zamore et al., 1992), defines the strength of the 3′ splice site (Moore, 2000; Reed, 1989). *In vitro* splicing of an intron with a strong polypyrimidine tract (AG-independent introns) requires only U2AF65 (Valcárcel et al., 1996; Zamore et al., 1992), whereas that with

Figure 5. A 3′ splice site m^{6}A inhibits splicing in human cells and blocks its recognition by U2AF35

(A) Worm transgene reporter constructs based on *sams-3* were transfected into human HeLa cells, and splicing patterns were analyzed by RT-PCR. A MUT construct with triple mutations (AAC→CUU) within the methylation consensus motif (in the exon 3 part) increases 3′ splice site use, producing lower amounts of the AS isoform. The barplot depicts the mean relative proportion of the AS isoform to the sum of all isoforms ± SD (n = 3). The p value was calculated using a t test. **p ≤ 0.01. See also Figure S5B for all replicates.

(B) *In vitro* splicing assay with HeLa S3 nuclear extracts. The human β-globin pre-mRNA substrate is spliced correctly, whereas the same substrate with an m^{6}A methylated 3′ splice site (ss) remains unspliced. The presence of the methyl mark on the exonic part does not inhibit splicing. A band corresponding to the lariat intermediate is visible in lanes where the substrate is spliced correctly. Substrates were incubated for different durations (time in minutes) with the extracts. See also Figure S5C.

(C) ITC experiments reveal that the full-length (FL) yeast U2AF35 (stabilized with a fragment of yeast U2AF65; STAR Methods) strongly binds an unmethylated RNA substrate mimicking the 3′ ss (AG), whereas the presence of an m^{6}A mark decreases affinity. The quality of the recombinant protein used is shown. See also Figure S6.

(D) Splicing assays with the MINX pre-mRNA substrate. 3′ ss m^{6}A does not inhibit splicing of this substrate, which has a strong polypyrimidine tract.

(E) Mutations that weaken the polypyrimidine tract in MINX pre-mRNA make it sensitive to inhibition by 3′ ss m^{6}A. The presence of the methyl mark on the exonic part does not inhibit splicing.

(F) Sequence of the 3′ end of the intron in the splicing substrates, showing the polypyrimidine tract (bold) and the 3′ ss. A similar region from worm *sams-3* pre-mRNA is also shown, with the consensus ss motif shown (bold).

(G) Model showing how 3′ ss m^{6}A methylation under nutrient-high conditions prevents binding of U2AF35, leading to inhibition of splicing of *sams* pre-mRNA in worms.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

a weak polypyrimidine tract (AG-dependent introns) additionally requires U2AF35, which recognizes the AG dinucleotide (Wu et al., 1999). Thus, although the conserved AG dinucleotide at the 3′ splice site is only required for the second step of splicing during exon ligation, AG-dependent introns require its recognition by U2AF35 early during spliceosome assembly and for the first step of splicing (Reed, 1989; Wu et al., 1999).

To determine whether splicing inhibition by 3′ splice site m^{6}A depends on the type of intron involved, we experimented with the MINX (an adenovirus major late pre-mRNA derivative) splicing substrate (Figures 5D and 5E). Compared with the β-globin pre-mRNA substrate, the MINX substrate has a strong polypyrimidine tract with a run of eight uridines (U_8), identifying the intron as AG independent (Figure 5F). When incubated with HeLa S3 extracts, the unmethylated MINX substrate is spliced, with the lariat intermediate and spliced product visible (Figure 5D). Interestingly, the MINX substrate with an m^{6}A-methylated 3′ splice site is also spliced, albeit with slightly lower efficiency (Figure 5D). This is contrary to the observation for the β-globin pre-mRNA substrate, where the 3′ splice site m^{6}A completely inhibits splicing (Figure 5B). This suggests that the inhibitory effect of 3′ splice site m^{6}A is dependent on the type of intron being regulated. To verify whether this is due to the presence of a strong polypyrimidine tract (U_8), we introduced mutations to convert the MINX construct into a substrate with only four uridines (U_4) (Figure 5F). Strikingly, splicing of such a MINX pre-mRNA substrate with a weakened polypyrimidine tract (effectively making it an AG-dependent intron) is abolished completely in the presence of a 3′ splice site m^{6}A (Figure 5E). An exonic methylation does not affect splicing of either substrate. This indicates that AG-dependent introns with a weakened polypyrimidine tract are sensitive to a 3′ splice site m^{6}A because they require recognition by U2AF35 of the AG dinucleotide for efficient U2AF recruitment.

In this context, it is worth mentioning that introns in *C. elegans* lack the polypyrimidine tract consensus sequence as in other metazoans but instead have a conserved consensus sequence, U_4CAG (Figure 5F), at the 3′ end (Blumenthal and Steward, 1997). The U_4C sequence in this consensus sequence is bound by worm U2AF65, but this association is enhanced by simultaneous binding of worm U2AF35 to the AG dinucleotide (Zorio and Blumenthal, 1999a). We show that splicing of AG-dependent introns, which rely on U2AF35 binding to the AG dinucleotide to recruit the U2AF complex, can be regulated by m^{6}A methylation of the 3′ splice site (Figure 5G).

Search for 3′ splice sites potentially regulated by mammalian METTL16

Although we demonstrated that splicing inhibition by 3′ splice site m^{6}A methylation is conserved in mammals, mammalian METTL16 was not shown to methylate 3′ splice sites of mammalian pre-mRNAs. Mammalian METTL16 regulates its conserved SAM synthetase *MAT2A* RNA target by promoting splicing via its non-catalytic C-terminal VCRs (Pendleton et al., 2017). Loss of *Mettl16* causes pre-implantation embryonic lethality in mice (Mendel et al., 2018). To examine the *in vivo* relevance of its catalytic activity, we created a knockin mouse mutant carrying mutations in the catalytic motif (Figures S7A and S7B; STAR Methods). Although the heterozygous mutants are viable and fertile, homozygous catalytic-dead *Mettl16* mutants are never recovered in the born litters, indicating developmental lethality (Figure 6A). Similarly, mutations designed to cause loss of RNA-binding activity also result in lethality (Figure 6A). METTL16 has a tissue-specific expression pattern in adult mice, with strong enrichment in the gonads (Figure 6B). To probe its relevance for fertility, we engineered conditional deletion of *Mettl16* in the mouse germline (Figure S7C; STAR Methods). Such conditional KO (cKO) males are infertile, as evidenced by atrophied testes (Figure 6C) and arrested germ cell development (Figure 6D). Taken together, our genetic analyses reveal an essential role of the catalytic activity of METTL16 during mouse development and show that the protein is also relevant outside of the embryonic stages.

Next we identified putative mammalian targets for METTL16-mediated 3′ splice site m^{6}A methylation (STAR Methods). Briefly, these sites overlap one of the METTL16/METT-10 methylation motifs (UAC**m^{6}A**GAGA or UAC**m^{6}A**GAAA) and are present within a stem-loop structure (Figure 6E). Direct testing of the top 10 such sequences with recombinant human METTL16 shows that several of these are methylated efficiently *in vitro* (Figure 6F). To examine whether any of these putative targets are regulated differentially in the absence of *Mettl16*, we used single-embryo RNA-seq datasets prepared from *Mettl 16* KO embryonic day 2.5 (E2.5) morulae and E3.5 blastocysts (Mendel et al., 2018). This identified *Sorbs1* and *Lrpprc* as two transcripts that have increased use of the 3′ splice site in the *Mettl16* KO (Figure 6G). Furthermore, the target splice sites in these transcripts can be methylated *in vitro* by recombinant METTL16 (Figure 6G). Although our computational and biochemical analyses reveal the existence of putative 3′ splice site targets for mammalian METTL16, it remains to be seen whether they are indeed regulated by METTL16 *in vivo*.

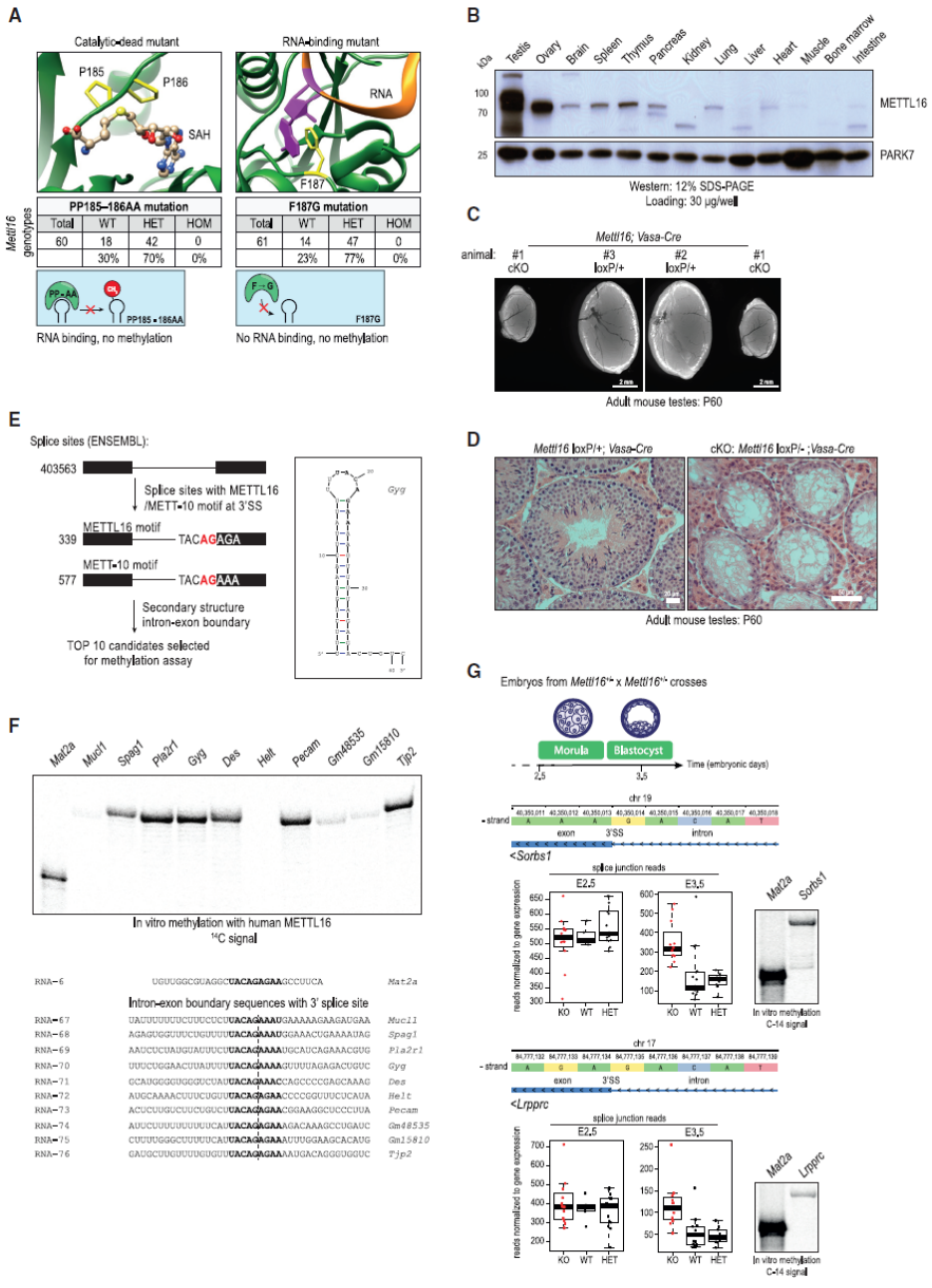
DISCUSSION

SAM is the major methyl donor for methylation reactions in the cell (Cantoni, 1975). Production of SAM from methionine and ATP via the methionine cycle is carried out by methionine adenosyltransferase (MAT) or SAM synthetase, which is conserved from prokaryotes to humans. One conserved principle for regulation of SAM synthetase gene expression is use of RNA structures. Prokaryotes use complex RNA structures, called riboswitches, present in the 5′ leader sequence of SAM synthetase mRNA for feedback regulation by inhibiting translation or attenuating transcription (Batey, 2011; Mandal and Breaker, 2004). Binding of SAM alters the RNA structure, leading to gene repression; for example, by occluding key features like the Shine-Dalgarno sequence required for translation initiation (Breaker, 2018). Even in eukaryotes, fission yeast SAM synthetase *sam1* mRNA has a tertiary structure feature in the 5′ UTR, which, upon SAM binding, undergoes structural transition to regulate translation (Zhang et al., 2020).

Mammals use a different strategy to regulate *MAT2A* SAM synthetase expression that does not involve direct binding of SAM. Six hairpin structures in the 3′ UTR of the *MAT2A* pre-mRNA bind the m^{6}A writer METTL16, which uses its non-catalytic C-terminal VCRs to enhance splicing of a frequently retained

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS



(legend on next page)

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

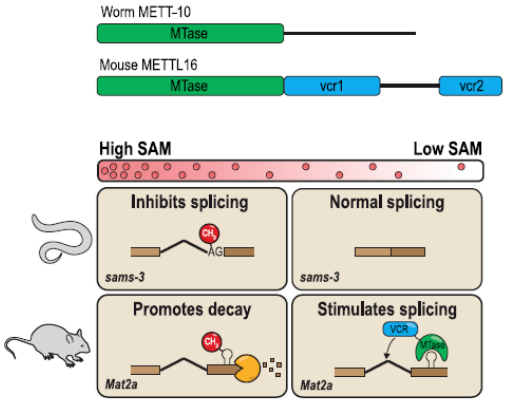


Figure 7. Conserved targets of METTL16-mediated m^{6}A methylation activity and specialization of the C-terminal VCR in vertebrates

SAM levels are highly regulated *in vivo*, and this is achieved by splicing regulation of the SAM synthetase RNA (*sams-3* or *MAT2A*). Under high-SAM conditions, METT-10 m^{6}A methylates a 3′ ss in *sams-3* pre-mRNA to directly inhibit splicing, whereas methylation in the 3′ UTR of *MAT2A* by mammalian METTL16 leads to intron retention/decay of the RNA. Under low-SAM conditions, mammalian METTL16 binds hairpins in the 3′ UTR of *MAT2A* and uses its C-terminal VCR to stimulate splicing of the terminal intron, whereas in nematodes, absence of ss methylation allows normal splicing to proceed. The different mechanisms also highlight the different approaches to regulation of SAM levels: nematode METT-10 turns off SAM production, whereas mammalian METTL16 actively turns on SAM production.

terminal intron (Pendleton et al., 2017). The role of SAM in this process is as a molecular regulator of METTL16's dwell time at the 3′ UTR; low levels increase dwell time, whereas high levels, which lead to m^{6}A methylation of the hairpins, rapidly evict the protein from the pre-mRNA (Figure 7). The VCR has also been shown to facilitate binding to the U6 snRNA (Aoyama et al., 2020). In this study, we show that the worm METTL16 ortholog METT-10 inhibits SAM synthetase pre-mRNA splicing via 3′ splice site m^{6}A methylation in response to a rich diet. Identification of the precise 3′ splice site is by its presence within a stem-loop structure. Conservation of these sequence elements within SAM synthetase genes implies that this type of regulation might be common in invertebrates (Figures 3E and 3F).

Our findings also highlight the different strategies used to regulate SAM synthetase expression. Prokaryotes and invertebrates negatively regulate SAM production in response to high intracellular SAM levels. On the other hand, mammalian systems have opted for a mechanism that allows them to increase SAM production, probably to suit the requirements of early embryonic development, where optimal SAM levels are critical for survival of embryonic stem cells (ESCs) (Shiraki et al., 2014) and development of embryos (Sun et al., 2019). This critical role explains why mice lacking METTL16 die early during embryogenesis (Mendel et al., 2018), whereas worms lacking METT-10 are viable (Dorsett et al., 2009) because they just need to cope with the aberrantly high levels of SAM after a rich diet (Figure 4E). Nevertheless, the observed phenotypes of the *mett-10* KO and the *sams* triple mutant lacking intron 2 (Figure 4G) demonstrate that the ability to dial down SAM production in response to a rich diet is important for ensuring normal fertility in worms (Figure 4G).

Previous studies have linked m^{6}A methylation to splicing regulation, and they document splicing changes in the absence of an m^{6}A writer (Haussmann et al., 2016; Lence et al., 2016), reader (Kasowitz et al., 2018; Xiao et al., 2016; Zhou et al., 2019), or eraser (Bartosovic et al., 2017; Zhao et al., 2014). Our study identifies a direct role of the modification in interfering with splicing via precise methylation of a key sequence feature used by the splicing machinery. This mechanism is direct because it repels an essential splicing factor, U2AF35, leading to an early spliceosome assembly defect.

Limitations of study

The strong fertility defect seen in the *mett-10* KO worms (Dorsett et al., 2009) sharply contrasts the relatively mild phenotype in the

Figure 6. RNA m^{6}A methylation activity of mouse METTL16 is essential for development and has the potential to methylate the 3′ ss of target RNAs

(A) Analysis of knockin (KI) mouse mutants for *Mettl16*, with mutations abolishing catalytic activity or RNA binding. A structural model of human METTL16 (PDB: 6GFK) shows the two prolines (PP185–PP186) of the NPPF catalytic motif close to the bound SAH molecule, and a model of human METTL16 in complex with bound *MAT2A* hairpin RNA (PDB: 6DU4) shows the F187 that flips in to interact with the target adenosine upon substrate RNA binding. Introduced mutations are indicated. See also Figures S7A and S7B. Shown are genotypes of animals recovered in born litters from crosses between heterozygous *Mettl16* knockin (KI) parents (*Mettl16*$^{KI/+}$). Homozygous KI mutants were not obtained for either mutation, indicating lethality. HET, heterozygous; HOM, homozygous KI.

(B) Multiple-tissue western blot showing tissue-specific expression of mouse METTL16. A loading control is provided by detection of PARK7.

(C) Representative picture of atrophied testes from a mouse with conditional (*Vasa*-Cre) deletion of *Mettl16* in the germline. Such animals are infertile. See also Figure S7C.

(D) Histology of adult mouse testes showing complete absence of germ cells in seminiferous tubules from mice with conditional (*Vasa*-Cre) deletion of *Mettl16* in the germline. cKO, conditional KO. The control HET testis shows all different stages of germ cells, including post-meiotic round spermatids and elongated spermatids.

(E) Scheme showing identification of putative targets of mammalian METTL16 on 3′ ss. The total numbers of 3′ ss checked and those recovered with the METTL16/METT-10 motifs are given. The predicted secondary structure of one such RNA (intron-exon boundary with 3′ ss) is shown.

(F) *In vitro* methylation assays with recombinant human METTL16 and the indicated RNAs. The RNA sequence for mouse *Mat2a* is from the 3′ UTR, whereas for other mouse genes it spans the intron-exon boundary (sequences are shown below). Reactions were resolved by PAGE, and the radioactivity (^{14}C) signal was detected.

(G) Two transcripts that show increased splice junction reads specifically in *Mettl16* KO embryos (morulae at E2.5 or blastocysts at E3.5), indicating increased use of that ss in the absence of METTL16. Genomic coordinates of the 3′ ss and the underlying sequence on the Crick strand are shown. *In vitro* methylation assays with RNAs spanning the intron-exon boundary show methylation of the 3′ ss by mammalian METTL16.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS

triple mutant, where intron 2 is deleted in the *sams-3/4/5* genes (Figure 4G). This points to the existence of additional targets for METT-10 that may contribute to fertility. Alternatively, although we did not observe any global splicing differences (Figure S1J), it is possible that loss of U6 snRNA m^{6}A methylation in the *mett-10* KO may affect splicing of specific genes. We show that the 3′ splice site m^{6}A inhibits splicing in *C. elegans*, and it is also active in human cell cultures and in *in vitro* HeLa splicing extracts, but there is no evidence of its actual use in splicing regulation in mammals. We identified several putative 3′ splice sites that are methylated by METTL16 *in vitro* or have, as an isolated RNA sequence, all requirements for methylation. It is possible that many of these sites are never methylated *in vivo* because transcription kinetics (Herzel et al., 2017) may affect the ability of the region to fold into the required stem-loop structure for methylation. Even for the two transcripts (*Sorbs1* and *Lrpprc*) that show altered splicing patterns in the *Mettl16* KO embryos, it is not clear whether this is actually due to m^{6}A methylation of the specific 3′ splice sites. A search for such mammalian targets will have to involve analysis of specific cell types or tissues under specific developmental or environmental conditions. Nevertheless, given the conservation of the mechanisms involved, our work identifies 3′ splice site methylation as an ancient strategy for splicing control.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animal Work
  - Generation of catalytic-dead and RNA-binding mutant Mettl16 mouse lines
  - Conditional Mettl16 knockout mouse generation
  - Genotyping
  - Nematode strains and growth conditions
  - Generation of *C. elegans* strains
  - Generation of *C. elegans* lines expressing transgene reporter constructs as transgenes
  - Collection of *C. elegans* for the RNA isolation
  - Cell lines
- METHOD DETAILS
  - Clones and constructs
  - Constructs for bacterial protein expression
  - Constructs for insect cell expression
  - Constructs for expression of *sams-3* transgene reporters in transgenic worms
  - Constructs for expression of transgene reporters in human cells
  - Antibodies
  - Recombinant protein production
  - Expression and purification of yeast U2AF35 protein
  - ITC experiment with yeast U2AF35
  - Worm total RNA purification
  - Poly(A)$^+$ RNA purification
  - Detection of m^{6}A methylation using SCARLET
  - Quantification of *sams-3* splicing by RT-PCR
  - Quantification of RNA modifications using LC-MS/MS
  - Metabolomics analyses of worm lysates
  - *In vitro* RNA methylation assay with human METTL16 and worm METT-10
  - Preparation of RNA substrates for *in vitro* splicing assay
  - Preparation of nuclear extracts
  - *In vitro* splicing reaction
  - Histology of mouse tissue sections
  - Protein extraction from mouse tissues
  - Protein extraction from worms
  - Western Blot
  - Preparation of RNA libraries
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Analysis of m^{6}A-IP-seq to compare m^{6}A levels in **mouse, worm and insects**
  - Analysis of m^{6}A-IP-seq comparing m^{6}A levels in WT and *mett-10* KO worms
  - Search for mouse genes with METTL16 methylation motif at their 3′ splice site

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2021.03.062.

ACKNOWLEDGMENTS

We thank the Caenorhabditis Genetics Center (CGC), which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440), for some worm strains. We thank Markus Stoffel for transgenic mouse lines. We thank Sebastian Fica for providing β-globin and MINX splicing constructs, Wojtek Galej for discussions, and Pascal Gos and Marina Berti for assistance with experiments. We thank the following facilities: the iGE3 Genomics Platform, University of Geneva; the Metabolomics Platform, University of Lausanne; and the EMBL Genomics core facility. M.M. is grateful for a Boehringer Ingelheim Fonds PhD fellowship and thanks the Novartis Foundation for Medical-Biological Research (19B138) for salary support. This work was supported by grants to R.S.P. from the Swiss National Science Foundation (ERC Transfer Grant CRETP3_166923, Project Grant 310030B_185386, and Sinergia Grant CRSII5_183524 and funding from the NCCR RNA & Disease 51NF40_182880). Work in the Pillai lab is supported by the Republic and Canton of Geneva.

AUTHOR CONTRIBUTIONS

M.M. performed all cell culture and biochemical experiments and analyses of transgene reporter assays from worms and human cells, created and analyzed mouse mutants, and coordinated worm experiments with K.D. K.D. prepared transgenic and knockin worm lines and conducted worm experiments with help from J.M.W. and under supervision of F.A.S. R.R.P. made m^{6}A-IP-seq libraries. K.-M.C. made recombinant proteins and carried out ITC measurements. C.B.V. conducted RNA mass spectrometry. D.H. conducted all computational analyses. R.S.P. and D.H. supervised the study and wrote the manuscript with input from everyone.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

Received: June 7, 2020
Revised: February 16, 2021
Accepted: March 30, 2021
Published: April 29, 2021

REFERENCES

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J. Mol. Biol. *215*, 403–410.

Aoyama, T., Yamashita, S., and Tomita, K. (2020). Mechanistic insights into m6A modification of U6 snRNA by human METTL16. Nucleic Acids Res. *48*, 5157–5168.

Arribere, J.A., Bell, R.T., Fu, B.X., Artiles, K.L., Hartman, P.S., and Fire, A.Z. (2014). Efficient marker-free recovery of custom genetic modifications with CRISPR/Cas9 in Caenorhabditis elegans. Genetics *198*, 837–846.

Bailey, T.L., and Elkan, C. (1994). Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. *2*, 28–36.

Bartosovic, M., Molares, H.C., Gregorova, P., Hrossova, D., Kudla, G., and Vanacova, S. (2017). N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3′-end processing. Nucleic Acids Res. *45*, 11356–11370.

Batey, R.T. (2011). Recognition of S-adenosylmethionine by riboswitches. Wiley Interdiscip. Rev. RNA *2*, 299–311.

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.

Bieniossek, C., Imasaki, T., Takagi, Y., and Berger, I. (2012). MultiBac: expanding the research toolbox for multiprotein complexes. Trends Biochem. Sci. *37*, 49–57.

Blumenthal, T., and Steward, K. (1997). C. elegans II (Cold Spring Harbor Laboratory Press).

Breaker, R.R. (2018). Riboswitches and translation control. Cold Spring Harb. Perspect. Biol. *10*, a032797.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. Genetics *77*, 71–94.

Brown, J.A., Kinzig, C.G., DeGregorio, S.J., and Steitz, J.A. (2016). Methyltransferase-like protein 16 binds the 3′-terminal triple helix of MALAT1 long noncoding RNA. Proc. Natl. Acad. Sci. USA *113*, 14013–14018.

Cantoni, G.L. (1975). Biological methylation: selected aspects. Annu. Rev. Biochem. *44*, 435–451.

Chevalier, C., Kieser, S., Çolakoğlu, M., Hadadi, N., Brun, J., Rigo, D., Suárez-Zamorano, N., Spiljar, M., Fabbiano, S., and Busse, B. (2020). Warmth prevents bone loss through the gut microbiota. Cell Metab. *32*, 575–590.e7.

Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc. Natl. Acad. Sci. USA *71*, 3971–3975.

Desrosiers, R.C., Friderici, K.H., and Rottman, F.M. (1975). Characterization of Novikoff hepatoma mRNA methylation and heterogeneity in the methylated 5′ terminus. Biochemistry *14*, 4367–4374.

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.

Dorsett, M., Westlund, B., and Schedl, T. (2009). METT-10, a putative methyltransferase, inhibits germ cell proliferative fate in Caenorhabditis elegans. Genetics *183*, 233–247.

Doxtader, K.A., Wang, P., Scarborough, A.M., Seo, D., Conrad, N.K., and Nam, Y. (2018). Structural Basis for Regulation of METTL16, an S-Adenosylmethionine Homeostasis Factor. Mol. Cell *71*, 1001–1011.e4.

Evans, T.C. (2006). Transformation and microinjection. WormBook, ed. (The C. elegans Research Community), https://doi.org/10.1895/wormbook.1.108.1, http://www.wormbook.org.

Félix, M.-A., and Duveau, F. (2012). Population dynamics and habitat sharing of natural populations of Caenorhabditis elegans and C. briggsae. BMC Biol. *10*, 59.

Fica, S.M., and Nagai, K. (2017). Cryo-electron microscopy snapshots of the spliceosome: structural insights into a dynamic ribonucleoprotein machine. Nat. Struct. Mol. Biol. *24*, 791–799.

Fu, Y., Dominissini, D., Rechavi, G., and He, C. (2014). Gene expression regulation mediated through reversible m^{6}A RNA methylation. Nat. Rev. Genet. *15*, 293–306.

Galej, W.P. (2018). Structural studies of the spliceosome: past, present and future perspectives. Biochem. Soc. Trans. *46*, 1407–1422.

Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., et al. (2015). Stem cells. m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science *347*, 1002–1006.

Haussmann, I.U., Bodi, Z., Sanchez-Moran, E., Mongan, N.P., Archer, N., Fray, R.G., and Soller, M. (2016). m^{6}A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. Nature *540*, 301–304.

Herzel, L., Ottoz, D.S.M., Alpert, T., and Neugebauer, K.M. (2017). Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. Nat. Rev. Mol. Cell Biol. *18*, 637–650.

Hsu, P.J., Zhu, Y., Ma, H., Guo, Y., Shi, X., Liu, Y., Qi, M., Lu, Z., Shi, H., Wang, J., et al. (2017). Ythdc2 is an N^6-methyladenosine binding protein that regulates mammalian spermatogenesis. Cell Res. *27*, 1115–1127.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. Nat. Methods *12*, 115–121.

Ivanova, I., Much, C., Di Giacomo, M., Azzi, C., Morgan, M., Moreira, P.N., Monahan, J., Carrieri, C., Enright, A.J., and O'Carroll, D. (2017). The RNA m^{6}A Reader YTHDF2 Is Essential for the Post-transcriptional Regulation of the Maternal Transcriptome and Oocyte Competence. Mol. Cell *67*, 1059–1067.e4.

Jain, D., Puno, M.R., Meydan, C., Lailler, N., Mason, C.E., Lima, C.D., Anderson, K.V., and Keeney, S. (2018). *ketu* mutant mice uncover an essential meiotic function for the ancient RNA helicase YTHDC2. eLife *7*, e30919.

Jalili, V., Matteucci, M., Masseroli, M., and Morelli, M.J. (2018). Using combined evidence from replicates to evaluate ChIP-seq peaks. Bioinformatics *34*, 2338.

Kan, L., Grozhik, A.V., Vedanayagam, J., Patil, D.P., Pang, N., Lim, K.-S., Huang, Y.-C., Joseph, B., Lin, C.-J., Despic, V., et al. (2017). The m^{6}A pathway facilitates sex determination in Drosophila. Nat. Commun. *8*, 15737.

Kasowitz, S.D., Ma, J., Anderson, S.J., Leu, N.A., Xu, Y., Gregory, B.D., Schultz, R.M., and Wang, P.J. (2018). Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. PLoS Genet. *14*, e1007412.

Kastner, B., Will, C.L., Stark, H., and Lührmann, R. (2019). Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. Cold Spring Harb. Perspect. Biol. *11*, a032417.

Ke, S., Alemu, E.A., Mertens, C., Gantman, E.C., Fak, J.J., Mele, A., Haripal, B., Zucker-Scharff, I., Moore, M.J., Park, C.Y., et al. (2015). A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes Dev. *29*, 2037–2053.

Kielkopf, C.L., Rodionova, N.A., Green, M.R., and Burley, S.K. (2001). A novel peptide recognition mode revealed by the X-ray structure of a core U2AF35/U2AF65 heterodimer. Cell *106*, 595–605.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

Krainer, A.R., Maniatis, T., Ruskin, B., and Green, M.R. (1984). Normal and mutant human beta-globin pre-mRNAs are faithfully and efficiently spliced in vitro. Cell *36*, 993–1005.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

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 m^{6}A reader proteins. Genes Dev. *34*, 1373–1391.

Lee, K.A., Bindereif, A., and Green, M.R. (1988). A small-scale procedure for preparation of nuclear extracts that support efficient transcription and pre-mRNA splicing. Gene Anal. Tech. *5*, 22–31.

Lence, T., Akhtar, J., Bayer, M., Schmid, K., Spindler, L., Ho, C.H., Kreim, N., Andrade-Navarro, M.A., Poeck, B., Helm, M., and Roignant, J.Y. (2016). m^{6}A modulates neuronal functions and sex determination in Drosophila. Nature *540*, 242–247.

Li, F., Zhao, D., Wu, J., and Shi, Y. (2014). Structure of the YTH domain of human YTHDF2 in complex with an m(6)A mononucleotide reveals an aromatic cage for m(6)A recognition. Cell Res. *24*, 1490–1492.

Liu, N., Parisien, M., Dai, Q., Zheng, G., He, C., and Pan, T. (2013). Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA *19*, 1848–1856.

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.

Lorenz, R., Bernhart, S.H., Höner Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol. Biol. *6*, 26.

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.

Mandal, M., and Breaker, R.R. (2004). Gene regulation by riboswitches. Nat. Rev. Mol. Cell Biol. *5*, 451–463.

Mayeda, A., and Krainer, A.R. (1999). Mammalian *in vitro* splicing assays. In RNA-Protein Interaction Protocols, S.R. Haynes, ed. (Springer), pp. 315–321.

Mendel, M., Chen, K.M., Homolka, D., Gos, P., Pandey, R.R., McCarthy, A.A., and Pillai, R.S. (2018). Methylation of Structured RNA by the m^{6}A Writer METTL16 Is Essential for Mouse Embryonic Development. Mol. Cell *71*, 986–1000.e11.

Merendino, L., Guth, S., Bilbao, D., Martínez, C., and Valcárcel, J. (1999). Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3′ splice site AG. Nature *402*, 838–841.

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.

Mon, H., Kobayashi, I., Ohkubo, S., Tomita, S., Lee, J., Sezutsu, H., Tamura, T., and Kusakabe, T. (2012). Effective RNA interference in cultured silkworm cells mediated by overexpression of Caenorhabditis elegans SID-1. RNA Biology 9, 40–46. https://doi.org/10.4161/rna.9.1.18084.

Moore, M.J. (2000). Intron recognition comes of AGe. Nat. Struct. Biol. *7*, 14–16.

Moore, M.J., and Sharp, P.A. (1992). Site-specific modification of pre-mRNA: the 2′-hydroxyl groups at the splice sites. Science *256*, 992–997.

Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F., and Sharp, P.A. (1984). Lariat RNA's as intermediates and products in the splicing of messenger RNA precursors. Science *225*, 898–903.

Pandey, R.R., Delfino, E., Homolka, D., Roithova, A., Chen, K.-M., Li, L., Franco, G., Vågbø, C.B., Taillebourg, E., Fauvarque, M.-O., and Pillai, R.S. (2020). The Mammalian Cap-Specific m^{6}Am RNA Methyltransferase PCIF1 Regulates Transcript Levels in Mouse Tissues. Cell Rep. *32*, 108038.

Patil, D.P., Pickering, B.F., and Jaffrey, S.R. (2018). Reading m^{6}A in the Transcriptome: m^{6}A-Binding Proteins. Trends Cell Biol. *28*, 113–127.

Patro, R., Duggal, G., Love, M.I., Irizarry, R.A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nat. Methods *14*, 417–419.

Pendleton, K.E., Chen, B., Liu, K., Hunter, O.V., Xie, Y., Tu, B.P., and Conrad, N.K. (2017). The U6 snRNA m^{6}A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. Cell *169*, 824–835.e14.

R Core Team (2017). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing).

Reed, R. (1989). The organization of 3′ splice-site sequences in mammalian introns. Genes Dev. *3* (12B), 2113–2123.

Revtovich, A.V., Lee, R., and Kirienko, N.V. (2019). Interplay between mitochondria and diet mediates pathogen and stress resistance in Caenorhabditis elegans. PLoS Genet. *15*, e1008011.

Roignant, J.Y., and Soller, M. (2017). m^{6}A in mRNA: An Ancient Mechanism for Fine-Tuning Gene Expression. Trends Genet. *33*, 380–390.

Rudner, D.Z., Kanaar, R., Breger, K.S., and Rio, D.C. (1996). Mutations in the small subunit of the Drosophila U2AF splicing factor cause lethality and developmental defects. Proc. Natl. Acad. Sci. USA *93*, 10333–10337.

Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell *38*, 317–331.

Schibler, U., Kelley, D.E., and Perry, R.P. (1977). Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. J. Mol. Biol. *115*, 695–714.

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. Nat. Methods *9*, 676–682.

Schwartz, S., Agarwala, S.D., Mumbach, M.R., Jovanovic, M., Mertins, P., Shishkin, A., Tabach, Y., Mikkelsen, T.S., Satija, R., Ruvkun, G., et al. (2013). High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. Cell *155*, 1409–1421.

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

Shi, Y. (2017). Mechanistic insights into precursor messenger RNA splicing by the spliceosome. Nat. Rev. Mol. Cell Biol. *18*, 655–670.

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 N^6-Adenosine Methylation and mRNA Degradation Involving METTL16 and YTHDC1. Cell Rep. *21*, 3354–3363.

Shiraki, N., Shiraki, Y., Tsuyama, T., Obata, F., Miura, M., Nagae, G., Aburatani, H., Kume, K., Endo, F., and Kume, S. (2014). Methionine metabolism regulates maintenance and differentiation of human pluripotent stem cells. Cell Metab. *19*, 780–794.

Shtonda, B.B., and Avery, L. (2006). Dietary choice behavior in Caenorhabditis elegans. J. Exp. Biol. *209*, 89–102.

Sickmier, E.A., Frato, K.E., Shen, H., Paranawithana, S.R., Green, M.R., and Kielkopf, C.L. (2006). Structural basis for polypyrimidine tract recognition by the essential pre-mRNA splicing factor U2AF65. Mol. Cell *23*, 49–59.

Singh, R., Valcárcel, J., and Green, M.R. (1995). Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. Science *268*, 1173–1176.

Śledź, P., and Jinek, M. (2016). Structural insights into the molecular mechanism of the m(6)A writer complex. eLife *5*, e18434.

Soares, L.M., Zanier, K., Mackereth, C., Sattler, M., and Valcárcel, J. (2006). Intron removal requires proofreading of U2AF/3′ splice site recognition by DEK. Science *312*, 1961–1965.

Sun, H., Kang, J., Su, J., Zhang, J., Zhang, L., Liu, X., Zhang, J., Wang, F., Lu, Z., Xing, X., et al. (2019). Methionine adenosyltransferase 2A regulates mouse zygotic genome activation and morula to blastocyst transition. Biol. Reprod. *100*, 601–617.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

Theler, D., Dominguez, C., Blatter, M., Boudet, J., and Allain, F.H. (2014). Solution structure of the YTH domain in complex with N6-methyladenosine RNA: a reader of methylated RNA. Nucleic Acids Res. *42*, 13911–13919.

Valcárcel, J., Gaur, R.K., Singh, R., and Green, M.R. (1996). Interaction of U2AF65 RS region with pre-mRNA branch point and promotion of base pairing with U2 snRNA [corrected]. Science *273*, 1706–1709.

van Delft, P., Akay, A., Huber, S.M., Bueschl, C., Rudolph, K.L.M., Di Domenico, T., Schuhmacher, R., Miska, E.A., and Balasubramanian, S. (2017). The profile and dynamics of RNA modifications in animals. ChemBioChem *18*, 979–984.

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

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 N^6-methyladenosine (m^{6}A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep. *18*, 2004–2014.

Wei, C.-M., and Moss, B. (1977). Nucleotide sequences at the N6-methyladenosine sites of HeLa cell messenger ribonucleic acid. Biochemistry *16*, 1672–1676.

Wei, C., Gershowitz, A., and Moss, B. (1975a). N6, O2′-dimethyladenosine a novel methylated ribonucleoside next to the 5′ terminal of animal cell and virus mRNAs. Nature *257*, 251–253.

Wei, C.M., Gershowitz, A., and Moss, B. (1975b). Methylated nucleotides block 5′ terminus of HeLa cell messenger RNA. Cell *4*, 379–386.

Will, C.L., and Lührmann, R. (2011). Spliceosome structure and function. Cold Spring Harb. Perspect. Biol. *3*, a003707.

Wojtas, M.N., Pandey, R.R., Mendel, M., Homolka, D., Sachidanandam, R., and Pillai, R.S. (2017). Regulation of m^{6}A Transcripts by the 3′→5′ RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline. Mol. Cell *68*, 374–387.e12.

Wu, S., Romfo, C.M., Nilsen, T.W., and Green, M.R. (1999). Functional recognition of the 3′ splice site AG by the splicing factor U2AF35. Nature *402*, 832–835.

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 m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol. Cell *61*, 507–519.

Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., et al. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. Nature *478*, 64–69.

Yoshida, H., Park, S.Y., Oda, T., Akiyoshi, T., Sato, M., Shirouzu, M., Tsuda, K., Kuwasako, K., Unzai, S., Muto, Y., et al. (2015). A novel 3′ splice site recognition by the two zinc fingers in the U2AF small subunit. Genes Dev. *29*, 1649–1660.

Zamore, P.D., and Green, M.R. (1989). Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. Proc. Natl. Acad. Sci. USA *86*, 9243–9247.

Zamore, P.D., Patton, J.G., and Green, M.R. (1992). Cloning and domain structure of the mammalian splicing factor U2AF. Nature *355*, 609–614.

Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

Zhang, Z., Theler, D., Kaminska, K.H., Hiller, M., de la Grange, P., Pudimat, R., Rafalska, I., Heinrich, B., Bujnicki, J.M., Allain, F.H., and Stamm, S. (2010). The YTH domain is a novel RNA binding domain. J. Biol. Chem. *285*, 14701–14710.

Zhang, X., Sun, W., Chen, D., and Murchie, A.I.H. (2020). Interactions between SAM and the 5′ UTR mRNA of the *sam1* gene regulate translation in *S. pombe*. RNA *26*, 150–161.

Zhao, X., Yang, Y., Sun, B.F., Shi, Y., Yang, X., Xiao, W., Hao, Y.J., Ping, X.L., Chen, Y.S., Wang, W.J., et al. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res. *24*, 1403–1419.

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, K.I., Shi, H., Lyu, R., Wylder, A.C., Matuszek, Z., Pan, J.N., He, C., Parisien, M., and Pan, T. (2019). Regulation of Co-transcriptional Pre-mRNA Splicing by m(6)A through the Low-Complexity Protein hnRNPG. Mol. Cell *76*, 70–81.e9.

Zorio, D.A., and Blumenthal, T. (1999a). Both subunits of U2AF recognize the 3′ splice site in Caenorhabditis elegans. Nature *402*, 835–838.

Zorio, D.A., and Blumenthal, T. (1999b). U2AF35 is encoded by an essential gene clustered in an operon with RRM/cyclophilin in Caenorhabditis elegans. RNA *5*, 487–494.

Zuo, P., and Maniatis, T. (1996). The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. Genes Dev. *10*, 1356–1368.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Polyclonal rabbit anti-m^{6}A	Synaptic Systems	Cat. no. 202003; RRID:AB_2279214
Polyclonal rabbit anti-METT10D (METTL16)	Abcam	Cat. no. ab186012
Polyclonal rabbit anti-PARK7	Invitrogen	Cat. no. PA5-13404, RRID:AB_2160112
Polyclonal rabbit anti-Histone H3	Abcam	Cat. no. ab1791, RRID:AB_302613
Monoclonal mouse anti-FLAG	Sigma	Cat. no. F3165; RRID: AB_259529
Monoclonal rat anti-TUBULIN	Abcam	Cat. no. ab6160, RRID:AB_305328
HRP-conjugated anti-rabbit IgG	GE Healthcare	Cat. no. NA934, RRID:AB_772206
HRP-conjugated anti-mouse IgG	Thermo Fisher	Cat. no. A27025, RRID:AB_2536089
HRP-conjugated anti-rat IgG	GE Healthcare	Cat. no. NA935, RRID:AB_772207
Bacterial and virus strains		
BL21(DE3) bacterial strain	NEB	C2527H
DH10EMBacY bacterial strain	Bieniossek et al., 2012	N/A
Chemicals, peptides, and recombinant proteins		
γ-^{32}P-Adenosine triphosphate	Perkin Elmer	Cat. no. NEG002A001MC
^{32}P-Cytidine 3′, 5′ bis(phosphate) [pCp]	Perkin Elmer	Cat. no. NEG019A250UC
^{14}C-S-Adenosyl-L-Methionine	Perkin Elmer	Cat. no. NEC363010UC
40% Acrylamide/Bis Solution 19:1	Bio-Rad	Cat. no. 1610144
30% acrylamide (37.5:1)	National Diagnostic	Cat. no. EC-890
N,N,N′,N′-Tetramethylethylendiamin	Merck	Cat. no. 1107320100
Amersham Prime Western Blotting Detection Reagent	GE Healthcare	Cat. no. RPN2232
Pierce ECL Plus Western Blotting Substrate	Thermo Fisher	Cat. no. 32134
Folic acid	Sigma	Cat. no. F8758-5G
Vitamin B12	Sigma	Cat. no. V6629-250MG
L-Methionine	Sigma	Cat. no. M5308-25G
L-Leucine	Sigma	Cat. no. L8912-25G
L-Cysteine	Sigma	Cat. no. C7602-25G
m7G(5′)ppp(5′)A RNA Cap Structure Analog	NEB	Cat. no. S1405S
Pierce Protease Inhibitor Tablets, EDTA-free	Thermo Fisher	Cat. no. A32965
Benzonase Nuclease	Santa Cruz Biotechnology	Cat. no. sc-202391
Nuclease P1	Sigma	Cat. no. N8630
RiboLock RNase Inhibitor	Thermo Fisher	Cat. no. EO0381
Chloroform	Merck	Cat. no. 102445
RNase H	Thermo Fisher	Cat. no. EN0201
FastAP	Thermo Fisher	Cat. no. EF0651
T4 PNK	NEB	Cat. no. M0201L
10 mM ATP	GE Healthcare	Cat. no. 27-2056-01
100% DMSO	Thermo Fisher	Cat. no. F-515
T4 DNA Ligase	NEB	Cat. no. M0202M
RNase T1	Thermo Fisher	Cat. no. EN0541
RNase A	Sigma	Cat. no. R6513
T4 RNA Ligase 1	NEB	Cat. no. M0204
Water-saturated phenol	AppliChem	Cat. no. A1624

(*Continued on next page*)

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Glycogen, RNA grade	Thermo Fisher	Cat. no. R0551
RiboRuler Low Range RNA Ladder	Thermo Fisher	Cat. no. SM1831
Bouin's solution	Sigma	Cat. no. HT10132
Critical commercial assays/kits		
NEBNext Multiplex Small RNA Library Prep Set for Illumina	NEB	Cat. no. E7300
MinElute Gel Extraction Kit	QIAGEN	Cat. no. 28604
MEGAshortscript T7 Transcription Kit	Life Technologies	Cat. no. AM1354
Dynabeads Protein A	Life Technologies	Cat. no. 10002D
Dynabeads mRNA purification kit	Life Technologies	Cat. no. 61006
Maxima H Minus First Strand cDNA kit	Thermo Fisher	Cat. no. K1682
Phire Green Hot Start II PCR Master Mix	Thermo Fisher	Cat. no. F126L
DC Protein Assay Kit II	Bio-Rad	Cat. no. 5000112
Deposited data		
Deep sequencing datasets	Mendel et al., 2018	GEO: GSE116329
Deep sequencing datasets	This study.	GEO: GSE146873
All raw gel data are deposited at Mendeley Data.	This study.	https://doi.org/10.17632/s92zgtbhjp.1
Experimental models: Cell lines		
Sf21 insect cells for protein production	Eukaryotic Expression Facility, EMBL Grenoble, France	N/A
High Five (Hi5) insect cells for protein production	Eukaryotic Expression Facility, EMBL Grenoble, France	N/A
HeLa cells	ECACC	Cat. no. 93021013
HeLa S3 cells	ECACC	Cat. no. 87110901
Bombyx cell line (BmN4-SID1)	(Mon et al., 2012)	RRID:CVCL_Z091
Experimental models: Organisms/strains		
Mouse: *Mettl16 knock-out*	Mendel et al., 2018	EMMA (EM: 12199)
Mouse: *Mettl16 F187G mutation*	This study	Available from Lead Contact
Mouse: *Mettl16 185PP→AA186 mutation*	This study	Available from Lead Contact
Mouse: *Mettl16 Floxed*	KOMP repository	*Mettl16*$^{tm1a(KOMP)Mbp}$
Mouse: *Ddx4-Cre*	The Jackson Laboratory	Cat. no. 006954, RRID:IMSR_JAX:006954
C. elegans: WT (N2 Bristol strain)	Caenorhabditis Genetics Center	
C. elegans: mett-10 KO Genotype: ZK1128.2(*ok2204*) III.	Caenorhabditis Genetics Center	Strain VC1743 WormBase: WBStrain00036838
C. elegans strains generated in the study	This study	See Table S5
Oligonucleotides		
DNA and RNA oligos	This study	See Table S3
Recombinant DNA		
pACEBac2	Bieniossek et al., 2012	N/A
Human *Mettl16* cDNA	Mendel et al., 2018	NP_076991; NM_024086
Worm *mett-10* cDNA	This study	NP_499247.2, NM_066846.4
Worm *sams-3* full-length gene	This study	Gene ID: 177355
S. pombe U2AF35 cDNA	This study	NP_594945.1, NM_001020376.2
S. pombe U2AF65 cDNA	This study	NP_595396.1, NM_001021303.2
Software and algorithms		
Cutadapt		https://doi.org/10.14806/ej.17.1.200
MEME - Motif discovery tool	Bailey and Elkan, 1994	https://meme-suite.org/meme/

(*Continued on next page*)

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
WebLogo		http://weblogo.berkeley.edu/
R	R Core Team, 2017	https://www.r-project.org
Bowtie	Langmead et al., 2009	http://bowtie-bio.sourceforge.net/
DESeq2	Love et al., 2014	https://bioconductor.org/packages/DESeq2
Bioconductor	Huber et al., 2015	https://www.bioconductor.org/
Salmon	Patro et al., 2017	https://combine-lab.github.io/salmon/
MACS2	Zhang et al., 2008	https://github.com/macs3-project/MACS
MSPC	Jalili et al., 2018	https://genometric.github.io/MSPC/
BLAST	Altschul et al., 1990	http://blast.ncbi.nlm.nih.gov//blast.ncbi.nlm.nih.gov/Blast.cgi
RNAfold	Lorenz et al., 2011	https://www.tbi.univie.ac.at/RNA/
STAR	Dobin et al., 2013	https://github.com/alexdobin/STAR
Other		
Chelating Sepharose Fast Flow beads	GE Healthcare	Cat. no. 17-0575-01
StrepTrap HP	GE Healthcare	Cat. no. 28-9075-46
Superdex S75 10/300 GL	GE Healthcare	Cat. no. 17-5174-01
Superdex 200 10/300 GL	GE Healthcare	Cat. no. 17-5175-01
MethaPhor agarose	Lonza	Cat. no. 50180
Amersham Protran 0.45 mm Nitrocellulose Membrane	GE Healthcare	Cat. no. 10600002
Amersham MicroSpin S-400 HR Columns	GE Healthcare	Cat. no. GE27-5140-01
TRIzol Reagent	Invitrogen	Cat. no. 15596026
TLC PEI Cellulose F plates	Merck	Cat. no. 1055790001
Phosphor Screen BAS IP MS 2025 E	GE Healthcare	Cat. no. 28956475
Peel-A-Way Embedding Mold S22	Polysciences	Cat. no. 18646A-1
Superfrost Plus microscope slides	Thermo Fisher	Cat. no. 4951PLUS4

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact Ramesh S. Pillai (ramesh.pillai@unige.ch).

Materials availability

All unique reagents including plasmids, animal models etc generated in this study are available from the Lead Contact without any restriction.

Data and code availability

The accession number for the deep sequencing data reported in this paper is GEO: GSE146873. Other deep sequencing data used (GEO: GSE116329) are already published. Original data have been deposited to Mendeley Data: https://doi.org/10.17632/s92zgtbhjp.1. Code used in the current study is available from the corresponding authors upon reasonable request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animal Work

Mutant mice were generated at the Transgenic Mouse Facility of University of Geneva or obtained from the Knockout Mouse Project (KOMP). The mice were bred in the Animal Facility of Sciences III, University of Geneva. The use of animals in research at the University of Geneva is regulated by the Animal Welfare Federal Law (LPA 2005), the Animal Welfare Ordinance (OPAn 2008) and the Animal Experimentation Ordinance (OEXA 2010). The Swiss legislation respects the Directive 2010/63/EU of the European Union. Any project involving animals has to be approved by the Direction Générale de la Santé and the official ethics committee of the

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

Canton of Geneva, performing a harm-benefit analysis of the project. Animals are treated with respect based on the 3Rs principle in the animal care facility of the University of Geneva. We use the lowest number of animals needed to conduct our specific research project. Discomfort, distress, pain and injury is limited to what is indispensable and anesthesia and analgesia is provided when necessary. Daily care and maintenance are ensured by fully trained and certified staff. All animals were housed 3-5 per cage and maintained on a 12-hour light/dark cycle, with water and food available *ad libitum*. The use of mice in this work was approved by the Canton of Geneva (GE/162/19, GE/16/19 and GE13).

Generation of catalytic-dead and RNA-binding mutant Mettl16 mouse lines

Mettl16 gene locus is located in mouse on chromosome 10 and consists of 10 exons (Figure 7A). The locus was modified (Transgenic Core Facility of the University of Geneva) using CRISPR/Cas9 technology to introduce mutations meant to destroy the RNA methylation activity (F187G) and RNA-binding activity (185PP-AA186) of METTL16 (Mendel et al., 2018). Mouse embryos of the B6D2F1/J hybrid line (also called B6D2; The Jackson Laboratory, stock no. 100006) were used. This line is a cross between C57BL/6J (B6) and DBA/2J (D2) and is heterozygous for all B6 and D2 alleles. Single-cell mouse embryos were injected with a single guide RNA (gRNA) and a 200 nucleotides long single-stranded DNA (ssDNA) repair template (IDT, Belgium). Sequences of the gRNA and ssDNA repair templates are provided (Table S3). Template for F187G mutation had a mutation CCTCCCTTT to CCTCCCGGC, while the template for 185PP-AA186 mutation had CCTCCCTTT to GCCGCATTT (Figure 7A).

Founder male mice were crossed with wild-type C57BL/6J (Janvier) female partners to obtain germline transmission. Heterozygous male and female animals from the F1 generation were crossed with each other to obtain homozygotes. To genotype the animals, we PCR-amplified the region around the mutations and sequenced the PCR products (Figure S7B). While male and female heterozygotes for both mutations were detected, there were no homozygotes at weaning age (P21), pointing to potential embryonic lethality due to loss of either catalytic activity or loss of RNA-binding ability of METTL16 (Figure 6A).

Conditional Mettl16 knockout mouse generation

Mettl16$^{tm1a(KOMP)Mbp}$ mouse was obtained from the Knockout Mouse Project (KOMP; https://www.komp.org/) repository at University of California, Davis, USA. *Mettl16*$^{tm1a(KOMP)Mbp}$ mouse has the L1L2_Bact_P gene-trapping cassette inserted between exon 6 and exon 9 of the *Mettl16* gene (Figure S7C). This cassette has multiple functionalities. By default, it functions as a gene-trap as it has a LacZ ORF with a polyadenylation signal that is preceded by a splice acceptor site. Thus, the upstream exons of *Mettl16* pre-mRNA will become spliced to the LacZ sequence, and the polyadenylation signal will ensure that transcription terminates prematurely on the *Mettl16* locus. In addition, the cassette also brings two *loxP* sites flanking exons 7 and 8 of *Mettl16*, allowing conditional knockout (cKO) of the gene.

To allow for the conditional knockout of *Mettl16* gene, the FRT-flanked gene-trap cassette is removed by crosses with a mouse line expressing FLP recombinase from the ubiquitous ROSA26 promoter [B6.129S4-Gt(ROSA)26Sor$^{tm1(FLP1)Dym}$/RainJ, The Jackson Laboratory]. Gene-trap cassette removal leaves behind only the *loxP* sites flanking exons 7 and 8, creating the male and female heterozygous floxed *Mettl16*$^{loxP/+}$ mice (Figure S7C). The *Mettl16*$^{loxP/+}$ male and female mice were crossed with each other to obtain homozygous male and female floxed *Mettl16*$^{loxP/loxP}$ mice. In order to delete *Mettl16* in the germline, a mouse line expressing Cre recombinase under germline-specific promoter (*Ddx4*) was obtained: FVB-Tg(Ddx4-cre)1Dcas/J (The Jackson Laboratory). This line (males or females) was first crossed with male or female *Mettl16*$^{+/-}$ heterozygous knockout animals carrying a null allele (Mendel et al., 2018), producing the male and female *Mettl16*$^{-/+}$*;Ddx4*-Cre animals. These male or female animals were next crossed with the male or female *Mettl16*$^{loxP/loxP}$ line, generating male and female animals with conditional knockout (cKO, *Mettl16*$^{loxP/-}$*;Ddx4-Cre*) of the floxed *Mettl16* allele in the germline and male and female *Mettl16* heterozygotes (*Mettl16*$^{loxP/+}$*;Ddx4-Cre*). The germline-specific expression of *Ddx4* starts approximately at embryonic day 16.5 (E16.5) in male embryos, leading to early deletion of *Mettl16* in the testes. Such cKO males were found to be infertile due to an early block in spermatogenesis and had atrophied testes (Figures 6C and 6D). The conditional deletion of *Mettl16* is also expected to take place in the female germline, but we did not examine impact on fertility in such cKO females.

Genotyping

Ear punches of weaned male and female animals (21 days-old) were digested for 120 min at 95°C in 100 µl of buffer containing 10 mM NaOH and 0.1 mM EDTA. After centrifugation at 3000 rpm for 10 min, 50 µL of supernatant was transferred to a new tube containing 50 µL of TE buffer (20mM Tris-HCl, pH 8.0 and 0.1 mM EDTA). 1.5 µl of digestion mix was used for PCR with Phire Greeen Hot Start II PCR Mix (F126L, Thermo Fisher). The annealing temperatures were calculated using Tm calculator (Thermo Fisher). Reactions were examined by 2% agarose gel (Promega, cat.no. V3125) electrophoresis (Figure S7C).

For genotyping the male and female conditional *Mettl16* knockout mice, the primers sequences are provided (Table S3), so is a representative gel showing the PCR products (Figure S7C). For genotyping the *Mettl16*$^{loxP/loxP}$ mice primers MM101 and MM102 were used (Table S3) to detect loxP inserted into the region (WT PCR product: 474 bp, loxP PCR: 439 bp). To genotype conditional *Mettl16* knockout mice (*Mettl16*$^{loxP/-}$*;Ddx4-Cre)* primers (oIMR7643, oIMR7644) detecting *Ddx4-Cre* (PCR product size: 240 bp) as well as detecting *Mettl16*$^{+/-}$ (MM314, MM315; PCR product size: 296 bp) were used. Representative gels showing the PCR products are shown (Figure S7C).

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

For genotyping the male and female point mutant *Mettl16* mice, a region of 308 bp around the mutation site was amplified using MM340 and MM341 primers (Table S3) and Phire Green Hit Start II polymerase producing a 308 bp PCR product. PCR conditions: 98°C for 1 min., 30 x (98°C for 5 s, 65°C for 10 s., 72°C for 15 s.), 72°C for 1 min., 4°C hold. PCR products were purified with QIAquick® PCR Purification Kit (cat.no. 28106, QIAGEN) and sent for Sanger sequencing (Fasteris SA, Geneva) (Figure S7B). Having confirmed the existence of the mutations in the genome, we then designed primers that allow routine genotyping by specific detection of the mutations by genomic PCR. To this end, primers MM342 and MM343 (269 bp PCR product) were used to detect F187G mutation, primers MM348 and MM349 (265 bp PCR product) to detect PP185-186AA mutation (Table S3). The PCR reaction conditions were identical for both mutations: 98°C for 1 min., 30 x (98°C for 5 s, 65°C for 10 s., 72°C for 15 s.), 72°C for 1 min., 4°C hold.

Nematode strains and growth conditions

C. elegans strains were grown under the standard OP50 conditions for maintenance (Brenner, 1974). For experiments, the worms were fed one of the two diets: nutrient-low media (*E. coli* bacterial strain OP50 on NGM plates) or nutrient-high media (*E. coli* strain NA22 on peptone-rich plates), as detailed (Table S4) and indicated for each experiment. N2 (Bristol strain) was used as wild-type, unless otherwise indicated. The list of strains used in this study can be found in Table S5. All of the *C. elegans* experiments and worm maintenance was carried out at 20°C.

For culture conditions were the nutrient-low media was supplemented with various components, the nutrient-low OP50 plates were prepared with additional methionine (Sigma, cat. no. M5308), leucine (Sigma, cat. no. L8912) or cysteine (Sigma, cat. no. C7602) (with the final concentration of 10 mM for each aminoacid), with folic acid (Sigma, cat. no. F8758) (100 μM concentration) or vitamin B12 (Sigma, cat. no. V6629) (73 nM concentration). The required additional components were added to the cooled autoclaved media, just before pouring plates.

Generation of *C. elegans* strains

All of the genome editing for strain creation was performed in the endogenous loci of *sams-3, sams-4, sams-5* and *mett-10* genes using CRISPR/Cas-9 technology as described in Arribere et al. (2014). In brief, Cas-9 and sgRNAs in the form of plasmids and repair templates as single-stranded oligonucleotides were delivered to the worm germ cells through microinjection into the gonad. Sequences of sgRNAs, repair templates and plasmids used to detect and sequence the edits are indicated in the Table S3. SAMS-3 was tagged with 2xHA on C terminus, METT-10 was tagged with 3xFLAG and 1xHA on the C terminus, intron 2 was removed from *sams-3*, *sams-4* and *sams-5* (Figure 4F). All of the edits were performed on the endogenous copy of the genes. For intron 2 deletions, multiple alleles were generated in different genetic backgrounds (Figure 4G).

Generation of *C. elegans* lines expressing transgene reporter constructs as transgenes

For the transgene reporter constructs experiments (Figures 3D and S3A), either the wild-type N2 (Bristol strain) or the *mett-10* knockout VC1743 strain was used for plasmid injections. Generation of extrachromosomal arrays was carried out via microinjection as described in WormBook" (Evans, 2006) Briefly, the plasmids containing WT or MUT transgene reporter constructs were co-injected along with reporter plasmids pRF4 [*rol-6*(*su1006*), causes roller phenotype due to a cuticle defect] and pCFJ421 [Pmyo-2::GFP::H2B (pharynx)] into the gonads of young adult wild-type or *mett-10* knockout worms. Concentrations of plasmid injected are 5 ng/μl (transgene reporter constructs), 2 ng/μl (pRF4) and 5 ng/μl (pCFJ421). Progenies displaying phenotypes induced by presence of reporter plasmids (rollers with strong pharyngeal GFP signal) were singled out. The constructs are expected to be maintained as extrachromosomal arrays. In the following generation, 6 lines showing the highest rate of array transmission were identified and the presence of WT or MUT plasmid was confirmed by RT-PCR using MM363 and MM364 primers (Table S3). The transgenic lines were maintained by picking rollers. For the splicing analysis, we used three independent lines for the triplicate repetitions. For each repetition, 30 rollers were picked into 100 ul of TRIzol (Thermo Fisher, cat.no. 15596026).

Collection of *C. elegans* for the RNA isolation

For analysis of wild-type and *mett-10* knock-out (VC1743) worms, synchronized adult population was washed off either from 15 cm plates (in case of NA22) or from 6 cm plate (in case of OP50). Worms were washed 3 times in M9 buffer (3 g KH_2PO_4, 6 g Na_2HPO_4, 5 g NaCl, 1 mL 1 M $MgSO_4$ in 1l H_2O), put in Trizol (3x volume of the worm pellet), flash-frozen in liquid nitrogen and stored at −80°C until the RNA isolation.

Cell lines

Cell lines were obtained from the European Collection of Authenticated Cell Cultures (ECACC). HeLa cells (ECACC, cat. no. 93021013) were grown in Dulbecco's modified Eagle Medium (DMEM; Invitrogen, cat. no. 21969-035) supplemented with 10% fetal bovine serum (Thermo Fisher; cat. no. 10270106), 1% Penicilline/Streptomycin (Thermo Fisher; cat. no. 15140122), 2 mM L-Glutamine (Thermo Fisher; cat. no. 25030024). HeLa S3 (ECACC, cat. no. 87110901) were grown in spinner flasks (Corning, cat. No. 4500-125) in MEM medium supplemented with 2 mM L-Glutamine, 1% Non Essential Amino Acids (NEAA) (Thermo Fisher, cat. no. 11140050), 1% Penicilline/Streptomycin and 10% fetal bovine serum (Thermo Fisher). Both cell types were maintained in an environment with 5% CO_2 at 37°C. HeLa cells were sub-cultured at 1:5 ratio every 3 to 4 days using 0.05% trypsin-EDTA (GIBCO, cat. no. 25300054), while HeLa S3 were counted daily using Neubauer chamber and maintained at the 200 000 – 500 000 cells/mL concentration by diluting the culture with growth media.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

METHOD DETAILS

Clones and constructs

The *C. elegans* METT-10 (Wormbase: CE31860), was cloned by RT-PCR amplification from adult worm total RNA. The *sams-3* SAM synthetase gene (Wormbase: CE03957) was PCR amplified from worm genomic DNA. The cDNAs for yeast (*S. pombe*) U2AF35 (U2AF23, UniProtKB/Swiss-Prot: Q09176.2) and yU2AF65 (U2AF59, UniProtKB/Swiss-Prot: P36629.1) were synthesized (Thermo Fisher).

Constructs for bacterial protein expression

The untagged full-length yeast (*S. pombe*) U2AF35 (U2AF23) and a tagged (6xHis-StrepII) fragment (93-161 aa) of yU2AF65 (U2AF59) were co-expressed in *E. coli* and purified as a complex. This minimal fragment of yU2AF65 (93-161 aa) is essential and sufficient for ensuring stability of yU2AF35 (Yoshida et al., 2015). The required ORFs were cloned into the pETDuet-1 vector (Novagen) for co-expression in *E. coli*. We also prepared complexes where the yU2AF35 has specific point mutations in the zinc finger 1 (Figures S6B–S6D). The Arginine 35 in ZF1 is proposed to be involved in recognition of the splice site adenosine (Yoshida et al., 2015). A conservative mutation to a positively-charged lysine (R35K) or to a non-conservative mutation to uncharged serine (R35S) were made. The logic of R35K mutation was to see if the shorter side-chain of Lysine could allow recognition of m^{6}A. We also made a mutation (S34Y) in the serine 34, which is frequently mutated in patients with myelodysplastic syndromes (Yoshida et al., 2011). The constructs used for recombinant protein production were verified by restriction digest, as well as by Sanger sequencing.

Constructs for insect cell expression

The worm METT-10 ORF was cloned into pACEBac2-His-StrepII-SUMO vector (Geneva Biotech) for expression in Sf21 or High Five (Hi5) insect cells as a 6xHis-StrepII-SUMO-tagged fusion. The constructs used for recombinant protein production were verified by restriction digest, as well as by Sanger sequencing. Expression construct for human METTL16 (hMETTL16) was previously reported (Mendel et al., 2018).

Constructs for expression of *sams-3* transgene reporters in transgenic worms

Transgene reporter constructs were based on modified full-length *C. elegans* SAM synthetase gene *sams-3* (Wormbase: CE03957). The whole *sams-3* gene (2189 nt, including 5′ and 3′UTR sequences) was amplified using DH298 and DH299 primers, but to distinguish the transgene reporter constructs from the endogenous *sams-3* transcript, 20 nt-long artificial sequences were placed into the exon2 (TGAACGACCGTGTTCTAGGG, DH300 and DH301) and exon3 (ACAGCCTACTTTGAGTGCGTA, DH302 and DH303), allowing for a specific PCR amplification (Table S3). The inserted 20-nt artificial sequences also rendered the reporters non-coding, as they were designed to cause a frameshift in the protein coding sequence. In addition, the wild-type (WT) METT-10 methylation consensus motif (UAC**AG**AAAC) that overlaps the 3′ splice site AG was mutated (MUT) in the part of the consensus that belongs to the exon3 (UAC**AG**ACUU, mutation is underlined). The mutation was introduced using MM320 and MM321 primers, and by amplifying the whole plasmid (Table S3). Such a mutation is demonstrated to prevent methylation by METT-10 *in vitro* (Figure 3B). Making the mutations on the exonic part is meant to reduce the disruption of binding sites for the key splicing factors. The constructs were cloned into the pUC19 plasmid for *C. elegans* expression under the *his-72* promoter to ensure ubiquitous expression of the construct. The constructs were verified by restriction digest, as well as by Sanger sequencing.

Constructs for expression of transgene reporters in human cells

Sams-3 full-length fragment was PCR-amplified from the pUC19-sams-3 plasmid prepared for worms injections, using MM415 and MM416 primers (Table S3). Forward primer introduced NotI restriction site, while reverse primer introduced NheI site. The amplified fragment was cloned into the mammalian expression vector phRL-TK (Promega), which was first digested with NotI-HF (NEB, cat. no. R3189S) and NheI-HF (NEB, cat. no. R3131S) restriction enzymes. This removes the whole luciferase (hRL) sequence. The phRL-|TK vector has the HSV Thymidine Kinase promoter, allowing low-level expression in mammalian cells. The final constructs used were verified by restriction digest, as well as by Sanger sequencing.

Antibodies

The polyclonal rabbit anti-m^{6}A (Synaptic Systems; cat. no. 202003) for m^{6}A-IP-seq, polyclonal rabbit anti-METTL16 (abcam; cat. no. ab186012) to detect mouse METTL16, polyclonal rabbit anti-PARK7 (Invitrogen, cat. no. PA5-13404) to detect mammalian PARK7 as a normalization control for western blots, anti-HA (a kind gift from Marc Bühler, clone #42F13) to detect HA-tag, anti-FLAG (Sigma, cat. no. F3165) to detect FLAG-tag, anti-TUBULIN (Abcam; cat. no. ab6160) to detect worm TUBULIN, and anti-H3 (Abcam; ab1791) to detect worm histone H3 were used. For secondary antibodies, the HRP-conjugated anti-rabbit IgG HRP-linked (GE Healthcare; cat. no. NA934), HRP-conjugated anti-mouse IgG (H+L) Superclonal Secondary Antibody (Thermo Fisher, cat. no. A27025) or HRP-conjugated anti-rat IgG (GE Healthcare, cat. no. NA935) were used.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS

Recombinant protein production

Production of full-length recombinant proteins was carried out either in insect cell lines using the baculovirus expression system or in the prokaryotic expression system using *E. coli*.

The insect ovary-derived cell lines used are: High Five (Hi5) insect cell line originating from the cabbage looper (*Trichoplusia ni*) and the Sf9 cells derived from the fall army worm *Spodoptera frugiperda*. Briefly, recombinant full-length hMETTL16 (Mendel et al., 2018) or worm METT-10 coding sequences were cloned into the pACEBac2-Sumo acceptor vector (His-Strep-Sumo tag) (Bieniossek et al., 2012). Plasmids were transformed into DH10EMBacY competent cells for recombination with the baculovirus genomic DNA (bacmid). The bacmid DNA was extracted and transfected with FuGENE HD (Promega, cat. no. E231A) into the *Sf*9 insect cells for virus production. The supernatant (V_0) containing the recombinant baculovirus was collected after 72 to 96 hours post-transfection. To expand the virus pool, 6.0 mL of the V_0 virus stock was added into 25 mL of *Sf*9 (0.5×10^6/mL) cells. The resulting cell culture supernatant (V_1) was collected 24 h post-proliferation arrest. For large-scale expression of the protein, Hi5 cells were infected with virus (V_1) and cells were harvested 72 h post-proliferation arrest.

For bacterial expression, plasmids were transformed into the *E. coli* BL21(DE3) strain and expression was initiated by addition of 0.7 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) when the culture density reached 0.6 (OD_{600}). The proteins were then expressed overnight at 20°C following induction.

Expression and purification of yeast U2AF35 protein

The U2AF heterodimer is formed by interactions between the large subunit U2AF65 and small subunit U2AF35 (Zamore and Green, 1989). The U2AF65 has three RNA recognition motifs (RRMs), while the U2AF35 has one RRM flanked by two CCCH-type zinc fingers (ZFs). Only a 28-amino acid fragment from U2AF65 is required for interaction with U2AF35 (Kielkopf et al., 2001). A soluble *S. pombe* U2AF complex consisting of full-length U2AF35 and a larger region (93-161 aa) encompassing the 28-amino acid proline-rich fragment from U2AF65, was previously described (Yoshida et al., 2015). This complex is shown to specifically recognize the AG dinucleotide, as single mutations in the RNA at these positions either abolish or greatly reduce binding to the RNA (Yoshida et al., 2015). Importantly, this complex does not bind a polypyrimidine stretch (U_{10}) (Yoshida et al., 2015). The ZFs in U2AF35 cooperatively bind the RNA (Yoshida et al., 2015).

We used this *S. pombe* U2AF complex (which has RNA-binding property only in U2AF35) for our studies, and for simplicity is referred to as full-length yeast U2AF35.

The ORFs for yU2AF35 and His-StrepII tag fused yU2AF65 (93-161 aa) were cloned into pETDuet-1 vector, and the plasmid was transformed into *E. coli* BL21 (DE3) strain for co-expression. The protein complex was expressed overnight at 20°C with 0.5 mM IPTG. Cells were harvested and lysed by sonication in lysis buffer (30 mM Tris-HCl, pH 8.0, 200 mM NaCl, 5 mM 2-mercaptoethanol, 5% Glycerol, 20 mM Imidazole and proteinase inhibitor). The supernatant after centrifugation was subjected to Ni-NTA column and StrepTrap column for obtaining His-StrepII tagged protein complex. After removal of the tags by overnight TEV cleavage, the untagged protein complex was further purified by size exclusion column (Superdex S75) in buffer (20 mM HEPES, pH 7.0, 100 mM NaCl).

ITC experiment with yeast U2AF35

ITC experiments were performed (at EMBL Grenoble, France; kind help of Dr. Andrew McCarthy) using a MicroCal iTC 200 (Malvern Panalytical) at 20°C. The yeast U2AF35 protein, as well as RNAs, were dialyzed overnight in the buffer (20 mM HEPES, pH 7.0, 100 mM NaCl). The sample cell was filled with 50 μM of either unmethylated or m^{6}A methylated RNA (CUAGG, methylated adenosine is underlined), and the syringe was filled with 500 μM yeast U2AF protein. Titrations were carried out with a first 0.4 μL injection followed by constant volume injections (19 injections of 2 μL) with 150 s spacing. Data analysis was performed using Origin software.

Worm total RNA purification

Worms were collected in TRIzol as described in the "Collection of C. elegans for the RNA isolation" section. To isolate RNA, worms were first kept in TRIzol at room temperature for 1 – 2 hours with frequent vortexing to ensure complete lysis. Then, samples were spun for 10 min. at 14000 x g at 4°C and supernatant was transferred to fresh tubes. Chloroform (Merck, cat. no. 102445) was added to TRIzol in 1 to 5 volume ratio and the tubes were first vortexed for 15 s. and then left at room temperature for 3 minutes. Next, the tubes were spun at 14000 x g for 15 min. at 4°C. The upper layer (aqueous phase) was transferred to fresh tubes and an equal amount of chloroform was added. The tubes were vortexed for 30 s. and spun at 14000 x g for 10 min. at 4°C. The upper layer (aqueous fraction) was transferred to fresh tubes, where 2.5 volume of 100% ethanol (VWR, cat. no. 20821.321) was added. The tubes were stored at −20°C for at least 1 hour to precipitate the RNA. After precipitation, the tubes were spun at 14000 x g for 30 min. at 4°C. The RNA pellet was washed once with 70% ethanol, dried at RT and resuspended in RNase-free water (Invitrogen, cat. no. 10977-05). The isolated RNA was stored at −80°C to avoid RNA degradation.

Poly(A)$^+$ RNA purification

Worm poly(A)$^+$ RNA was purified using Dynabeads mRNA Purification Kit (Invitrogen, cat. no. 61006). All the step were performed accordingly to the protocol provided by the manufacturer. Briefly, 225 μg of RNA was diluted in RNase-free water to a final volume of 300 μl. RNA was then heated to 65°C for 3 minutes and placed on ice. At the same time, 600 μl of resuspended Dynabeads Oligo

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

$(dT)_{25}$ beads were transferred to fresh tubes, which were placed on a magnetic stand. The supernatant was removed and 300 µl of Binding Buffer was added to equilibrate the beads. Tubes were again placed on a magnetic stand and supernatant was removed. Another 300 µl of Binding Buffer was added and the beads were mixed with the RNA (1:1 volume ratio). The RNA was incubated with the beads for 15 min. at RT with rocking. Then, tubes were placed at the magnetic stand and supernatant was removed. Beads were washed twice with 600 µl Washing Buffer B. Next, beads were resuspended in 40 µl of water, heated to 75°C for 2 minutes and immediately placed on ice. The supernatant with eluted mRNA was transferred into fresh tubes and stored at −80°C to avoid RNA degradation.

Detection of m^{6}A methylation using SCARLET

We followed the method described previously (Liu et al., 2013) to produce the data presented in Figures 1J and S1H. The oligonucleotides used are provided in Table S3. In short, 1 µg of total RNA or mRNA, isolated as previously described, was mixed with 3 pmol of chimeric oligo in a total volume of 3 µl of 30 mM Tris-HCl, pH 7.5. We tested different lengths (17/18 nt, 20 nt, 23 nt, 30 nt) of chimeric oligos for both U6 snRNA and *sams-3* targets. Generally, the shortest chimeric oligos (18 nt) were performing the best, although in the case of U6 snRNA the difference wasn't dramatic. On the other hand, in the case of the *sams-3* target, there was a huge increase in cleavage efficiency from 20 nt to 18 nt chimeric oligo. Thus, we recommend using short chimeric oligos and testing different sizes.

The mix was heated for 1 min. at 95°C followed by incubation at room temperature for 3 min. before placing on ice. Then, 1 µl of 5x RNase H mix (2 x T4 polynucleotide kinase buffer (NEB, cat. no. B0201S), 1 unit/µl of RNase H (Thermo Fisher, cat. no. EN0201)) and 1µl of FastAP (Thermosensitive Alkaline Phosphatase, 1U/µl, Thermo Scientific, cat. no. EF0651) were added to the tube (total volume was 5 µl). Samples were incubated for 1 hour at 44°C and then heated for 5 min. to 75°C in order to inactivate RNase H and FastAP. Immediately after the heating, samples were placed on ice. We have tested multiple RNase H enzymes: NEB, cat. no. M0297S; Thermo Scientific, cat. no. EN0201; Sigma, cat. no. R6501; Invitrogen, cat. no. 18021014; Takara, cat. no. 2150A. Although all of them worked, RNase H from Thermo Scientific (cat. no. EN0201) and RNase H from Sigma (cat. no. R6501) seemed to be the most efficient in our hands. After the RNase H digest, the RNA was 5′-end labeled with ^{32}P by adding 1 µl of 6 x T4 PNK mix (1 x T4 PNK buffer, 6 units/µL of T4 PNK (NEB, cat. no. M0201L) and 28 µCi/µL [γ-^{32}P] ATP (PerkinElmer, NEG002A001MC)). The mix was incubated for 1 hour at 37°C, then heated for 5 min. at 75°C in order to inactivate T4 PNK and immediately put on ice.

We took 1.5 µl mix of 4 pmol of splint oligo/5 pmol of ssDNA universal oligo (ssDNA-116/MM437) was added to the tubes. The mix was annealed by heating for 3 min. at 75°C and cooling down for 3 min. at room temperature, then it was put on ice. 2.5 µL of 4 x ligation mix (1.4x T4 PNK buffer, 0.27 mM ATP (GE Healthcare, cat. no. 27-2056-01), 57% DMSO (ThermoScientific, F-515), 80 U/µL T4 DNA ligase (NEB, M0202M)) was added and RNA was ligated for 3.5 hours at 37°C. The reaction was stopped with 2 x RNA loading buffer (9 M urea, 100 mM EDTA, xylene and bromophenol dye) and 1 µL of RNase T1/A mix (160 U/µL RNase T1 (ThermoScientific, cat. no. EN0541) and 0.16 mg/mL RNase A (Sigma, cat. no. R6513-10MG)) was added. The RNA was incubated overnight at 37°C. Next, it was run on 10% Urea-PAGE gel together with a ^{32}P labeled ssDNA universal oligo (ssDNA-116/MM437) used as a marker. The band corresponding to 117-/118-bp was cut out of the gel. RNA was eluted from the gel for 4 h at 25°C with 750 rpm shaking using RNA extraction buffer (300 mM sodium acetate, 1 mM EDTA, 0.1% SDS). Supernatant was transferred to a fresh tube and RNA was isolated with phenol/chloroform. RNA pellet was then resuspended in 3 µl of Nuclease P mix (0.33 U/µL of Nuclease P1 (Sigma, cat. no. N8630) in 30 mM sodium acetate pH 4.8) and incubated for 2 h at 37°C.

After digestion, 1 µL of digested RNA was transferred into TLC PEI Cellulose F plate (Merck, cat. no. 1055790001) and was resolved for 14 h in a mix of isopropanol:HCl:water (70:15:15, v/v/v). After that, the TLC plate was dried at RT for 30 – 60 min., wrapped in a plastic film and exposed to a phosphor screen BAS IP MS 2025 E (GE Healthcare, cat. no. 28956475). Phosphor screen was scanned in a Typhoon FLA 9500 laser scanner (GE Healthcare) at 700V and 100 µm pixel size.

Quantification of *sams-3* splicing by RT-PCR

The total RNA was isolated from various sources: WT worms grown on various food sources, or transgenic worms with WT or MUT transgene reporter constructs or HeLa cells expressing worm reporter constructs. The RNA was reverse transcribed using Maxima H Minus First Strand cDNA (Thermo Fisher, K1682) with random primers. The cDNA was diluted to 50 µl, and of this a 2 µl aliquot was used for PCR. Primers used for the PCR were MM395 and MM396 in the case of endogenous *sams*-3 transcript or MM363 and MM364 in the case of transgene reporter constructs construct (Table S3). Transgene reporter constructs PCR amplification (MM363, MM364) generates three bands: intron-retained transcript (517 bp), alternatively spliced transcript (239 bp) and fully spliced isoform (148 bp), while for the endogenous *sams-3* transcript PCR generates: intron-retained transcript (502 bp), alternatively spliced transcript (225 bp) and fully spliced isoform (133 bp). The PCR products were resolved in a 2% agarose gel (Figures 3D, S3A, 5A, S4C, and S5B), stained with ethidium bromide and visualized under UV light in a gel visualization system (Vilber Lourmat E-Box VX2). Gel pictures were analyzed using Fiji software package (Schindelin et al., 2012). The intensity of each transcript isoform was calculated using the gel analyzer function. It was then normalized to the total intensity of all the transcript isoforms within one sample, allowing for internal normalization and comparison of different samples. Results obtained by RT-PCR quantification of worms grown on different food sources are in agreement with RNA-seq quantification.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS

Quantification of RNA modifications using LC-MS/MS

Total RNA was isolated by Trizol extraction from adult *C. elegans* (worm), the *Bombyx mori* (Silkmoth) BmN4 cell line (insect) and adult mouse testes (mouse), as indicated in Figure 1B. These RNAs were also used to purify poly(A)$^+$ RNA. The RNAs were hydrolyzed to ribonucleosides by 20 U Benzonase® Nuclease (Santa Cruz Biotech, cat. no. sc-202391) and 0.2 U Nuclease P1 (Sigma, cat. no. N8630-1VL) in 10 mM ammonium acetate pH 6.0 and 1 mM magnesium chloride at 40°C for 1 h. After that, ammonium bicarbonate to 50 mM, 0.002 U phosphodiesterase I and 0.1 U alkaline phosphatase (Sigma) were added, and incubated further at 37°C for 1 h. The hydrolysates were mixed with 3 volumes of acetonitrile and centrifuged (16,000 x g, 30 min, 4°C). The supernatants were dried and dissolved in 50 µL water for LC-MS/MS analysis of modified and unmodified ribonucleosides. Chromatographic separation was performed using an Agilent 1290 Infinity II UHPLC system with an ZORBAX RRHD Eclipse Plus C18 150 × 2.1 mm ID (1.8 µm) column protected with an ZORBAX RRHD Eclipse Plus C18 5 × 2.1 mm ID (1.8 µm) guard column (Agilent). The mobile phase consisted of water and methanol (both added 0.1% formic acid) run at 0.23 mL/min. For modifications, starting with 5% methanol for 0.5 min followed by a 2.5 min gradient of 5%–15% methanol, a 3 min gradient of 15%–95% methanol and 4 min re-equilibration with 5% methanol. A portion of each sample was diluted for the analysis of unmodified ribonucleosides which was chromatographed isocratically with 20% methanol. Mass spectrometric detection was performed using an Agilent 6495 Triple Quadrupole system, monitoring the mass transitions 268.1-136.1 (A), 284.1-152.1 (G), 244.1-112.1 (C), 245.1-113.1 (U), 296.1-150.1 (m^{6}Am), 282.1-150.1 (m^{6}A and m^{1}A), 282.1-136.1 (Am), 296.1-164.1 (m$^6{}_2$A), 283.1-151.1 (m^{1}I), 298.1-166.1 (m^{7}G), 312.1-180.1 (m^{2,7}G), 326.1-194.1 (m^{2,2,7}G), 258.1-126.1 (m^{3}C and m^{5}C), 274.1-142.1 (hm^{5}C), 286.1-154.1 (ac^{4}C), 259.1-139.1 (m1Ψ), 333.1-201.1 (5-methoxycarbonylmethyl-2-thiouridine, mcm5s2U), and 333.1-183.1 ((*S*)- and (*R*)-5-methoxycarbonylhydroxymethyluridine, S-mchm5U and R-mchm5U) in positive electrospray ionization mode, and 267.1-135.1 (inosine, I) and 243.1-153.1 (pseudouridine, Ψ) in negative electrospray ionization mode. Modifications detected in a mock control (containing only the hydrolytic enzymes) were subtracted from modifications detected in the RNA samples. In general, the mock control contained at least 1000-fold less RNA than the RNA samples and gave negligible background in the modification analyses.

Metabolomics analyses of worm lysates

Approximately, 3-4 L4 hermaphrodite worms were placed on 6 cm plates containing either nutrient-low media (NGM medium and seeded with OP50 bacteria) or nutrient-high media (peptone-rich medium and seeded with NA22 bacteria) (Table S4). After around 5 days of incubation at 20°C, when the progeny reached adult stage, the worms were washed off the plates with PBS, washed 2 more times with PBS and once with water. The excess of the liquid was discarded and packed worm pellet was immediately snap-frozen in liquid nitrogen and stored at −80°C. Then we shipped the lysates in dry ice to the Metabolomics Platform, Faculty of Biology and Medicine, University of Lausanne, Switzerland. Analysis was conducted as previously described for polyamines concentration measurement (Chevalier et al., 2020).

In vitro RNA methylation assay with human METTL16 and worm METT-10

Methylation assays were carried using chemically synthetized RNA oligos (IDT, Belgium or Microsynth, Switzerland) (Table S3). Recombinant untagged full-length human METTL16 or full-length *C. elegans* METT-10 were used (Figure 3B).

Before the experiment, RNA was refolded by heating the 100 µM RNA solution in H_2O or 50 mM NaCl to 65°C in a Thermoblock (Eppendorf) for 5 min, and allowed to slowly cool down to the room temperature. All methylation reactions were performed in a 50 mM Tris-HCl, pH 7.5, 100 mM KCl, 5 mM MgCl2, 2 mM DTT buffer with 10 µM of refolded single-stranded RNA, 5 µg of recombinant protein, 1 µl of RiboLock RNase Inhibitor (Thermo Fisher, cat. no. EO0381) and 0.02 µCi of ^{14}C-SAM (Perkin Elmer, NEC363010UC) in a total volume of 20 µl. Unless otherwise indicated, all reactions were performed overnight at room temperature. RNA was subsequently extracted using phenol/chloroform extraction protocol, resuspended in 15 µl of 2x RNA loading buffer (90% formamide, 0.02% SDS, 1 mM EDTA, 0.02% bromophenol blue, 0.02% xylene cyanol), heated for 5 min. at 70°C, cooled down to room temperature and loaded on a 15% Urea-PAGE gel.

The 15% Urea-PAGE gel was prepared by mixing 12.6 g of urea, 3 mL of 10x TBE (1 M Tris base, 1 M boric acid, 0.02 M EDTA), 11.25 mL of 40% acrylamide (19:1) and 6.75 mL of H2O. To catalyze gel polymerization, 240 µl of APS and 24 µl of TEMED (Merck, cat. no. 1107320100) were added. Gel was left to polymerize for 40 min. at room temperature. Wells were washed with 1xTBE to remove urea deposits and gel was pre-run in 1x TBE at 20 W for 25 min. RNA markers 5′ end-labeled with ^{32}P-γ-ATP and composed of four single-stranded RNA oligos (RP_RNA_19: 40 nt, RP_RNA_1: 30 nt, RP_RNA_3: 28 nt, RP_RNA_18: 16 nt; Table S3) were loaded into the gel, together with 10 µl of RNA samples from the *in vitro* methylation assay. Gel was run at 12 W for 1 h 30 min.

After running, to visualize RNA bands, the gel was stained with a methylene blue solution [0.2% (w/v) methylene blue in a 1:1 solution of 0.4M sodium acetate and 0.4M acetic acid] for 10 min. This staining can be done by carefully sealing the gel in a plastic bag and shaking it on a rocking platform/frequently mixing the contents. Next, the gel was destained with 1xTBE and scanned using Epson Perfection 3200 Photo scanner. The destaining is also done by removing the staining solution, injecting the wash buffer into the bag and carefully mixing the contents. After scanning, the gel was dried in a gel dryer (Bio-Rad, model 583) with a gradual heating program, 80°C for 1.5 h. The dried gel was transferred to a cassette and exposed with a phosphor screen BAS (GE Healthcare) for 24 h. The phosphor screen was scanned in a Typhoon FLA 9500 laser scanner (GE Healthcare) at 700V and at 100 µm pixel size. Control software used for Typhoon FLA 9500 is the 1.1 version. Scans were analyzed using ImageQuant TL 8.1 software (GE Healthcare).

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

Preparation of RNA substrates for *in vitro* splicing assay

Both the human β-globin and adenovirus-based MINX splicing constructs were prepared by splint ligation (Moore and Sharp, 1992) of two RNA fragments. A longer T7 transcribed 5′ fragment that has 5′ exon and most of the intron, while a shorter 3′ synthetic RNA fragment (IDT, Belgium) that has the 3′ splice site and the 3′ exon. This allows introduction of either a methylated or unmethylated adenosine at the splice site by chemical synthesis.

To obtain the 5′ RNA fragment, both the β-globin and MINX DNA fragments were amplified by PCR. Forward primers contained T7 promoter sequence, while reverse primers contained 2′-O-methyl residues at the last two nucleotides to prevent non-template nucleotide addition by the T7 polymerase. In addition, the MINX construct was mutated in order to create a version with a weaker polypyrimidine tract. The mutation was introduced by using a modified reverse primer (MM583) for the PCR reaction (Table S3). The PCR fragments were purified and used for T7 transcription with an m^{7}G cap analog (m^{7}G(5)ppp(5)G, NEB, Cat. No. S1404S) used in 4:1 ratio to GTP. Using m^{7}G analog is essential as only capped transcripts are efficiently spliced. After the *in vitro* transcription, RNA was first purified with MicroSpin G-25 size exclusion columns (GE Healthcare, cat. no. 27-5325-01), and then it was extracted with phenol/chloroform.

The 3′ synthetic fragments without methylation, or with m^{6}A at the 3′ splice site, or with m^{6}A within the 3′ exon were purchased (IDT, Table S3). 50 pmol of the T7 transcribed RNA and 50 pmol of the synthetic RNA were mixed with 50 pmol of an antisense DNA oligo splint (that has extensive complementarity to the two RNA ends that need to be joined) in a total volume of 15 μl of an annealing buffer (10 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA). The mix was heated in a thermocycler to 90°C and cooled down to 25°C (a program with gradual cooling at −0.1°C/second was used). Next, the RNA-DNA hybrid was incubated with 6μl of T4 DNA ligase mix (1x T4 ligase buffer, 1 μl of high-concentration T4 DNA ligase (2,000U/μl; NEB, cat.no. M0202M), 1 μl of RiboLock RNase Inhibitor, 2 μl PEG8000 (50%)) for 4 h at 37°C. The reaction was subsequently purified with phenol/chloroform and resuspended in 10 μl of H_2O.

Approximately, 1 – 2 μl of ligated RNA was used for 3′ end labeling reaction, where it was mixed with 13 – 14 μl of pCp ligation mix(1 x T4 RNA Ligase 1 reation buffer, 1 mM ATP, 10% DMSO, 1 μl of T4 RNA Ligase 1 (NEB, cat.no. M0204), 1 μl of ^{32}P-labeled cytidine 3′, 5′ bis(phosphate) (pCp, PerkinElmer, cat.no. NEG019A250UC), 1 μl of RiboLock RNase Inhibitor and water) and incubated overnight at 4°C. This reaction was then loaded on a 5% Urea-PAGE gel and the band corresponding to ~200 nt ligated RNA was cut out, eluted and resuspended in 20 – 30 μl of H_2O. This gives a splicing pre-mRNA substrate that is protected at the 5′ end with an m^{7}G cap and radioactively marked at the 3′ end with ^{32}P-labeled pCp.

Preparation of nuclear extracts

Splicing extracts were prepared as described before (Lee et al., 1988) using HeLa S3 cells. The cells were maintained at the concentration of 200 000 – 500 000 cells/mL and prior to collection were expanded to 600 000 – 800 000 cells/mL. It is important to harvest cells at the logarythimc growth stage. 400 mL of HeLa S3 suspension culture was collected (1200 x g, 5 min, 4°C), washed twice with 1xPBS and spun down (1200 x g, 5 min, 4°C) to assess packed cell volume (PCV). 300 mL of cell culture resulted in approximately 1 mL of PCV. All subsequent steps were performed on ice in a cold room (4°C). All buffers (placed in an ice bucket), pipette tips, Eppendorf tubes, etc. were pre-chilled in the cold room before use. Cells were gently resuspended in buffer A (10 mM HEPES-KOH, pH 7.9, 1.5 mM MgCl2, 10 mM KCl, 0.5 mM DTT) in a volume equal to PCV. After 15 min. cells were passed 6-times through a 23 gauge syringe (vigorous passage) and centrifuged at 12 000 x g for 20 s. at 4°C. The supernatant was removed and crude nuclear pellet was resuspended in buffer C (20 mM HEPES-KOH pH 7.9, 25% glycerol, 420 mM NaCl, 1.5 mM MgCl2, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT) in the volume of 2/3rd of PCV. Resuspended nuclear pellet was mixed on a tube revolver (Thermo Scientific, cat. no. 88881002) with 15 rpm speed for 30 min. at 4°C. Next, it was spun down at 12 000 x g for 5 min. at 4°C. Supernatant was transferred to a fresh tube and spun again (12 000 x g for 5 min. at 4°C), while the remaining pellet was removed. Supernatant (nuclear extract) was dialyzed twice for 2 hours in 100-times the volume of extract in buffer D (20 mM HEPES-KOH pH 7.9, 20% glycerol, 100 mM KCl, 0.2 mM EDTA, 0.5 mM PMSF, 0.5 mM DTT). Protease inhibitor PMSF was always added fresh, just before using the buffers. Extracts were flash frozen in liquid nitrogen and stored at −80°C. Quality of the lysates was verified by retrieving a frozen aliquot for use in a splicing assay.

In vitro splicing reaction

In vitro splicing reaction was done as described previously (Mayeda and Krainer, 1999) In short, 15 μl of nuclear extract was mixed on ice with 10 μl of a mix containing the RNA substrate [1 μl of RNA, 1.25 mM ATP, 10 mM creatine phosphate, 8 mM MgCl2, 50 mM HEPES-KOH, pH 7.3, 6.5% polyvinyl alcohol (PVA)] and incubated at 30°C for 0, 1 and 2 hours (for β-globin substrate) or 0, 15, 30 and 60 min. (for MINX substrate). At each time point, the reaction was stopped by adding 180 μl of splicing stop solution [0.3 M sodium acetate, pH 5.2, 0.1% (w/v) sodium dodecyl sulfate (SDS), 62.5 μg/mL tRNA (Sigma, cat. no. R-9001)] and kept at 4°C until all samples were ready. RNA was extracted with equal volume of water-saturated phenol (AppliChem; cat.no. A1624) and centrifuged for 10 min. at 12 000 x g at 4°C. Do not use chloroform at any point of the RNA extraction as it forms a very large interphase with PVA. After centrifugation, supernatant was transferred to a fresh tube and 1 μl of RNA grade glycogen (20 μg/μL, Thermo Scientific, cat.no. R0551) as well as 100% ethanol (2.5x the volume of supernatant) was added and RNA was precipitated for at least 1 hour at −20°C. The tubes were centrifuged at 14'000 x g for 30 min. at 4°C and RNA was resuspended in 15 μl of 2x RNA loading buffer (90% formamide, 0.02% SDS, 1 mM EDTA, 0.02% bromophenol blue, 0.02% xylene cyanol). 10 μl was loaded on 8% Urea-PAGE gel and run at 12W for 1 hour. ^{32}P-labeled RiboRuler Low Range RNA Ladder (Thermo Scientific, cat. no. SM1831) was loaded as a

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

molecular-weight size marker. Gel was dried and exposed with phosphor screen BAS (GE Healthcare). The phosphor screen was scanned in a Typhoon FLA 9500 laser scanner (GE Healthcare) at 700V and 100 μm pixel size.

Histology of mouse tissue sections

Adult males (post-natal day 60, P60) were euthanized using CO2 and testes were isolated. Pictures of freshly isolated testes were taken with SteREO Discovery V12 (Zeiss) (Figure 6C). To prepare the paraffin sections, isolated testes were fixed in Bouin's solution (Sigma, cat. no. HT10132) for 48h at 4°C. Next, testes were washed in PBS for 48 h at 4°C, with frequent PBS changes. Samples were transferred into the embedding cassettes (Simport; cat. no. M508-3) and sent to the histology platform of University of Geneva. The samples were dehydrated in 70% (2 × 2h), 90% (1h), 95% (1h) and 100% ethanol (3 × 30 min) followed by incubation (3 times for 30 min) in HistoSAV2 (Biosystems). The solution was removed, replaced with paraffin, and incubated at 56-58°C. Testes were then transferred into plastic molds (Peel-A-Way® Embedding Mold S22, Polysciences, ; cat. no. 18646A-1) filled with paraffin and left at room temperature for paraffin to solidify. The sections (5 μm thickness) were cut using microtome (Leica RM2135) and mounted on the Superfrost Plus microscope slides (Thermo Fisher; cat no. 4951PLUS4). The sections were allowed to stretch for 24 h at 42°C and then were stored at room temperature.

Next, sections were stained using Hematoxylin and Eosin (H&E) stain protocol. The slides containing the paraffin sections were placed in a glass slide holder filled with HistoSAV2 (3 × 5 min) to remove the paraffin. For rehydration, the slides were incubated in 3x 100% ethanol, 96% ethanol, 70% ethanol, 50% ethanol and water (3 min for each step). Sections were stained with Hematoxylin solution (Merck) for 3-5 min and rinsed in running tap water. Then, sections were stained with Eosin Y solution (Sigma Aldrich; cat. no. E4382) for 3 to 5 min and washed with water. For dehydration, the sections were incubated in 50% (30 s), 70% (30 s), 96% (30 s), 100% ethanol (2 min) and HistoSAV (3 × 3 min). Neo-Mount (Merck) was put on the sections and immediately covered with coverslips. Pictures were taken using microscope AXIO Imager M2 (Zeiss).

Protein extraction from mouse tissues

Adult (P60) wild-type C57BL/6J male and female were euthanized using CO2 and various different tissues were isolated. Tissues were washed with PBS and immediately flash-frozen in liquid nitrogen. For lysate preparation, a piece the frozen tissue was cut out on a metal block placed on dry ice. The tissue piece was homogenized in 1 mL lysis buffer [50 mM Tris pH 7.4, 150 mM NaCl, 0.5% Triton X-100, 0.5% sodium deoxycholate, 1 mM DTT, Complete Protease Inhibitor Cocktail Tablet (Roche, Cat. No. 5056489001)]. The lysate was transferred to a 1.5 mL Eppendorf tube, centrifuged at 14000 x g for 30 min, and the supernatant was collected. Protein concentration was measured using the detergent-compatible colorimetric assay using a kit DC Protein Assay (Bio-Rad, 5000112). The reaction is similar to Lowry assay. The lysate concentration was normalized to 1 mg/mL using lysis buffer. Protein extracts were stored at −80°C.

Protein extraction from worms

Worms from three 6 cm plates containing synchronized adult population were washed off with M9 buffer. Worm pellets were washed additional 2 times, the excess of liquid was discarded and packed worm pellets were resuspended in lysis buffer (8 mM Na2HPO4, 2 mM KH2PO4, 137 mM NaCl, 100 mM KCl-1mM MgCl2, 1 mM EGTA, 10% Glycerol, 1% CHAPS, PMSF), 3x the volume of the worm pellet. The samples were immediately snap-frozen in liquid nitrogen and stored at −80°C. To obtain worm lysates, the samples were sonicated (10x 30 s ON, and 30 s OFF) with occasional snap-freezing in liquid nitrogen to break the worm cuticle. The lysates were clarified (20min spin 21'000 g at 4°C) and supernatant was transferred to fresh tubes. Protein concentration was measured immidiatelly after the protein extraction, using DC Protein Assay (Bio-Rad, 5000112). Protein lysate was mixed with 5X Laemmli Sample Buffer and water to a final concentration of μg/μl protein, boiled for 5 minutes at 95°C and frozen. Samples were stored at −80°C prior to the analysis.

Western Blot

Mouse whole tissue lysates (30 μg/well) or worm lysates were separated on SDS-PAGE gels prepared using Ultra-Pure ProtoGel 30% acrylamide (37.5:1) (National Diagnostic; cat. no. EC-890) mixed with ultra-pure water and resolving gel buffer, to obtain 12% resolving gel (0.375 M Tris, 0.1% SDS, pH 8.8), or with stacking gel buffer to obtain 8% stacking gel (0.125 M Tris, 0.1% SDS, pH 6.8). N,N,N',N'-Tetramethylethylendiamin (Merck, cat. no. 1107320100). The gel was polymerized by addition of 10% ammonium persulfate (AppliChem, cat.no. A1142). Gel electrophoresis was performed at 90 V for 30 min. and then at 120 V for 90 min. After separation, proteins were blotted on the Amersham Protran 0.45 mm nitrocellulose membrane (GE Healthcare; cat. no. 10600002) overnight at 5 V at room temperature using Trans-Blot SD. Semi-Dry Transfer Cell system (Bio-Rad; cat. no. 1703940). After transfer, membranes were washed with Tris-buffered saline (TBS, 20 mM Tris, 150 mM NaCl, pH 7.6) and blocked for 30 min. at room temperature with 5% dry milk in TBS with 0.05% Tween20 (TTBS) (SIGMA; cat. no. P7949). After 30 min. membranes were incubated with primary antibody: 1:500 anti-METTL16 (abcam; ab186012) or 1:100 anti-PARK7 (Invitrogen, PA5-13404) for 1 h at RT in 5% milk with TTBS. Membranes were then washed 5 times for 5 minutes with TTBS and incubated with HRP-conjugated anti-rabbit IgG HRP-linked (GE Healthcare; NA934) secondary antibody at 1:10,000 dilution for 1 h at RT in 5% milk in TTBS. After 1 h, membranes were washed 5 times for 5 minutes with TTBS followed by 3 washes for 5 minutes with TBS and incubated with one of the detection reagents: Amersham Prime Western Blotting Detection Reagent (GE Healthcare; RPN2232), SuperSignal West Femto Maximum

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

Sensitivity Substrate (Thermo Fisher; cat. no. 34095) or Pierce ECL 2 Substrate (Thermo Fisher; cat. no. 1896433A) for 5 min. at room temperature. The chemiluminiscence signal was detected using Amersham Hyperfilm ECL (GE Healthcare; cat. no. 28906837). The processed films were scanned using Perfection 3200 Photo scanner (Epson) with XSane image scanning software (ver. 0999).

Preparation of RNA libraries

m^{6}A-IP-seq to map m^{6}A transcriptome-wide in mouse, worm and insects

To compare the extent of m^{6}A RNA methylation between the species and map their location in the respective transcriptomes, we carried out m^{6}A-IP-seq using pre-mixed RNAs from the different organisms. To this end, total or poly(A)$^+$ RNA was isolated from the adult worm (*C. elegans*) grown on nutrient-high media (peptone-rich media+ NA22 *E. coli* strain, Table S4), adult mouse testis (P30) and Silkworm (*Bombyx mori*) BmN4-SID1 insect cell line. The RNAs were pre-mixed before further processing for RNA fragmentation and m^{6}A-IP-seq. In this mixed sample, the mouse RNA serves as an internal control for efficient m^{6}A immunoprecipitation via the unambiguous detection of m^{6}A peaks that are already reported (Dominissini et al., 2012; Ke et al., 2015; Wojtas et al., 2017).

Poly(A)$^+$ transcripts were purified from 75 μg of total RNA using the Dynabeads mRNA purification kit (Life Technologies; cat. no 61006). For total RNA fragmentation, 5 μg of total RNA each from mouse testis, adult worms and BmN4-SID1 cells was mixed with 2 μl of fragmentation reagent (AM8740, Thermo Fisher Scientific) in a final volume of 20 μl in a PCR tube. The reaction mix was incubated at 75°C for 12 minutes in a PCR machine. The tube was then transferred on ice immediately, and the reaction was stopped by adding 2.2 μl of stop solution provided with fragmentation reagent. Similarly, for poly(A)$^+$ RNA fragmentation, 2 μg of poly(A)$^+$ selected RNA each from mouse testis, adult worms and BmN4-SID1 was fragmented with fragmentation reagent as above. Denaturing urea-PAGE confirmed that majority of the RNA fragments were in the size range of 20-80 nts. A small portion (10%) of fragmented RNA from each sample was kept aside as input, while the remainder was subjected to immunoprecipitation.

The m^{6}A immunoprecipitation was performed as described (Ke et al., 2015). Briefly, Protein A Dynabeads were washed once in PXL buffer (1 × PBS, 0.1% SDS, 0.5% sodium deoxycholate, 0.5% NP-40) followed by pre-treatment with BSA (final concentration 1 μg/ μL) in 200 μL PXL buffer for 45 minutes at RT. BSA pre-treated beads was then conjugated with m^{6}A rabbit polyclonal antibody (20 μg; Synaptic Systems, catalog no. 202003) in 200 mL PXL buffer supplemented with 4 mL of RNasin RNase inhibitor (Promega; N2611) for one hour at RT on a rotating wheel. Dynabeads were further washed twice with PXL buffer, and finally, beads were resuspended in 400 mL of PXL buffer and 5 mL of RNasin. Fragmented RNA was added to the beads and incubated 4°C for 2 hours on a rotating wheel. After two hours incubation, the beads were washed twice by ice-cold Nelson low-salt buffer (15 mM Tris at pH 7.5, 5 mM EDTA), once by ice-cold Nelson high-salt buffer (15 mM Tris at pH 7.5, 5 mM EDTA, 2.5 mM EGTA, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 1 M NaCl), once by ice-cold Nelson stringent wash buffer (15 mM Tris at pH 7.5, 5 mM EDTA, 2.5 mM EGTA, 1% Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 120 mM NaCl, 25 mM KCl), and last by ice-cold NT-2 buffer (50 mM Tris at pH 7.4, 150 mM NaCl, 1 mM MgCl2, 0.05% NP-40). Antibody bound RNAs were eluted by incubating the beads with 0.5 mg/mL N^6-methyl adenosine (Sigma-Aldrich; M2780) in NT2 buffer for one hour at 4°C. The eluted RNAs were precipitated with ethanol and glycogen and dissolved in RNase-free water. The input and IP RNAs were first 3′ end dephosphorylated with T4 PNK (NEB; M0201S, 10 U/mL) in the absence of ATP at 37°C for 45 minutes (40 mL reaction: 35.5 mL RNA, 4 mL 10X T4 PNK buffer, 0.5 mL of T4 PNK) followed by phosphorylation of 5′ end (50 mL reaction: 40 mL dephosphorylated RNA, 6.5 mL water, 1 mL RNasin, 0.5 mL 100 mM ATP, 1 mL 10X T4 PNK buffer 1 mL T4 PNK) at 37°C for 45 minutes. RNAs were phenol chloroform-extracted, ethanol precipitated and resuspended in 6 mL of RNase free water. The input RNA fragments and the immunopurified RNAs after the phosphorylation step were directly used for strand-specific library preparation (barcoded at 3′ end) using NEBNext® Multiplex Small RNA Library Prep Set for Illumina® (NEB; catalog No. E7560L) following manufacturer's instructions. The libraries were resolved on 3% high-resolution MethaPhor agarose (Lonza; catalog. No. 50180) gels in 1X TAE buffer at 70 V. Fragments in the size-range of ~150-250 bp were gel-extracted with the use of MinElute Gel Extraction Kit (QIAGEN; cat No. 28604). Multiple libraries with different barcodes (at 3′ end) were mixed in equimolar ratios and sequenced with the NextSeq Illumina® Platform (EMBL Gene Core facility, Heidelberg). The maximum sequencing length was 75 nucleotides. The list of sequencing libraries generated is provided in Table S1.

m^{6}A-IP-seq to compare m^{6}A levels in WT and mett-10 worms

For identification of m^{6}A targets of METT-10 (Figures 1F and 1G), total RNA from biological triplicates of adult wild-type and *mett-10* KO worms, grown on nutrient-high media (Table S4), was isolated using TRIzol (Thermo Fisher Scientific) according to manufacturer instructions. Poly(A)$^+$ transcripts were purified from 75 μg of total RNA using the Dynabeads mRNA purification kit (Life Technologies; cat. no 61006). The RNA fragmentation and m^{6}A-IP-seq protocol, as described above, was followed for total RNA from WT, and poly(A)$^+$ RNA from both WT and KO (in biological triplicates). For comparison between WT and the *mett-10* KO was made under different diet conditions (Figure 4B), the animals were fed on either nutrient-high or nutrient-low media (Table S4). Triplicate biological replicates were processed for m^{6}A-IP-seq using poly(A)$^+$ RNA. Library preparation was as described above.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical methods are indicated in the figure legends.

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

CellPress
OPEN ACCESS

Analysis of m^{6}A-IP-seq to compare m^{6}A levels in mouse, worm and insects

The reads were sorted into individual libraries based on the barcodes and clipped using cutadapt (parameters: -a AGATCGGAAGAG CACACGTCT -m 15 -e 0.2 -O 4 -q 10–match-read-wildcards). The clipped reads were aligned to the mouse (GRCm38 – Ensembl release 95) or worm (WBcel235 – Ensembl release 95) genome using STAR (parameters:–outFilterType BySJout–limitOutSJcollapsed 50000000–limitIObufferSize 1500000000). *Bombyx* m^{6}A distribution was not analyzed. We detected similar ratio of m^{6}A/input reads for both species (Figure S1B). The m^{6}A peak calling was done separately for reads mapped to mouse or worm genome using MACS2 (macs2 callpeak -f BAM -q 0.01–nomodel–extsize 50–call-summits). Consensus peaks from the biological replicates were identified using MSPC (parameters: -r bio -w 1e-4 -s 1e-8). Unlike in the mouse, only very low number of m^{6}A peaks was identified in the worm. Peaks were annotated based on their overlap with the annotated features described in Ensembl gtf files (Figures S1C and S1D) and the enriched motifs were searched using MEME (parameters: -brief 50000 -nmotifs 5 -dna -revcomp -mod zoops -oc). Top motif is shown in Figures 1D and S1E. To investigate the m^{6}A distribution along the individual transcripts, the adaptor trimmed reads from poly(A)$^+$ libraries were mapped to individual ENSEMBL mRNAs using bowtie (parameters: -v 0 -a–best –strata). The read counts were divided by number of transcripts they mapped to and coverage was calculated using IRanges:: coverage function along the transcripts. We focused only on transcripts from genes which showed significant (adjusted p value $\leq$ 0.1) increase in m^{6}A/input ratio. These were identified using mapping the reads using SALMON (parameters: -l A -p 10–gcBias–validateMappings) followed by DESeq2 analysis. For the transcripts of these genes metaplots were created comparing the coverage along the 5′ UTRs, CDS and 3′ UTRs longer than 100 nt. Each part was divided into 100 pieces for which the mean coverage was calculated. In the mouse we observed expected m^{6}A enrichment at the start of 5′ UTRs (due to m^{6}Am) and at the end of CDS (Figure 1E). The worm distribution was rather uniform. The mean m^{6}A enrichment was also plotted for 0.5 kb vicinity of the STOP codon, which also showed m^{6}A enrichment only for the mouse.

Analysis of m^{6}A-IP-seq comparing m^{6}A levels in WT and *mett-10* KO worms

The reads were sorted into individual libraries based on the barcodes and clipped using cutadapt (parameters: -a AGATCGGAAGAG CACACGTCT -m 15 -e 0.2 -O 4 -q 10–match-read-wildcards). The clipped reads were aligned to worm transcripts (WBcel235 – Ensembl release 95) using bowtie (parameters: -v 0 -a -k 10–allow-contain). The read counts were divided by number of transcripts they mapped to and the counts were summarized to gene levels. To find the genes whose transcripts lose m^{6}A methylation in the absence of METT-10 we looked for genes with significant (padj $\leq$ 0.1) decrease of m^{6}A-IP/input ratio in *mett-10* KO using DESeq2 likelihood ratio test (LRT) where we compared the full model (~input_or_m^{6}A_IP + genotype + input_or_m^{6}A_IP:genotype) to the reduced model (~input_or_m^{6}A_IP + genotype) (Figure 1G). We identified the transcripts of U6 snRNA genes and three highly similar SAM-synthetase genes (*sams-3*, *sams-4*, *sams-5*) as the main targets of METT-10. To identify the precise adenosine which is methylated in U6 snRNA transcripts, we mapped the reads to the worm consensus sequence of U6 snRNA (Figure S1G), which was obtained by BLAST of mouse *Rnu6* to worm genome, using bowtie (parameters: -v 0 -a -k 10–allow-contain) and plotted the normalized (reads per million - rpm) coverage which confirmed that the worm METT-10 recognizes the same motif as the human METTL16. The methylation was completely gone in the *mett-10* KO (Figure 1I). Plotting the read counts normalized to library sizes (rpm) showed that although the U6 snRNAs are not polyadenylated we still were able to obtain enough reads in the poly(A)$^+$ libraries – but less than in the total RNA library from WT (Figures 1H and S1I). Comparison of normalized (rpm) read counts in input samples discovered a bit higher U6 snRNA levels in the KO (Figure S1H). To find out whether the loss of U6 snRNA m^{6}A methylation in the KO affects general splicing, we aligned the reads to the genome (WBcel235 – Ensembl release 95) using STAR (parametres:–outFilterType BySJout –limitBAMsortRAM = 40000000000–outSAMattributes All) and used the STAR generated SJ.out.tab files to count the reads spanning the annotated and novel splice junctions. The read counts were normalized to library sizes (rpm). We did not observe any decrease in the input KO samples which would suggest the negative impact on splicing (Figure S1J).

Investigation of normalized (rpm) m^{6}A coverage along SAM synthetase genes (*sams-3, sams-4, sams-5*), which showed decreased m^{6}A/input ratio in the KO, revealed the m^{6}A peak in the WT which overlaps the exon-intron boundary and is completely gone in the *mett-10* KO (Figure S2A). Comparison of DESeq2 normalized read counts in input samples showed increased expression of these genes (Figures S1E and S2C). Stronger WT m^{6}A signal with clear peak summit was obtained from *sams-3* and *sams-4* which harbor identical sequence in this region. We used bowtie to align the reads specifically to this consensus sequence and plotted the normalized (rpm) coverage (Figure 2A), which identified the adenosine of the 3′ splice site to be methylated in the WT. Interestingly, when changing the plates on which the worms were grown from NA22 plates to OP50 plates, the methylation in the WT was strikingly decreased which was apparent from both the coverage along the exon-intron boundary and also the amount of reads mapping to the boundary (Figures 4B and S4A).

To visualize the read coverages along selected genomic loci, we calculated thee normalized read coverages (rpm) of STAR mapped reads for individual samples. Plotting of the normalized read coverage (rpm) of input samples along the *mett-10* locus demonstrated the loss of reads from the 5′ portion of *mett-10* in the KO (Figure S1F).

To investigate the effect of m^{6}A loss on *sams-3* expression we plotted the mean coverage of its exons and introns from individual input samples (Figure 2B). This showed the general increased coverage of the exons in the KO, together with decreased coverage of intron 2 which contains the 3′ splice site methylated in the WT. To specifically compare the individual isoforms of *sams-3*: canonically spliced protein coding (PC), alternatively spliced non-coding (AS) and non-coding intron retained isoform (IR), following criteria were used. The PC abundance was estimated based on spliced read counts spanning the second intron (chrIV: 5848949-5849317), AS

Please cite this article in press as: Mendel et al., Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing, Cell (2021), https://doi.org/10.1016/j.cell.2021.03.062

isoform was quantified based on the counts of spliced reads spanning alternative intron (chrIV: 5848949-5849224) and the IR variant abundance was calculated as mean coverage of the canonical intron minus 10 nucleotides from both sites. All counts were normalized to library sizes (rpm) (Figures 2C and S4B). Deep sequencing data generated during the study (Table S1) are deposited with the Gene Expression Omnibus (GEO: GSE146873).

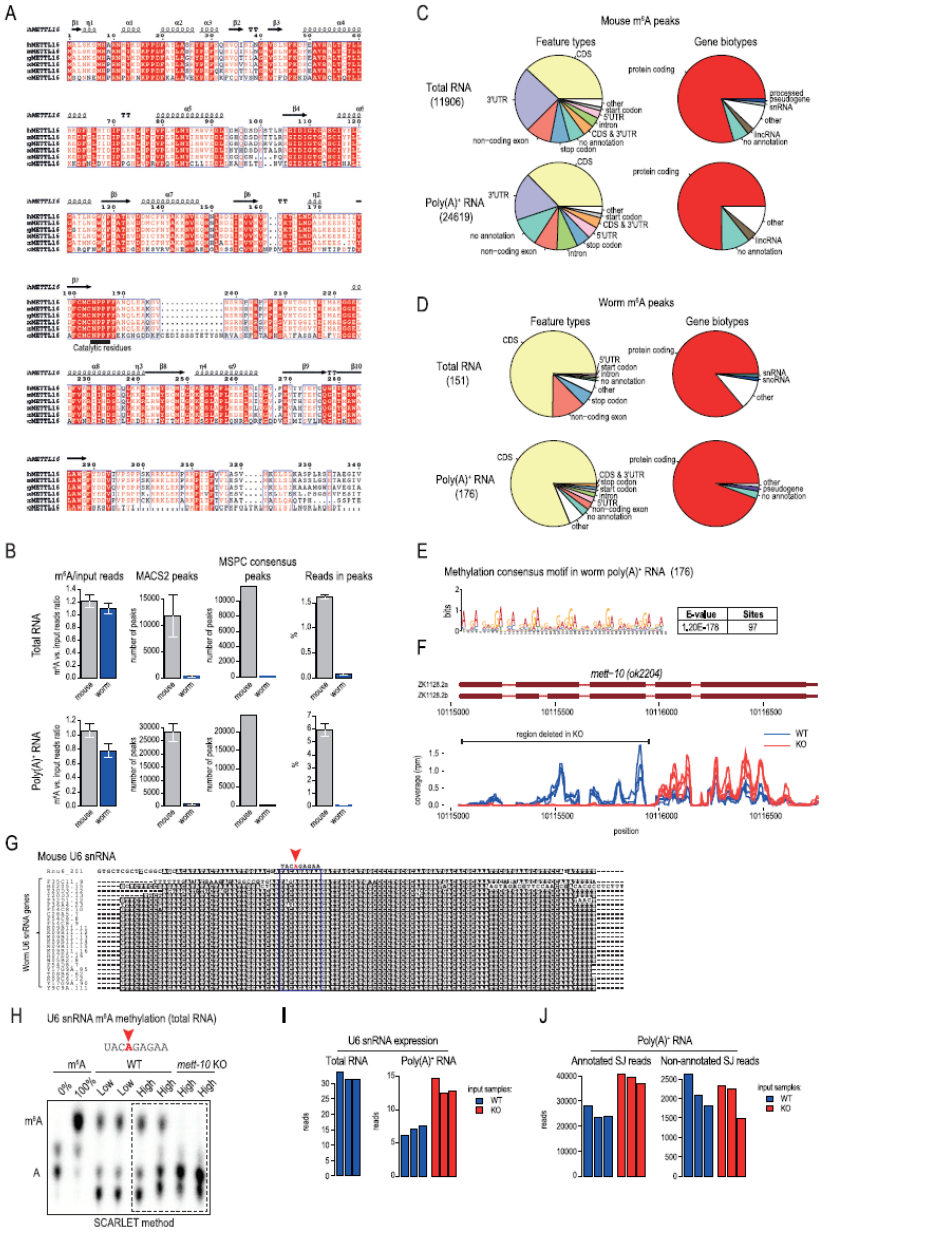
Search for mouse genes with METTL16 methylation motif at their 3′ splice site

From 403563 annotated mouse 3′ splice sites in the ENSEMBL database, 916 were found to overlap with one of the METTL16 methylation motifs (UAC**m^{6}AG**AGA or UAC**m^{6}AG**AAA). Next, we extracted a 41 nt sequence centered around the target adenosine and predicted the secondary structure taken by the sequences using RNAfold. We then ranked the sequences based on the similarity of their predicted structures to that taken up by the bonafide METTL16/METT-10 target hairpin in the worm *sams-3* (same sequence as in RNA-1 in Table S3). The similarity score was calculated as a sum of nucleotide positions represented as "(" or ")" in the dot-bracket notation of the secondary structure in both test sequences and in RNA-1 sequence. Several of the top-ranked sequences were used for *in vitro* methylation assay with recombinant human METTL16 (Figures 6E and 6F).

To see whether any of the METTL16 motif-containing 3′ splice sites display increased usage in *Mettl16*$^{-/-}$ mice, pointing toward their methylation dependent regulation, we searched our dataset (GEO: GSE116329) of E2.5 and E3.5 embryos (Mendel et al., 2018). We normalized the counts of splice junction reads spliced at the 3′ splice sites to gene expression levels, and searched for those 3′ splice sites with increased relative usage in *Mettl16*$^{-/-}$. Two such sites with increased usage in E3.5 embryos were identified which localize to chr17:84777132-84777139 (in *Lrpprc* gene) and chr19:40350011-40350018 (in *Sorbs1* gene) (Figure 6G).

CellPress
OPEN ACCESS

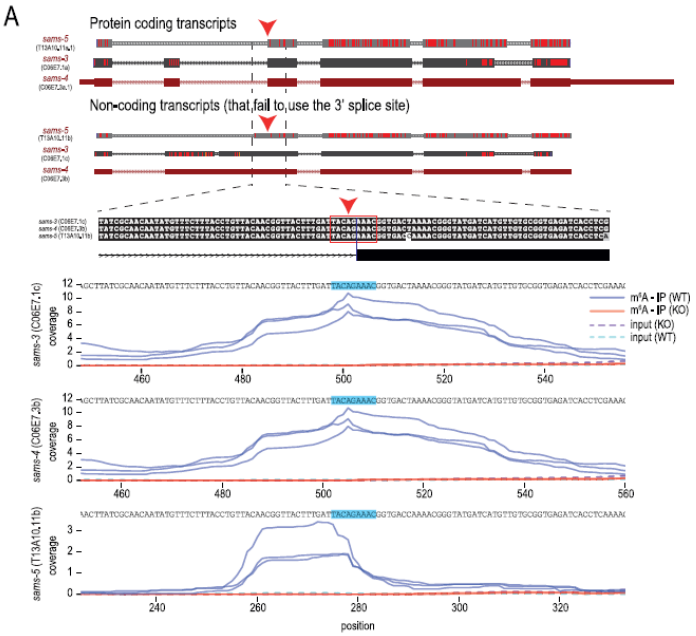
Supplemental figures

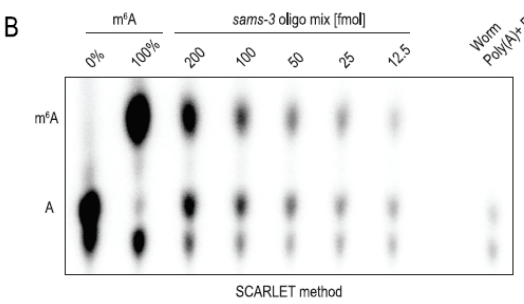


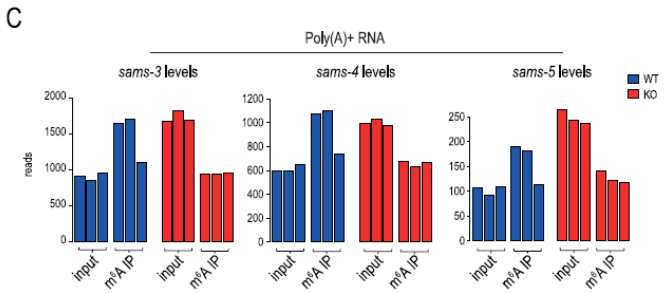
(legend on next page)

Figure S1. Distribution of m^{6}A in the worm transcriptome, related to Figure 1

(A) Protein sequence alignment of the methyltransferase domain of METTL16. h, *Homo sapiens* (NP_076991.3); m, *Mus musculus* (NP_080473.1); g, *Gallus gallus* (NP_001026773.1); x, *Xenopus laevis* (NP_001085334.1); z, *Danio rerio* (NP_001003611.1); c, *Caenorhabditis elegans* (NP_499247.2). Secondary structure features from the human METTL16 core methyltransferase domain (PDB: 6GT5) are indicated: α helices, β strands and η-3_{10} helix. (B) Equimolar amounts of total or poly(A)+ RNA from the adult mouse testes and adult worms (*C. elegans*) were pre-mixed together before performing m^{6}A-IP-seq. This allowed us to compare the m^{6}A distribution between the species. The worm and the mouse RNAs reveal a similar amount of m^{6}A-enriched sequences but only very low number of worm reads pile up as m^{6}A peaks. Mean values ± s.d. are plotted (n = 3). (C) Analysis of mouse m^{6}A peaks (peak counts are indicated within brackets). (D) Analysis of worm m^{6}A peaks (peak counts are indicated within brackets). (E) A consensus sequence identified in the small number (176) of m^{6}A peaks identified in worm poly(A)+ RNA. Its significance is not known. (F) RNA-seq analysis of wild-type (WT) and *mett-10* (*ok2204*) knockout (KO) mutant worms showing loss of RNA coverage from the 5′ end of the *mett-10* gene in the KO, consistent with the genomic deletion in the mutant. Biological replicas (n = 3) are plotted separately. (G) Multiple worm U6 snRNA transcripts were identified based on sequence homology to mouse *Rnu6*. The METTL16/METT-10 methylation consensus sequence and position of m^{6}A (red arrowhead) are indicated. (H) Detection of m^{6}A methylation in U6 snRNA from total RNA using the SCARLET method (STAR Methods). The method allows interrogation of site-specific methylation status (red arrowhead indicates the nucleotide position we examined). The thin-layer-chromatography (TLC) assay used in the protocol is shown. The total RNA is from wild-type (WT) or *mett-10* KO worms, grown on nutrient-high or nutrient-low plates. m^{6}A, refers to synthetic RNA oligos without (0%) or with (100%) m^{6}A (Table S3), used here as positive controls for the experiment (see STAR Methods). A part (dotted box) of this image is reproduced as Figure 1J. (I) The loss of U6 snRNA methylation in the *mett-10* KO results in slight increase of cellular U6 snRNA levels. Three input replicas are plotted separately for each tested genotype. (J) The loss of U6 snRNA methylation in the *mett-10* KO does not result in overall change in counts of reads covering splice junctions, therefore has no drastic effect on general splicing. Three input replicas are plotted separately for each genotype.







(legend on next page)

Figure S2. The worm m^{6}A writer METT-10 methylates the 3′ ss in an intron of the three *sams* homologs, related to Figure 2

(A) Three highly similar *sams* duplicated genes are present in the *C. elegans* genome. Splicing isoforms that differ in utilization of the methylated 3′ splice site (indicated with red arrowhead) are annotated in ENSEMBL. The cartoon shows the *sams-4* genomic locus and the *sams-3/-5* transcripts mapped to the *sams-4* locus using BLAT. All *sams-3/-4/-5* loci encode for the protein-coding isoform which uses the methylated 3′ splice site for splicing, but also non-coding variants where this splice site is not used. Coverage of m^{6}A along the intron-exon boundary identifies the methylated adenosine at the 3′ splice site of the three SAM synthetase homologs *sams-3*, *sams-4* and *sams-5*. The m^{6}A-IP-seq coverage has insufficient resolution to identify the methylated adenosine (red arrowhead) in *sams-5*. However, methylation is completely absent for all three homologs in the *mett-10* KO. Biological replicas (n = 3) are plotted separately. The METT-10 methylation consensus motif is highlighted. (B) Detection of m^{6}A methylation at the 3′ splice site in the *sams-3* pre-mRNA using poly(A)+ RNA from adult worms grown on a nutrient-high diet. The SCARLET method was used to specifically probe the 3′ splice site adenosine in intron2, but was undetectable. The thin-layer-chromatography (TLC) assay used in the protocol is shown. The same procedure was carried out with a dilution series of control synthetic RNA oligos (an equal mixture of oligos were the target adenosine is either methylated or unmethylated) mimicking the *sams-3* target sequence. Methylation of the oligos can be detected, but efficiency drops with a decrease in the amount of the oligos used. (C) Loss of 3′ splice site methylation results in increased expression of SAM synthetase (*sams*) genes. Compare input reads in WT versus *mett-10* KO from the m^{6}A-IP-seq experiment. Biological replicas (n = 3) are plotted separately. The input data from this plot is reproduced in Figure 2D.

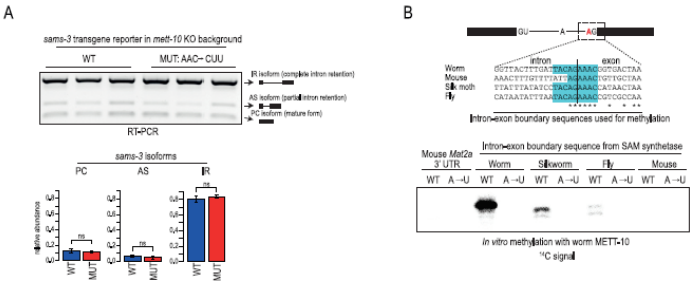
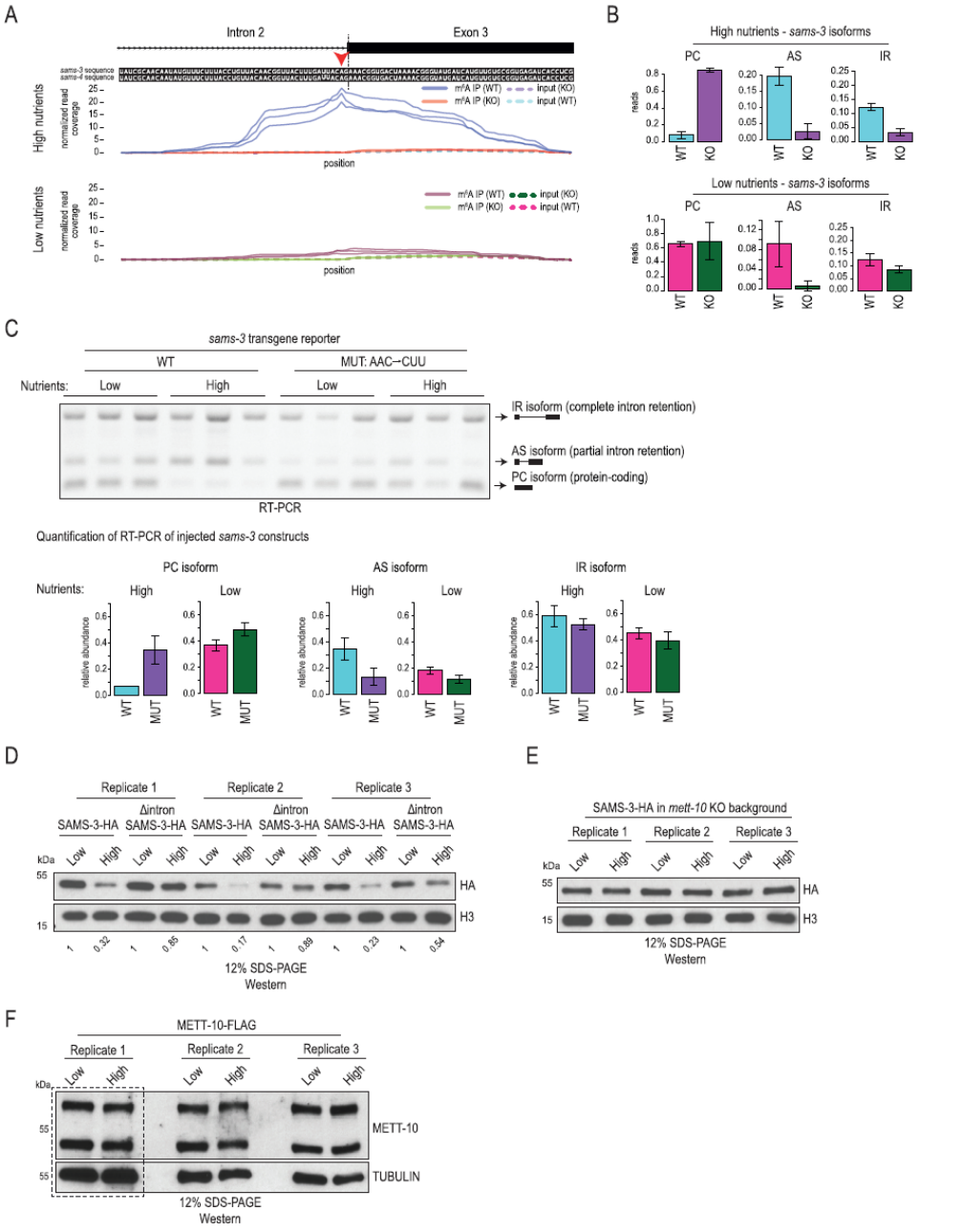


Figure S3. m^{6}A methylation of a specific 3′ ss in SAM synthetase pre-mRNA requires a stem-loop structure, related to Figure 3

(A) Wild-type (WT) transgene reporter constructs based on the *sams-3* gene were injected into *mett-10* KO worm gonads and multiple independent progeny lines stably expressing them were derived. Splicing patterns were analyzed by RT-PCR analysis using primers specific to the reporter. A mutated (MUT) construct with triple mutations (AAC →CUU) within the methylation consensus motif (in the exon3) was also tested. Lack of 3′ splice site m^{6}A methylation in the KO worms results in similar isoform levels from both WT and MUT constructs. Barplots depict mean relative proportion of individual isoforms ± s.d. (n = 3). See also Figure 3D. PC, protein-coding; AS, alternatively spliced; IR, intron-retained. (B) *In vitro* methylation assay using recombinant worm METT-10 protein and synthetic RNAs. The RNAs correspond to the intron-exon boundary of the SAM synthetase pre-mRNA from the indicated organisms, where the 3′ splice has the METT16/METT-10 methylation consensus motif. Note that the corresponding intron-exon boundary sequence in mouse *Mat2a* pre-mRNA has no consensus motif, unlike the confirmed METTL16 target site in its 3′ UTR. See also Figure 3F for the *in vitro* methylations with human METTL16. It appears that the worm METT-10 is inefficient on targets other than its own *sams* target site, while human METTL16 is active on all targets carrying the methylation consensus motif.



(legend on next page)

Figure S4. Diet-dependent change in m^{6}A RNA methylation of the 3′ ss of SAM synthetase pre-mRNA in *C. elegans* regulates SAMS protein levels, related to Figure 4

(A) The m^{6}A-IP-seq read coverages over the identical *sams-3* and *sams-4* intron 2/exon 3 boundary are shown. It shows a difference in methylation between the wild-type (WT) worms grown on nutrient-high and nutrient-low diets. Only the WT worms grown on nutrient-high plates show strong methylation of the 3′ splice site. When WT worms are grown on nutrient-low plates, the methylation is strikingly reduced. In *mett-10* KO the methylation is absent. (B) Quantification of the RNA-seq reads mapping to the various *sams-3* splice isoforms. PC, protein-coding isoform produced by correct use of the 3′ splice site in intron2; AS, alternatively spliced isoform due to use of an upstream 3′ cryptic splice site in intron2; IR, intron-retained isoform due to failure to use the 3′ splice site leading to intron2 being retained. All counts were normalized to library sizes (reads per million, rpm). Mean values are plotted ± s.d. (n = 3). Since there is almost no 3′ splice site methylation in WT worms grown on the nutrient-low plates, the removal of *mett-10* therefore has little effect on *sams-3* splicing (PC isoform). Consequently, both WT and *mett-10* KO worms (under nutrient-low conditions) use the site for splicing and produce predominantly the correctly spliced protein-coding (PC) version of *sams-3*, at levels comparable to KO worms grown on nutrient-high plates. (C) Transgene reporter constructs based on worm *sams-3* sequence were injected into wild-type worms to establish transgenic lines with stable expression. The constructs used had the wild-type (WT) *sams-3* sequence or had mutations (MUT: AAC→CUU) in the methylation consensus motif (on the part that sits on exon 3). This mutation is shown to abolish 3′ splice site m^{6}A methylation by recombinant worm METT-10 *in vitro* (Figure 3B). Three independent transgenic isolates expressing the constructs were used in the experiment. The worms were grown on either nutrient-high or nutrient-low plates. Splicing patterns were analyzed by RT-PCR with transgene-specific primers, and quantifications are shown below where mean relative proportions of individual isoforms are plotted ± s.d. A representative ethidium bromide-stained agarose gel showing the resolved cDNA products is shown. On nutrient-high plates, levels of the protein-coding (PC) isoform from WT construct is lower than that seen from the MUT construct, presumably due to 3′ splice site m^{6}A methylation in the former. The levels of the PC isoform from both constructs are similar in the nutrient-low plates, presumably, as the former is not methylated under these conditions. (D) Western analysis of SAMS-3-HA expressed from knock-in worm lines with or without intron2 in the *sams-3-HA* genomic locus. The worms were grown in nutrient-high or nutrient-low plates (Table S4). Three biological replicates are shown. Quantified HA signal normalized to that from endogenous histone H3 levels is shown below, with the value in nutrient-low diet being set to 1. Levels of SAMS-3-HA is reduced in nutrient-high diet condition, and this reduction is attenuated in the absence of intron2 in the *sams-3-HA* locus. The lysate from replicate #2 was re-run in the gel shown in Figure 4F. (E) Western analysis of SAMS-3-HA expressed from knock-in worm lines, in the *mett-10* background. Worms were grown on a nutrient-high or nutrient-low diet. Three biological replicates are shown. The lysate from replicate #2 was re-run in the gel shown in Figure 4F. (F) Western analysis of METT-10-FLAG expressed from knock-in worms grown in nutrient-high and nutrient-low plates. Three biological replicates are shown. Part of this image (replicate #1, dotted box) was reproduced in Figure 4F.

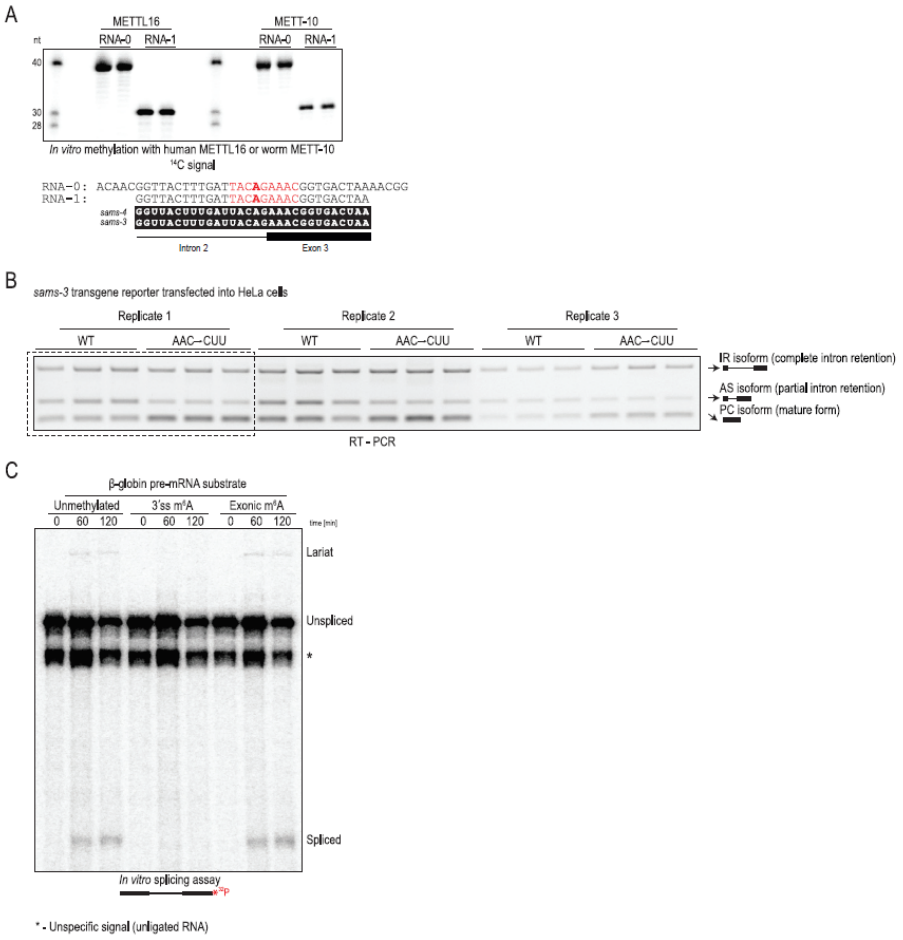
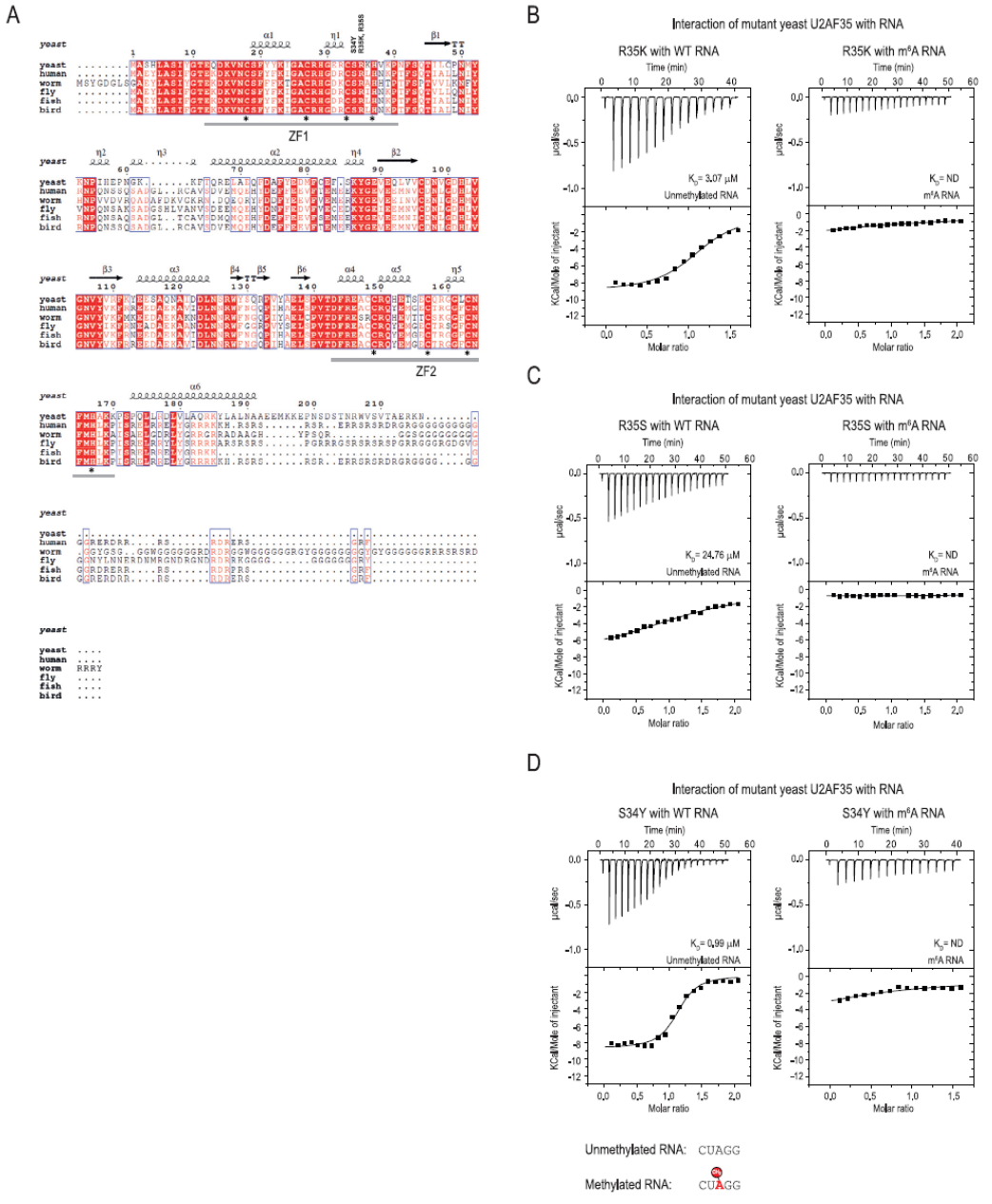


Figure S5. The 3′ ss m^{6}A methylation-mediated splicing inhibition is conserved in human cells, related to Figure 5

(A) *In vitro* methylation assay using recombinant human METTL16 or worm METT-10 proteins and radioactive ^{14}C-SAM as the methyl donor, using RNAs (two different lengths) corresponding to the intron 2-exon 3 boundary of the worm *sams-3* gene. The methylation consensus motif (red) and target adenosine (in bold) are shown. The reaction products were resolved by PAGE, the gel was stained with Methylene Blue to reveal the RNAs (to assure similar levels), and exposed to detect the radioactivity signal (^{14}C). The human METTL16 is able to recognize and methylate the worm *sams-3* target site, allowing us to test worm transgene reporter constructs in human HeLa cells. See also Figure 5A, and below. (B) RT-PCR analysis of the transcripts produced from worm *sams-3* transgene construct transfected into HeLa cells. Wild-type (WT) construct with the 3′ splice site which can be methylated by human METTL16 shows different splicing pattern when compared to the construct with mutations (MUT: AAC→CUU) in the methylation consensus motif (on the part that sits on exon 3). Compare ratios of alternatively spliced (AS) and correctly spliced protein-coding (PC) isoforms. Three biological replicates, each with three technical replicates, were used to quantify the individual isoforms and produce the barplot in Figure 5A. Part of this panel (replicate #1, dotted box) is reproduced in Figure 5A. (C) *In vitro* splicing assay shows that an artificially introduced 3′ splice site (3′ ss) m^{6}A within the human beta-globin pre-mRNA abolishes its splicing in human HeLa nuclear extracts, with neither the fully spliced product nor the lariat intermediate being detected. Presence of a single exonic m^{6}A has no effect on splicing. See also Figure 5B. A major RNA band (indicated with an asterisk) below the unspliced RNA substrate is an irrelevant non-ligated species leftover from production of the splint-ligated RNA substrate (see STAR Methods)



(legend on next page)

Figure S6. 3′ ss m^{6}A methylation blocks splicing by hindering its recognition by the U2AF35 splicing factor, Related to Figure 5

(A) Comparison of U2AF35 protein sequence among different species. The protein complex used for ITC experiments consists of the full-length U2AF23 (*S. pombe* U2AF35) and 93-161 aa of U2AF59 (*S. pombe* U2AF65). The Zinc Finger 1 (ZF1) and ZF2 domains in the yeast protein are highly similar to that in other organisms. The secondary structure features of *S. pombe* U2AF23 (PDB: 4YH8) is shown above the alignment: α helices, β strands and η-3_{10} helix. The asterisks at the bottom of the alignment indicate residues coordinating the zinc ion, while the residues we mutated are indicated on the top. (B-D) Isothermal calorimetry (ITC) experiments reveal that the yeast U2AF35 (in complex with a fragment of U2AF65) strongly binds an unmethylated RNA substrate (5′CU**AG**G) mimicking the 3′ splice site AG, while presence of an m^{6}A mark decreases the affinity. See Figure 5C. Two mutations of Arginine 35 that is involved in recognition of the splice site adenosine were made. A conservative mutation to a positively charged lysine (R35K) or to a non-conservative mutation to uncharged serine (R35S). The R35K mutation was made to see if the shorter side-chain of lysine could allow recognition of m^{6}A. We also made a mutation in the serine 34, which is frequently mutated to phenylalanine or tyrosine in human cancers, so we tested the S34Y mutant. Importantly, the two mutations replacing arginine 35 reduced binding to the unmethylated RNA, all three mutations did not bind to the methylated RNA.

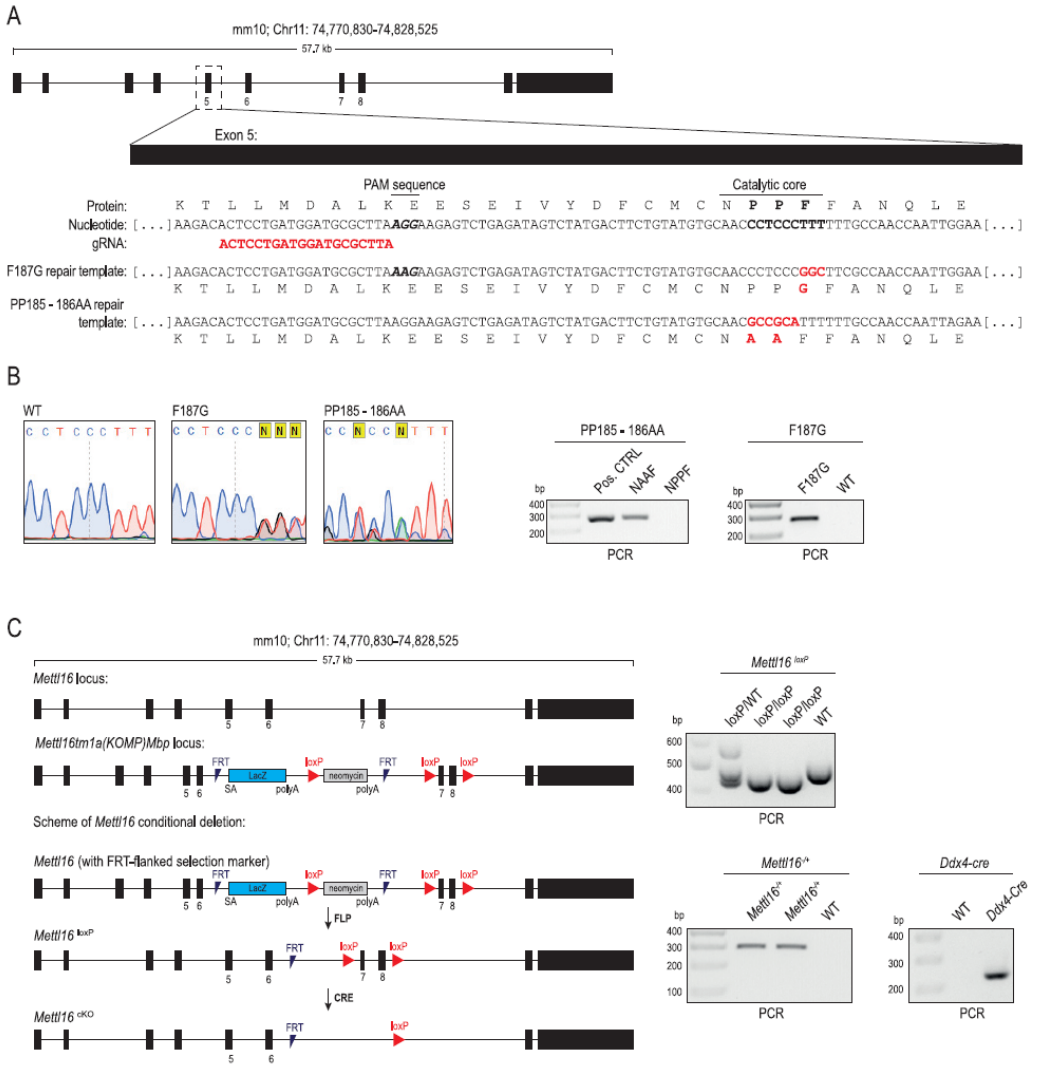


Figure S7. Creation of a mouse KI and cKO mutants for *Mettl16*, related to Figure 6

(A) Strategy for generation of the *Mettl16* knockin (KI) point mutant mice. A part of the genomic sequence of the *Mettl16* exon 5 and the predicted protein sequence encoded are shown. A single guide RNA (gRNA) targeting this region was used to guide Cas9 endonuclease activity and homology-mediated repair to introduce nucleic acid mutations that eventually result in the following amino acid changes: F187G (RNA binding mutant) and PP185-186AA (catalytic dead mutant). Sequence of part of the repair templates bringing the mutations (in red) are shown. (B) Examples of Sanger sequencing of genomic PCR to detect the WT, F187G and PP185-186AA *Mettl16* alleles (from mouse tail DNA). Representative ethidium bromide-stained agarose gels showing resolved PCR products is shown. Primer sequences are provided in Table S3. (C) Strategy for creation of the floxed *Mettl16* allele. The mouse line with floxed *Mettl16* allele and an inserted FRT-flanked selection markers cassette (LacZ and neomycin) was obtained from the KOMP repository at UC, Davis. Animals were crossed to remove the selection markers (STAR Methods). Using further crosses, we then brought together the floxed (*Mettl16* ^{loxP}) allele and the *Mettl16* null allele (*Mettl16* -). Crosses between *Mettl16* $^{loxP/-}$ and *Mettl16* $^{loxP/+}$; *vasa-Cre* partners gave us the *Mettl16* $^{loxP/-}$; *vasa-Cre* mice = conditional knockout (cKO) mutant. In the cKO, the gene is deleted in the male and female germline (starting from embryonic day E14.5 in the male germline). Representative ethidium bromide-stained agarose gels showing resolved PCR products detecting the different alleles and Cre driver is shown. Primer sequences are provided in Table S3.

Table S1. List of all deep-sequencing libraries created in this study. Related to STAR Methods and Figures 1-4.

Data is available from Gene Expression Omnibus (GEO: GSE146873).

sample	experiment	description	reads	sequencing type
RR723	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	INPUT-Total RNA-1	78748292	single-end
RR724	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	INPUT-Total RNA-2	109227705	single-end
RR725	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	INPUT-Total RNA-3	98509826	single-end
RR726	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	INPUT-PolyA+ RNA-1	111319107	single-end
RR727	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	INPUT-PolyA+ RNA-2	111369816	single-end
RR728	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	INPUT-PolyA+ RNA-3	104925272	single-end
RR729	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	m6A IP-Total RNA-1	97238279	single-end
RR730	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	m6A IP-Total RNA-2	116365247	single-end
RR731	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	m6A IP-Total RNA-3	103682866	single-end
RR732	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	m6A IP-PolyA+ RNA-1	99423158	single-end
RR733	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	m6A IP-PolyA+ RNA-2	90657183	single-end
RR734	m6A profiling on RNA mix from mouse, worm and BmN4-SID1	m6A IP-PolyA+ RNA-3	105299920	single-end
RR618	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_totalRNA_WT-1	107952378	pair-end
RR619	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_totalRNA_WT-2	104276535	pair-end
RR620	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_totalRNA_WT-3	99043875	pair-end
RR621	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_totalRNA_WT-1	88177640	pair-end
RR622	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_totalRNA_WT-2	92603501	pair-end
RR623	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_totalRNA_WT-3	80301018	pair-end
RR624	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_polyA_2_WT-1	51187060	pair-end
RR625	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_polyA_2_WT-2	52291052	pair-end
RR626	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_polyA_2_WT-3	33117898	pair-end
RR627	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_polyA_2_KO-1	39691391	pair-end
RR628	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_polyA_2_KO-2	35240657	pair-end
RR629	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_polyA_2_KO-3	39991008	pair-end
RR630	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_polyA_2_WT-1	12022240	pair-end
RR631	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_polyA_2_WT-2	36551759	pair-end
RR632	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_polyA_2_WT-3	25683360	pair-end
RR633	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_polyA_2_KO-1	54841480	pair-end
RR634	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_polyA_2_KO-2	44893258	pair-end
RR635	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_polyA_2_KO-3	43057847	pair-end
RR600	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_polyA_1_WT-1	22391643	pair-end
RR601	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_polyA_1_WT-2	23096755	pair-end
RR602	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	input_polyA_1_WT-3	19042425	pair-end
RR603	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_polyA_1_WT-1	18362889	pair-end
RR604	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_polyA_1_WT-2	18425770	pair-end
RR605	m6A profiling of the worm Mett-10 KO and WT on NA22 plates	m6A_IP_polyA_1_WT-3	18620691	pair-end
RR892	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	input_polyA_WT-1	44602975	single-end
RR893	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	input_polyA_WT-2	57428958	single-end
RR894	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	input_polyA_WT-3	63145019	single-end
RR895	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	input_polyA_KO-1	56220445	single-end
RR896	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	input_polyA_KO-2	46443279	single-end
RR897	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	input_polyA_KO-3	42188533	single-end
RR898	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	m6A_IP_polyA_WT-1	11536036	single-end
RR899	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	m6A_IP_polyA_WT-2	11339449	single-end
RR900	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	m6A_IP_polyA_WT-3	11065165	single-end
RR901	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	m6A_IP_polyA_KO-1	14331371	single-end
RR902	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	m6A_IP_polyA_KO-2	13228517	single-end
RR903	m6A profiling of the worm Mett-10 KO and WT on OP50 plates	m6A_IP_polyA_KO-3	12318564	single-end

Table S3. DNA primers and RNA oligonucleotides used in this study. Related to STAR Methods and Figures 1-6.

Name	Sequence	Name in the study/Comments
Worm mutants		
oKDMM42	GGTGATTGTTGTTCGAGTGC	Reverse primer for detecting and sequencing of intron 2 deletion in *sams-3*. If edit is present, it creates 473 bp with oKDMM43
oKDMM43	GTGACCAGGTAAGTAACCAACG	Forward primer for detecting and sequencing of intron 2 deletion in *sams-3*. If edit is present, it creates 473 bp with oKDMM42
oKDMM41	GACTCCAGCAGCGATTTCAG	Forward primer for detecting and sequencing of intron 2 deletion in *sams-4*. If edit is present, it creates 433 bp with oKDMM44
oKDMM44	ACGTTATCTCCGAAGCGTAG	Reverse primer for detecting and sequencing of intron 2 deletion in *sams-4*. If edit is present, it creates 433 bp with oKDMM41
oKDMM12	GCAACGTGTCGCTAAAGAGG	Forward primer for detecting and sequencing of intron 2 deletion in *sams-5*. If edit is present, it creates 526 bp with oKDMM39
oKDMM39	TCCGTCTGCTTACCTTTCGAG	Reverse primer for detecting and sequencing of intron 2 deletion in *sams-5*. If edit is present, it ceeates 526 bp with oKDMM12
oKDMM17	CCGTCATTCCCAAACTCCATC	Forward primer to detect and sequence C-terminal HA tag in *sams-3*. If edit is present, it creates 554 bp band with oKDMM17
oKDMM18	ATGAACCAACCGCTGAGAAC	Reverse primer to detect and sequence C-terminal HA tag in *sams-3*. If edit is present, it creates 554 bp band oKDMM17 to be sequenced
oFSa0131	TGTGGTGTTTGCACAGGTAG	Forward primer to detect and sequence deletion of *mett-10*. If edit is present, it creates around 260 bp band with oFSa0195
oFSa0195	ATGTCCATGCCAACATCCAGC	Reverse primer to detect and sequence deletion of *mett-10*. If edit is present, it creates around 260 bp band with oFSa0131
oKDMM5	AAACTTGATTGTACAATGTCTGAG	sgRNAs for C-terminal HA tagging of *sams-3*
oKDMM6	TCTTGTCAGACATTGTACAATCAA	
oKDMM7	CGATCTCAAACGGAAATGGAATTGCTCATGGAGGAGGATCAATGGGAGGATACCCATATGAT GTTCCAGATTACGCTGGCAGTTACCCATATGATGTTCCAGATTACGCTGGCAGTTAATGATT GTACAATGTCTGATTCATTTGTTG	Repair template for the C-terminal tagging of *sams-3*
oKDMM19	AAACTTTCTGTAATCAAAGTAACC	sgRNA for removing intron 2 of *sams-3, sams-4* and *sams-5*
oKDMM20	TCTTGGTTACTTTGATTACAGAAA	
oKDMM21	AAACTTGTAACAGGTAAAGAAACC	sgRNA for removing intron 2 of *sams-3, sams-4* and *sams-5*
oKDMM22	TCTTGGTTTCTTTACCTGTTACAA	
oKDMM23	AAACCCTGAATCACTCACCACAAGCAC	sgRNA for removing intron 2 of *sams-4*
oKDMM24	TCTTGTGCTTGTGGTGAGTGATTCAGG	
oKDMM25	AAACCAAGCAACCTTAGCGTGAGC	sgRNA for removing intron 2 of *sams-4*
oKDMM26	TCTTGCTCACGCTAAGGTTGCTTG	
oKDMM27	TCTTGCTCACGCAAAGGTTGCATG	sgRNA for removing intron 2 of *sams-3*
oKDMM28	AAACCATGCAACCTTTGCGTGAG	
oKDMM29	TCTTGAAAGGTTGCATGTGGTAAG	sgRNA for removing intron 2 of *sams-3*
oKDMM30	AAACCTTACCACATGCAACCTTTC	
oKDMM31	AAACACAGTTTCTGTAAAACTTAC	sgRNA for removing intron 2 of *sams-5*
oKDMM32	TCTTGTAAGTTTTACAGAAACTGT	
oKDMM33	AAACCAAGCAACCTTGGCGTGAGC	sgRNA for removing intron 2 of *sams-5*
oKDMM34	TCTTGCTCACGCCAAGGTTGCTTG	
oKDMM35	CGACAGCTTTCGAGGTGATCTCACCGCACAACATGATCATACCCGTTTTAGTCACtGTTTCA CAAGCAACCTTAGCGTGAGGGTCTTGAGCCAGATGAGCG	Repair template for removing intron 2 in *sams-4*
oKDMM36	CGCTGTTCTTGATGCTCATTTGGCTCAAGATCCTCACGCAAAGGTTGCATGTGAAACAGTGA CTAAAACGGGTATGATCATGTTGTGCGGTGAGATCACCTC	Repair template for removing intron 2 in *sams-3*
oKDMM37	CAGCTTTTGAGGTGATCTCACCGCACAACATGATCATACCCGTTTTGGTCACTGTTTCACAA GCAACCTTGGCGTGAGGATCCTGAGCCAAGTGAGCATC	Repair template for rem+A22:C38oving intron 2 in sams-5
sams-3 artificial constructs		
DH298	ACGAAGTGTTTCAACCTT	Forward primer for sams-3 FL amplification
DH299	CTCAATAAATGATTTTTTATAAG	Reverse primer for sams-3 FL amplification
DH300	CCCTAGAACACGGTCGTTCACATCAAGAACAGCGTCCGAGAT	Primer to introduce specific sequences into exon2
DH301	TGAACGACCGTGTTCTAGGGACGCAAAGGTTGCATGTG	Primer to introduce specific sequences into exon2
DH302	TACGCACTCAAAGTAGGCTGTCGGACGAGAACCTGGTAGTC	Primer to introduce specific sequences into exon3
DH303	ACAGCCTACTTTGAGTGCGTAATACGATGATTCCTCGAAAG	Primer to introduce specific sequences into exon3
DH304	CTTATAAAAAAATCATTTATTGAGCCTGCAGGACGCTGGCATGC	Forward primer for pUC19 fragment amplification
DH305	AAGGTTGAAACACTTCGTTGTTGTTCTGGAAATTGAGA	Reverse primer for pUC19 fragment amplification
MM320	GATTACAGACTTGGTGACTAAAACGGGTATG	Forward primer for METT-10 consensus motif mutation
MM321	TTAGTCACCAAGTCTGTAATCAAAGTAACCG	Reverse primer for METT-10 consensus motif mutation
HeLa splicing constructs		
MM415	ATGGCTAGCCACCACGAAGTGTTTCAACCTTACAAC	Forward primer for sams-3 FL clonning into phRL-TK vector
MM416	CATGCGGCCGCTCTAGAACTCAATAAATGATTTTTTATAAGGTC	Forward primer for sams-3 FL clonning into phRL-TK vector
Splicing		
MM378	GGTGAACGTGGATGAAGTTGGTGGTGAGGCCCTGGGCAGGTTGGTATCAAGGTTACAAGACA GGTTTAAGGAGACCAATAGAAACTGGGCATGTGGAGACAGAGAAGACTCTTGGCCTCGAGGT TTCTGATAGGCACTGACTCTCTCTGCCTATTGGTCTATTTTCCCACCCTTAGGCTGCTGGTG GTCTACCCTTGGACCCAGAGGTTCTTTGAGGATATCCGTACACCATCAGGGTACGAGCTAGC CCATGGCGTACACCATCAGGGTACGACTAGTAGATCTCGTACACCATCAGGGTACGGAATTC	Human B-globin with T7 promoter
MM379	TAATACGACTCACTATAGGGTGAACGTGGATGAAGTTG	Forward primer to amplify 3' fragment of B-globin
MM460	/52MOErG//i2MOErG/GAAAATAGACCAATAGGCAG	Reverse primer to amplify 3' fragment of B-globin
MM_RNA_59	/5phos/ACCCTTAGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAG	Unmethylated oligo for B-globin pre-mRNA ligation

MM_RNA_60	/5phos/ACCCTTm6AGGCTGCTGGTGGTCTACCCTTGGACCCAGAGGTTCTTTGAG	Splice site methylated oligo for B-globin pre-mRNA ligation
MM_RNA_61	/5phos/ACCCTTAGGCTGCTGGTGGTCTm6ACCCTTGGACCCAGAGGTTCTTTGAG	Exon methylated oligo for B-globin pre-mRNA ligation
MM453	CTGGGTCCAAGGGTAGACCACCAGCAGCCTAAGGGTGGGAAAATAGACCAATAGGCAGA	Splint oligo for B-globin ligation
MINX_FL	CCACTCTTGGATCGGAAACCCGTCGGCCTCCGAACGGTAAGAGCCTAGCATGTAGAACTGGT TACCTGCAGCCCAAGCTTGCTGCACGTCTAGGGCGCAGTAGTCCAGGGTTTCCTTGATGATG TCATACTTATCCTGTCCCTTTTTTTTCCACAGCTCGCGGTTGAGGACGACAAACTCTTCGCG GTCTTTCCAGTGGGG	MINX pre-mRNA sequence
MM_RNA_85	/5phos/cacagctcgcggttgaggacaaactcttcgcggtctttccagtgg	Unmethylated oligo for MINX pre-mRNA ligation
MM_RNA_86	/5phos/cacm6Agctcgcggttgaggacaaactcttcgcggtctttccagtgg	Splice site methylated oligo for MINX pre-mRNA ligation
MM_RNA_87	/5phos/cacagctcgcggttgagg/iN6Me-rA/caaactcttcgcggtctttccagtgg	Exon methylated oligo for MINX pre-mRNA ligation
MM578	gagtttgtcctcaaccgcgagctgtggaaaaaaaagg	Splint oligo for MINX ligation
MM617	TAATACGACTCACTATAGGGCCACTCTTGGATCGGAAACCC	Forward primer to amplify 3' fragment of MINX
MM582	/52MOErG//i2MOErA/aaaaaaaggggacaggaTaagtatgac	Reverse primer to amplify 3' fragment of MINX
MM583	/52MOErG//i2MOErT/gtgttaaaagggacaggaTaagtatg	Reverse primer to amplify 3' fragment of the weak polypyrimidine tract MINX
MM585	gagtttgtcctcaaccgcgagctgtggtgtgttaaaaggg	Splint oligo for the weak polypyrimidine tract MINX ligation
***sams-3* RT-PCR**		
MM395	CTCGGACGCTGTTCTTGATG	PCR for intron retention, endogenous sams-3
MM396	AACCTGGTAGTCGACGACAG	PCR for intron retention, endogenous sams-3
MM363	TGTGAACGACCGTGTTCTAGG	PCR for intron retention in mini-gene constructs
MM364	CGTATTACGCACTCAAAGTAGGC	PCR for intron retention in mini-gene constructs
***Mettl16* knock-in mice**		
tracerRNA		tracerRNA, 67 bases, cat no 1072533, IDT
CD.Cas9.PGDV2419.AB, RNA	CTCAAAGGAACCACTCAGCG GUUUUAGAGCUAUGCU	gRNA for cat-dead and RNA-binding mutant *Mettl16* mouse
MM233	TCATCTCCCTAGTGGTGAAAGTGCCGCAGAAGACACTCCTGATGGATGCGCTTAAAGAAGAG TCTGAGATAGTCTATGACTTCTGTATGTGCAACCCTCCCGGCTTCGCCAACCAATTGGAAGC CAAGGTACAGTACCTTTAAAGTTACCAAATGCTTAGCTAGCTTTTACATCATTGCAGAGATA TCAACAAAAGGCAG	ssDNA template for homology recombination for F187G modification
MM232	TCATCTCCCTAGTGGTGAAAGTGCCGCAGAAGACACTCCTGATGGATGCGCTTAAAGAAGAG TCTGAGATAGTCTATGATTTCTGTATGTGCAACGCCGCATTTTTTGCCAACCAATTGGAAGC CAAGGTACAGTACCTTTAAAGTTACCAAATGCTTAGCTAGCTTTTACATCATTGCAGAGATA AGAACAAAAGGCAG	ssDNA template for homology recombination for 185AA-PP186 modification
MM342	TCTCCATTGCTGCAGGTCTG	Forward primer for F187G genotyping
MM343	CTTCCAATTGGTTGGCGAAGC	Reverse primer for F187G genotyping
MM348	CGTCTCCATTGCTGCAGGTC	Forward primer for PP185-186AA genotyping
MM349	ATTGGTTGGCAAAAAATGCGGC	Reverse primer for PP185-186AA genotyping
MM340	AACTTGATGCGTCTCCATTG	Forward primer to amplify the whole catalytic site region for sequencing
MM341	TGGTAACTTTAAAGGTACTGTACC	Reverse primer to amplify the whole catalytic site region for sequencing
Mettl16 conditional KO		
MM101	aaagatttgccttggcgtttcc	Forward primer to detect loxP insertion in the *Mettl16* loxus
MM102	cccaggaagtggcactatcaag	Reverese primer to detect loxP insertion in the *Mettl16* locus
oIMR7643	CACGTGCAGCCGTTTAAGCCGCGT	Forward primer for *Ddx4-cre* genotoping
oIMR7644	TTCCCATTCTAAACAACACCCTGAA	Reverse primer for *Ddx4-cre* genotyping
MM314	AAATTCAAATGGGCATTGGG	Forward primer for *Mettl16 deletion genotyping*
MM315	CCGAATCCTCATCTATTTACATTG	Reverse primer for *Mettl16 deletion genotyping*
In vitro methylation		**Name in the study/Comments**
MM_RNA_5	CCCCCCCGUAGGCUACAGAGAAGGGGGGG	Mouse 3'UTR ; 29nt
MM_RNA_6	UGUUGGCGUAGGCUACAGAGAAGCCUUCA	Mouse 3'UTR mut ; 29nt
MM_RNA_31	ACAACGGUUACUUUGAUUACAGAAACGGUGACUAAAACGG	RNA-0; 40nt
MM_RNA_32	GGUUACUUUGAUUACAGAAACGGUGACUAA	RNA-1 ; 30nt
MM_RNA_34	GGUUACUUUGAUUACUGAAACGGUGACUAA	RNA-2 ; 30nt
MM_RNA_33	GGUUACUUUGGUUACAGAAACGGUGACUAA	RNA-3 ; 30nt
MM_RNA_35	GGUUACUUUGAUUACAGACUUGGUGACUAA	RNA-4 ; 30nt
MM_RNA_36	GGUUACUUUGAUUACAGUAACGGUGACUAA	RNA-5 ; 30nt
MM_RNA_37	GGUUACUUUGAUUACAGAGACUGUGACUAA	RNA-6 ; 30nt
MM_RNA_43	GGUUACUUUGAUUACAGAGACGGUGACUAA	RNA-7 ; 30nt
MM_RNA_38	CCCCCCCCUACAGAAACCCCCCCCC	RNA-8 ; 25nt
MM_RNA_39	CCCCCCCCUACAGAAACGGGGGGGG	RNA-9 ; 25nt
MM_RNA_100	GGUUACUUUGAUUACAGAAACUCAUUGGAA	RNA-10 ; 30nt
MM_RNA_101	UCAGUGGUUGAUUACAGAAACUCAUUGGAA	RNA-11 ; 30nt
MM_RNA_120	GGUUACUUUGAUUACAGAAACGGUGUGUAA	RNA-12 ; 30nt
MM_RNA_121	GCAUACUUUGAUUACAGAAACGGUGUGUAA	RNA-13 ; 30nt
MM_RNA_122	GGUUACUUUGAUUACAGAAACGCAGACUAA	RNA-14 ; 30nt
MM_RNA_123	GCAUUGUUUGAUUACAGAAACGCAGUGUAA	RNA-15 ; 30nt
MM_RNA_124	CCCCCCUUUGAUUACAGAAACGGGGGGGAA	RNA-16 ; 30nt
MM_RNA_44	AAACUUUGUUUUAUUAGAAACUGUUGCUAA	Mouse ; 30nt
MM_RNA_45	AAACUUUGUUUUAUUUGAAACUGUUGCUAA	Mouse mut ; 30nt
MM_RNA_46	UUAUUUAUAUCCUACAGAAACCAUAACUAA	Silkworm ; 30nt
MM_RNA_47	UUAUUUAUAUCCUACUGAAACCAUAACUAA	Silkworm mut ; 30nt
MM_RNA_48	CAUAAUAUUUAAUACAGAAACCGUCGCCAA	Fly ; 30nt
MM_RNA_49	CAUAAUAUUUAAUACUGAAACCGUCGCCAA	Fly mut ; 30nt
MM_RNA_67	UAUUUUUUUCUUUCUCUUACAGAAAUGAAAAAGAAGAUGAA	RNA-67/Mucl1 ; 40nt
MM_RNA_68	AGAGUGGUUUCUGUUUUUACAGAAAUGGAAACUGAAAAUAG	RNA-68/Spag1 ; 40nt
MM_RNA_69	AAUCUCUAUGUAUUUCUUACAGAAAAUGCAUCAGAAACGUG	RNA-69/Pla2r1 ; 40nt
MM_RNA_70	UUUCUGGAACUUAUUUUUACAGAAAAGUUUUAGAGACUGUC	RNA-70/Gyg ; 40nt
MM_RNA_71	GCAUGGGGUGGGUCUAUUACAGAAACCAGCCCCGAGCAAAG	RNA-71/Des ; 40nt
MM_RNA_72	AUGCAAAACUUUCUGUUUACAGAGAACCCCGGUUUCUCAUA	RNA-72/Helt ; 40nt
MM_RNA_73	ACUCUUGUCUUCUGUCUUACAGAGAACGGAAGGCUCCCUUA	RNA-73/Pecam ; 40nt
MM_RNA_74	AUUCUUUUUUUUUUCAUUACAGAGAAAGACAAAGCCUGAUC	RNA-74/Gm48535 ; 40nt
MM_RNA_75	CUUUUGGGCUUUUUCAUUACAGAGAAAUUUGGAAGCACAUG	RNA-75/Gm15810 ; 40nt
MM_RNA_76	GAUGCUUGUUUUGUGUUUACAGAGAAAAUGACAGGGUGGUC	RNA-76/Tjp2 ; 40nt
MM_RNA_80	UAGGAUUUUUUUUCUUCUACAGAAACGACUGGGAGCCCCCAGAUA	Sorbs1 ; 45nt
MM_RNA_81	UAAUUGUGUUUCAUUUUACAGAGAGCGUCUGAGCCAAGCC	Lrpprc ; 40nt

RP_RNA_19	GCGUCUUUACGGUGCUUAAAACAAAACAAAACAAAACAAA	40 nt RNA marker
RP_RNA_1	TGACATGAACACAGGTGCTCAGATAGCTTT	30 nt RNA marker
RP_RNA_3	UGACAUGAACACAGGUGCUCAGAUAGCU	28 nt RNA marker
RP_RNA_18	AGCACCGUAAAGACGC	16 nt RNA marker
ITC		
KM_RNA_10	CUAGG	ITC experiment; 5 nt
KM_RNA_11	CUm6AGG	ITC experiment; 5 nt
SCARLET		
MM437	GGAGAGACAACTTAAAGAGACTTAAAAGATTAATTTAAAATTTATCAAAAAGAGTATTGACT TAAAGTCTAACCTATAGGATACTTACAGCCATCGCCGGTCCTGTGAGTTAATAG	Universal DNA oligo for splint ligation
MM438	TGATCATACCCGTTTTAGTCACCGTTTCTCTATTAACTCACAGGACCGGCGATGGCTG	sams-3 splint oligo
MM698	mCmGmUmUmUmCmU GTAA mUmCmAmAmAmGmU	sams-3 chimera
MM614	CCTTGCGCAGGGGCCATGCTAATCTTCTCTCTATTAACTCACAGGACCGGCGATGGCTG	U6snRNA splint oligo
MM702	mCmUmUmCmUmCmU GTAT mUmGmUmUmCmC	U6 snRNA chimera
MM616	CCACTTTTTATATCTATAGCGCGCGTTGTCTATTAACTCACAGGACCGGCGATGGCTG	Positive control splint oligo
MM615	mGmCmGmUmUmGmU TCAG mAmGmUmCmUmAmA	Positive control chimera
MM_RNA_98	UUAGACUCUGAACAACGCGCGCUAUAGAUAUAAAAAGUGG	Positive control oligo - unmethylated
MM_RNA_99	UUAGACUCUGAm6ACAACGCGCGCUAUAGAUAUAAAAAGUGG	Positive control oligo - methylated
MM_RNA_96	GTTACAACGGTTACTTTGATTACAGAAACGGTGACTAAAACGGGT	sams-3 control oligo - unmethylated
MM_RNA_97	GTTACAACGGTTACTTTGATTACm6AGAAACGGTGACTAAAACGGGT	sams-3 control oligo - methylated

Table S4. Composition of nutrient-low and nutrient-high plates used for *C. elegans* culture. Related to Figure 1-4

Nutrient-low media

NGM (nematode growth medium)	**2 litres**
Agar	35 g
Peptone	5 g
Phosphate Buffer pH 6,0 + NaCl	50 ml phosphate buffer + NaCl
MgSO4 1 M	2 ml
CaCl2 1 M	2 ml
Cholesterol in ETOH (5 mg/ml)	2 ml
Buffer NGM Phosphate + Nacl 40X	**2 litres**
1 M Phosphate Buffer pH:6.0	72 g K2HPO4
	216 g KH2PO4
2 M NaCl	240 g NaCl

Nutrient-high media

Pepton-rich plates	**2 litres**
Agar	50 g
Peptone	40 g
Phosphate Buffer pH: 6,0 + NaCl	50 ml phosphate buffer + NaCl
MgSO4 1M	2 ml
Cholesterol in ETOH (5 mg/ml)	2 ml
Buffer Pepton Rich Phosphate + Nacl 40X	**2 litres**
1 M Phosphate Buffer	40 g K2HPO4
	240 g KH2PO4
820 mM NaCl	96 g NaCl

Table S5: List of worm strains used in this study. Related to Figure 1-4.

Strain	Genotype	Source	Comments
N2	wild type	CGC	wild type strain
VC1743	*mett-10 (ok2204) III*	CGC	*Mett-10* deletion
FAS198	*sams-3* (*uge125*) IV	This study	*sams-3* intron 2 deletion
FAS199	*sams-4* (*uge126*) IV	This study	*sams-4* intron 2 deletion
FAS200	*sams-3* (*uge127*) IV; *sams-4 (uge126)* IV	This study	*sams-3* intron 2 deletion and *sams-4* intron 2 deletion
FAS201	*sams-4* (*uge126*) IV; *sams-5 (uge128)* IV	This study	*sams-4* intron 2 deletion and *sams-5* intron 2 deletion
FAS202	*sams-3* (*uge129*) IV; *sams-4 (uge126)* IV; *sams-5 (uge128)* IV	This study	*sams3 i*ntron 2 deletion, *sams-4* intron 2 deletion and *sams-5* intron 2 deletion
FAS203	*sams-3* ((*uge130*) [*sams-3* ::HA]) IV	This study	HA-tagged *sams-3*
FAS204	*sams-3 (uge131)* IV; *sams-3* ((*uge130*) [*sams-3* ::HA]) IV	This study	HA-tagged *sams-3* with intron 2 deleted
FAS206	*sams-3* ((*uge130*) [*sams-3* ::HA]) IV; *sams-1 ((uge132)* [OLLAS::*sams-1*]) X; *mett-10 (ok2204)* III	This study	HA-tagged *sams-3* with *mett-10* deletion NOTE: This strain has *sams-1* tagged with OLLAS in the background, which was not used in this study

8. Discussion

The aim of my PhD work was to understand the biological role of m^6A methyltransferase METTL16 homologues in mice (METTL16) and *C. elegans* (METT-10) as well as to understand the structural basis of the RNA substrate recognition by METTL16. We discovered that surprisingly, despite millions of years of evolutionary difference, METTL16 and METT-10 have the same function in mouse (Chapter I) and worm (Chapter II) and methylate the same targets: U6 snRNA and transcripts encoding for SAM synthetases, *Mat2a* mRNA in mice and *sams-3,4,5* in *C. elegans.* In both cases, methylation of SAM synthetase transcripts regulates the stability of the transcript in response to changes in SAM levels, being essential for SAM homeostasis. However, the mechanism of this regulation is different. In mice, METTL16 stimulates intron splicing to stabilize *Mat2a*, while in *C. elegans,* METT-10 methylates the 3′SS, blocking splice site usage and inhibiting splicing. In both organisms, METTL16/METT-10 is important for proper development. In mice, it's essential for early embryonic development and male germ cells differentiation, while in worms, it controls germ cells proliferation.

Next, we obtained a crystal structure of the METTL16 methyltransferase domain and defined structural features crucial for RNA binding and methylation. Additionally, we explored RNA substrate requirements and showed that both METTL16 and METT-10 recognize specific consensus motif placed in the context of structured RNA.

Finally, we uncovered a new mechanism of splicing regulation through methylation of 3′SS, which blocks U2AF35 splicing factor binding and splice site recognition. Methylation of this site is dynamically regulated based on the *C. elegans* diet and is essential for SAM levels control. In mammals, SAM synthetase transcript splicing is regulated differently, but the mechanism of U2AF35 inhibition by m^6A is conserved.

In the following chapters, I would like to discuss the potential implications of these results in more detail.

8.1. METTL16 is essential for mouse embryonic development

Initially, METTL16 was shown to regulate SAM synthetase *MAT2A* mRNA levels in cell lines (Pendleton et al., 2017; Shima et al., 2017), but its role in mouse development was unknown. To understand the physiological role of METTL16 in mice, we generated the *Mettl16* KO mouse line and demonstrated that the lack of METTL16 leads to early embryonic lethality, between embryonic days 3.5 to 6.5 (Chapter I, Fig. S4D). Sequencing of E2.5 embryos showed only four transcripts downregulated, with a substantial decrease in *Mat2a* mRNA levels (Chapter I, Fig. 4). MAT2A is the only SAM synthetase in the developing embryo, and thus decrease in *Mat2a* mRNA levels most likely led to SAM deficiency. However, due to the low amount of material, we could not measure the levels themselves. SAM is the primary methyl donor for all methylation reactions, and thus, its deficiency would lead to widespread demethylation of RNA, DNA and proteins, including histones.

The reason for embryo lethality is most likely failure to regulate gene expression both at the level of epigenetic regulation, which is reintroduced during the embryo implantation phase (Reik et al., 2001), as well as on post-transcriptional and post-translational level. It was shown that removal of SAM for more than 24 hours leads to apoptosis of human pluripotent stem cells (Shiraki et al., 2014). Similarly, inhibition of MAT2A in bovine embryos (Ikeda et al., 2017) or inhibition of methylation with cycloleucine in pig embryos (Yu et al., 2021) leads to embryo death. Therefore, it is not surprising that *Mettl16* KO E3.5 embryos showed transcriptome-wide changes, making them unfit for further development. Interestingly, also *Mettl3* KO mice die at a similar time, around E5.5 to E7.5 (Geula et al., 2015).

Finally, *Mettl16* KO E2.5 embryos sequencing results, showing a strong decrease in *Mat2a* pre-mRNA levels combined with the last intron inclusion, confirmed previous cell culture experiments (Pendleton et al., 2017) and brought a better understanding of the METTL16 mechanism of action in mice.

8.1.1. Regulation of *Mat2a* pre-mRNA in mouse

There are six conserved hairpins, named hp1 to hp6, recognized and methylated by METTL16 in the 3′UTR of *MAT2A* pre-mRNA (Pendleton et al., 2017). METTL16 was proposed to regulate *MAT2A* pre-mRNA levels both by retained intron splicing stimulation, through the VCR domains (Pendleton et al., 2017), and transcript stability, through m^6A methylation of the hairpin structures recognised by YTHDC1, which leads to transcript degradation (Shima et al., 2017). In high SAM conditions, *MAT2A* pre-mRNA would be methylated and degraded through the interaction with YTHDC1, while in low SAM conditions, METTL16 dwell time on the transcript would increase, leading to splicing stimulation (Pendleton et al., 2017; Shima et al., 2017). Is one of the modes of action more important than the other? Do they act synergistically?

Sequencing of *Mettl16* knock-out mouse embryos showed a significant decrease in *Mat2a* pre-mRNA levels and simultaneous increase in reads coming from the retained intron (Chapter I, Figure 4), suggesting that splicing stimulation by METTL16 has a more important role in the developing embryo than m^6A-mediated transcript degradation. Methylation of hp2 – hp6 might have a role in further fine-tuning of *MAT2A* mRNA levels; however, the extent of this regulation remains to be seen. In the original study, YTHDC1-depletion affected only the luciferase reporter levels, while it did not affect the endogenous *MAT2A* mRNA levels (Shima et al., 2017). What is more, YTHDC1 is believed to regulate splicing and alternative polyadenylation, but its role in RNA degradation has never been reported before. On the other hand, YTHDF proteins have a well-established function in transcript degradation, making them better candidates for drivers of *Mat2a*

mRNA degradation. However, there was also no effect of YTDF2-depletion on *MAT2A* mRNA levels. The recent reports that YTHDF proteins work synergistically (Lasman et al., 2020; Zaccara and Jaffrey, 2020) might explain why there was no effect, as YTHDF2-depletion might have been compensated by the presence of YTHDF1 or YTHDF3 proteins.

8.1.2. The catalytic activity of METTL16 is essential for mouse development

Analysis of *Mettl16* KO mouse embryos showed a decrease in splicing of the last intron of *Mat2a* pre-mRNA, which is believed to be mediated through METTL16 VCR domains (Pendleton et al., 2017). In addition, cell lines experiments showed that tethering of catalytically-dead METTL16 is sufficient for *MAT2A* splicing (Pendleton et al., 2017). We assumed that if *Mat2a* mRNA levels are mainly controlled at the splicing level, then the catalytically-dead (but RNA binding) mutants of METTL16 should still stimulate splicing of *Mat2a* pre-mRNA and result in viable animals. At the same time, they would allow to distinguish between VCR-dependent and methylation-dependent roles.

To explore that, we decided to generate mouse lines with mutations in the catalytic domain of METTL16. Two mutations, F187G and PP185/186AA, were selected. According to the published data, while both mutants are catalytically inactive, PP185/186AA retains RNA biding ability and stimulates splicing of MAT2A retained intron (although with lower efficiency than WT METTL16), while F187G does not bind RNA at all (Pendleton et al., 2017). If *Mettl16* KO lethality phenotype is solely due to the inability to regulate *Mat2a* splicing, PP185/186AA mouse mutants should be still viable, while F187G mutation should be lethal. Unfortunately, both mutations led to embryonic lethality (Chapter II, Fig. 6A). Analysis of single embryo sequencing (E2.5 and E3.5 embryos) of PP185/186AA animals showed decreased *Mat2a* mRNA levels; however, due to the low number of

reads, we could not see reads coming from the last intron region to assess whether PP185/186AA mutant stimulates splicing of *Mat2a* pre-mRNA.

There are several potential explanations why PP185/186AA mutant (catalytic dead, able to bind RNA) fails to stimulate *Mat2a* pre-mRNA splicing and results in the same phenotype as *Mettl16* KO. The first is that mutated protein might be unstable *in vivo,* which would be very hard to assess due to the low amount of material available in the E2.5 and E3.5 embryos. The second is that the lower efficiency of splicing stimulation initially reported for the PP185/186AA mutant (Pendleton et al., 2017) might not be enough to stabilize *Mat2a* pre-mRNA *in vivo*. Third, methylation of *Mat2a* pre-mRNA might be essential for its stabilisation, but that would be the opposite mechanism to the one previously reported. Finally, the methylation activity of METTL16 might be important for the regulation of other targets required for embryonic development, including U6 snRNA.

8.1.3. Additional targets of METTL16 in mouse

While sequencing of E2.5 embryos showed four targets (*Mettl16, Mat2a, Ccdc92b* and *Gm15698)* being downregulated (Chapter I, Fig. 4 and Fig.S5) (Mendel et al., 2018), we focused our attention on *Mat2a* mRNA deficiency as the reason for embryonic lethality. Although the most probable, our hypothesis does not include the potential influence of the two other downregulated genes: *Gm15698* and *Ccdc92b*. Both genes are expressed in the early embryos, with *Gm15698* expressed only in the developing embryo and *Ccdc92b* expressed in the early embryo, brain and developing liver (Tapial et al., 2017). Unfortunately, there is very little information about the genes' role, making their role in the developing embryo challenging to assess. Nevertheless, while *Gm15698 (ElobI)* lacks METTL16 binding motif and so that it's hard to explain a significant drop in expression levels in *Mettl16* KO embryos, *Ccdc92b* gene is much more interesting. It has a perfect METTL16 consensus motif in the 3'UTR (TACAGAGAA), and the whole region around the consensus motif forms a secondary

structure identical to *Mat2a* hp1 structure. Further studies are necessary to establish the link between METTL16 and both genes.

Furthermore, it is important to mention that METTL16 was described to bind (but not methylate) to ENE+A triple helix structures in *MALAT1* (Brown et al., 2016). The role of this interaction is unknown, but as *Malat1* KO mice are viable, its loss shouldn't contribute to the *Mettl16* KO mouse phenotype (Nakagawa et al., 2012; Zhang et al., 2012). However, as ENE elements protect viral or cellular non-coding RNAs from rapid degradation by sequestering the 3′ poly(A) tail inside the RNA secondary structure (Conrad et al., 2007), it is interesting whether METTL16 could also bind other cellular or viral ENE sequences and for example contribute to viral infection.

Finally, Warda et al. showed that METTL16 binds to over 400 different RNAs in HEK293T cells, including 355 mRNA, 68 lncRNAs and 9 ncRNAs (Warda et al., 2017). However, the overwhelming majority of the sites located in mRNA (90%) and lncRNAs (75%) do not overlap with known m^{6}A sites, making them questionable targets. As it was not shown whether these sites have a consensus motif or potential to form a secondary structure needed for METTL16 binding (Warda et al., 2017), their association with METTL16 remains to be experimentally validated. One potential explanation for the mismatch between binding and the lack of methylation might come from a recent study that showed that some targets could be bound, but still not methylated by METTL16 (Doxtader et al., 2018).

8.2. METTL16 function is conserved in *C. elegans*

My next goal was to understand the role of METTL16 homologue, METT-10 in *C. elegans*. As METT-10 has a well-conserved methyltransferase domain (Dorsett et al., 2009), we hypothesised that it might be m^{6}A methyltransferase in *C. elegans*. In Chapter II, I describe the analysis of m^{6}A levels in WT and *mett-10* KO worms, which showed that METT-10 methylates several transcripts,

mainly U6 snRNA and SAM synthetase transcripts: *sams-3*, *sams-4* and *sams-5* (Chapter II, Fig. 1). That makes METT-10 the first confirmed mRNA m^6A methyltransferase in *C. elegans*, with targets conserved between mouse and worm. However, unlike in mice, where there are six m^6A peaks in the 5′UTR of *Mat2a* pre-mRNA, in worms, the *sams* transcripts are methylated at a single site, precisely at 3′SS of intron 2 of *sams* transcripts (Chapter II, Fig. 2). This methylation inhibits splice site usage by blocking U2AF35 splicing factor binding and leads to alternative splice site usage, resulting in transcript degradation. The whole mechanism is regulated by diet, acting as a feedback loop to regulate SAM levels in different environmental conditions (Chapter II, Fig. 4). It is the first instance where the presence of m^6A is shown to directly impact splicing, which will be discussed in more detail later.

Although the function of METTL16 is conserved between mice and worms, there are important differences in the mode of action. Comparison of *Mettl16* and *mett-10* KO outcomes shows that in mouse METTL16 is needed for SAM synthetase transcript stabilization, while in *C. elegans* METT-10 is important for its degradation. The source of this difference is perhaps different environments in which worms and mice live. As production of SAM is energetically costly, requiring methionine and ATP, mammals might need to control it tightly. At the same time, worms might have enough nutrients to produce SAM constantly, and only too high levels are dangerous.

Another interesting difference is the mechanism of SAM synthetase transcript regulation. In mice, it was proposed that METTL16 binding promotes splicing through the recruitment of additional factors, which recently were identified to be CFI_m25 protein. In worms, splicing is regulated directly by the m^6A methylation of the 3′SS. The additional factors in mice might allow for further fine-tuning of splicing, which might not be necessary for worms. The second possibility is that worms do not have m^6A-binding proteins and thus, the only potential mode of m^6A action is to directly prevent binding of splicing factors.

8.2.1. METT-10 importance besides SAM level control

The previous studies showed that METT-10 is important for *C. elegans* germline development, with *mett-10* KO germ cells unable to progress through the mitotic cell cycle, resulting in sterility phenotype at 25°C (Dorsett et al., 2009). Our experiments showed decreased brood size in *mett-10* KO worms at 20°C compared to WT, with a very significant drop in nutrient-rich conditions (Chapter II, Figure 4G). This drop in brood size was correlated with the increased SAM levels as measured by metabolomics (Chapter II, Fig. 4E). We hypothesised that the phenotype of *mett-10* KO is caused by the inability to regulate SAM levels, which at some point become toxic for worms. To test that, we generated *C. elegans* mutants, where we removed regulated intron in *sams-3*, *sams-4* and *sams-5* (so-called Δintron worms), having single mutants or combining them to generate a triple Δintron *sams-3/-4/-5* mutant. Worms with Δintron produced SAM synthetases, but lost the ability to control *sams* intron splicing through METT-10 methylation, leading to increased SAM synthetase protein levels in nutrient-high conditions as shown by Western Blot (Chapter II, Fig. 4F). Surprisingly, despite upregulated SAM synthetase protein levels, we saw a very modest decrease in brood size in Δintron lines. The reduction was significant only in comparison between WT and the triple Δintron (*sams-3/-4/-5)* worms, with the effect being much smaller than between *mett-10* KO and WT worms. This result indicates that METT-10, besides the SAM synthetase transcripts, have additional targets and functions in *C. elegans*. What could be the other targets?

U6 snRNA m^6A methylation could be the first candidate, with m^6A in the middle of the crucial ACAGA motif, which was shown to be important for splicing in *S. pombe* (Ishigami et al., 2021). However, we have not seen any major splicing defects in *mett-10* KO worms (Chapter II, Fig. S1), so either U6 snRNA methylation is not critical in worms, or it is specific only to a subset of targets. Another potential reason for the *mett-10* KO phenotype severity might be additional targets of METT-10. Analysis of m^6A-IP-seq between *mett-10* KO and WT worms showed that while U6

snRNA and *sams-3/-4/-5* pre-mRNAs are the main targets of METT-10, there are over 171 other transcripts with significant depletion of m^6A in *mett-10* KO. Besides U6 snRNA and *sams* transcripts, the most m^6A-depleted transcript is *mtce.24* encoding for one of the mitochondrial tRNAs (Chapter II, Fig. 1). However, it is unclear whether *mtce.24* is a direct target of METT-10, as it lacks the consensus motif and the secondary structure needed for METT-10 methylation. In addition, the function of *mtce.24* in *C. elegans* remains unknown.

8.2.2. m^6A methylation in *C. elegans*

The presence of m^6A on worm mRNA was for a long time controversial as *C. elegans* lacks all the main proteins involved in the deposition or recognition of mRNA m^6A, including METTL3/14 methyltransferase. In addition, recent studies showed that in worms, mRNA m^6A is either very low abundant (0.0008% of *C. elegans* mRNA) (Liberman et al., 2020) or is not present at all (Sendinc et al., 2020). In Chapter II, we show that despite the low levels, mRNA in *C. elegans* is m^6A methylated, with METT-10 being the m^6A mRNA methyltransferase that methylates *sams-3/-4/-5* pre-mRNA transcripts. Comparison of *mett-10* KO and WT worms identified few additional targets in the *mett-10* KO; however, most of them lacked METT-10 consensus motif, so the drop in m^6A levels might be due to secondary effects. At the same time, our broader analysis of m^6A levels by LC-MS/MS showed that the amount of m^6A on poly(A)+ RNA is much higher than previously reported, being approximately five times lower in *C. elegans* when compared to mice (Chapter II, Fig. 1B). Sequencing of equimolar amounts of worm and mouse poly(A)+ RNA revealed similar amounts of m^6A, but with striking differences in distribution. While in mice, m^6A was distributed in well-defined peaks, in *C. elegans,* almost all m^6A seems to be distributed randomly (Chapter II, Fig, 1E). Only for a few sites were we able to define a long consensus sequence rich in purines (GA repeats). Interestingly, the random distribution and the very same motif were reported recently, where they were attributed to the unspecific binding of the m^6A antibody (Sendinc et al., 2020).

Despite the overall high m^6A levels in poly(A)-enriched fraction, the seemingly random distribution of m^6A and lack of defined consensus motif raises questions about the source of m^6A in worms. Is there another yet undiscovered m^6A mRNA methyltransferase? Is m^6A coming from bacteria in the *C. elegans* diet, either as sample contamination or randomly incorporated into pre-mRNA by RNA polymerase II? It was shown that bacteria also have m^6A methylation, distributed uniformly along the transcripts at GCCAG consensus motifs (Deng et al., 2015). As RNA polymerase II can incorporate m^6A into pre-mRNA, plants and mammals use the ADAL enzyme for m^6A removal and recycling (Chen et al., 2018). A homologue of mammalian ADAL, adal-1, is also present in worms (Harris et al., 2020), so it is unlikely that bacterial m^6A might be randomly incorporated into the pre-mRNA, pointing more toward sample contamination with the bacterial mRNA. Or maybe both the levels of m^6A, as well as seemingly random distribution, are caused by a combination of contamination with other RNA species and unspecific m^6A antibody binding, as suggested recently (Sendinc et al., 2020)? The presence of RNA modifications not found on mRNA, like specific for rRNA N^6,N^6-dimethyladenosine ($m^6{}_2A$), might suggest rRNA contamination, however it doesn't explain the GA repeats consensus. To summarize, more research is needed to understand better the presence, distribution and role of worm mRNA m^6A.

8.3. m^6A methylation of the 3′SS directly regulates splicing

Since the 1980s and 1990s, m^6A was implicated in splicing regulation, with the early studies showing that inhibition of methylation leads to accumulation of unspliced pre-mRNAs in chicken embryo fibroblasts (Stoltzfus and Dane, 1982) and in chinese hamster ovary (CHO) cells (Carroll et al., 1990). However, as the studies used general methylation inhibitors, the results were more likely caused by general RNA processing defects, including lack of cap methylation, which is needed for recognition by the cap binding complex (CBC) (Worch et al., 2005). More recently, there is growing evidence for the role of m^6A in splicing regulation, with multiple studies finding widespread changes

in splicing patterns following depletion of components involved in m^{6}A methylation: METTL3 writer (Dominissini et al., 2012; Geula et al., 2015), FTO (Bartosovic et al., 2017; Zhao et al., 2014) and ALKBH5 (Tang et al., 2017) erasers as well as YTHDC1 reader (Kasowitz et al., 2018).

Our initial observation of the correlation between 3′SS m^{6}A methylation and impaired splice site usage in *C. elegans* led us to a discovery of a previously unknown mechanism of splicing regulation, where m^{6}A methylation can directly block U2AF35 splicing factor binding to inhibit splicing (Chapter II). Using *in vitro* splicing assays, we showed that the exact mechanism is conserved in mammals and is limited to U2AF35-dependent splice sites (so-called weak splice sites) (Chapter II, Fig. 5). To identify mammalian splice sites potentially regulated by m^{6}A methylation, we performed *in silico* analysis of all 3′SS in mice to select those having METTL16 binding motif and forming secondary structures around the 3′SS (Chapter II, Fig. 6E). Next, we confirmed that some of these sites can be *in vitro* methylated by METTL16 (Chapter II, Fig. 6F). Finally, we showed that two potentially m^{6}A methylated splice sites have different splice site usage in *Mettl16* KO mouse embryos, suggesting that this regulation might be active during mouse embryonic development (Chapter II, Fig. 6G). To summarize, we discovered an evolutionarily conserved mechanism of splicing regulation where m^{6}A methylation of 3′SS blocks U2AF35 binding to inhibit splicing. In addition, the same mechanism was very recently reported by another group, further strengthening our conclusions (Watabe et al., 2021).

A recent crystal structure of U2AF35 bound to its target explains how m^{6}A blocks U2AF25 binding. Amino-acid residues of U2AF35 very closely surround the -2 adenosine (the one which is methylated by METT-10), thus there is no space to accommodate the methyl group. Isothermal titration calorimetry (ITC) measurement of binding affinity to non-methylated and m^{6}A methylated RNA confirmed our results, with over 100x times weaker biding for methylated RNA (K_d value being 0.47 μM and 53.8 μM, respectively) (Yoshida et al., 2020). However, this type of splicing inhibition

would work only for the U2AF35-dependent transcripts. Indeed, the m^6A methylated 3′SS in *C. elegans sams-3/-4/-5* transcripts have mutated consensus motif, with AUUACAG/R sequence instead of the typical U_4CAG/R. As U2AF65 was shown to bind to the U_4 consensus motif (Hollins et al., 2005), these mutations move the balance towards U2AF35-driven splicing.

It is worth mentioning that our study is not the first one to implicate a direct role of RNA modification in splicing regulation. It was previously shown that pseudouridine deposited on particular sites in the polypyrimidine tract (PPT) could block U2AF65 binding and inhibit *in vivo* splicing (Chen et al., 2010). A recent study expanded on this research, showing that pseudouridine is deposited co-transcriptionally and that intronic pseudouridines are associated with alternatively spliced introns (Martinez et al., 2020). Depletion of pseudouridine synthetases PUS1, RPUSD4 and PUS7 results in widespread alternative splicing changes (Martinez et al., 2020). Interestingly, *in vitro* splicing with pseudouridinylated substrates enhanced splicing (Martinez et al., 2020), unlike the previous study, where pseudouridinylation inhibited the splice site usage (Chen et al., 2010). It suggests that the effect of pseudouridine on splicing might be dependent either on the pre-mRNA substrate or the site of pseudouridinylation.

Finally, both our study and the previous studies describing the role of pseudouridine in splicing regulation suggest that a more widespread regulation of splicing by a direct interaction between different RNA modifications and splicing factors is possible and might contribute to alternative splicing regulation.

8.4. The role of U6 snRNA m^6A in splicing

The second conserved target of METTL16/METT-10 is U6 snRNA, methylated at the ACAGA motif, which is essential for the 5′ splice site recognition and splicing. The METTL16 consensus motif in U6 snRNA is conserved from yeast (*S. pombe*) to humans (Pendleton et al., 2017), with

methylation present in both species. Conservation of m^6A methylation at the crucial splicing motif suggests that m^6A might be important for splicing. Indeed, removal of *S. pombe* homologue of METTL16, *Duf890/mtl16* led to slower growth compared to the WT yeast (Pendleton et al., 2017). Recently it was shown that U6 snRNA m^6A methylation in *S. pombe* is needed for efficient splicing of a subset of introns, which are not strongly bound by U5 snRNA (Ishigami et al., 2021). These introns are characterized by adenosine as the 4^{th} nucleotide (A4 introns) with the exonic 5′ sequence being BBH instead of U5 snRNA consensus AAG (Ishigami et al., 2021). The lack of U6 snRNA m^6A leads to intron retention and significant changes in pre-mRNA splicing (Ishigami et al., 2021). It is the first mechanistic example of U6 snRNA m^6A role in splicing. However, the question remains about the role of U6 snRNA m^6A methylation in other organisms.

In our case, we haven't observed any major splicing defects, both in mouse embryos depleted of METTL16 (Chapter I) as well as in *mett-10* KO *C. elegans* (Chapter II). While in mouse the reason might be the maternal contribution of U6 snRNA, which was described to have a long half-life of about 24 hours (Fury and Zieve, 1996), that could not be the case in worms, where U6 snRNA m^6A methylation is completely absent as confirmed by the SCARLET method (Chapter II, Fig. 1), which allows for single-nucleotide resolution analysis of RNA modifications (Liu et al., 2013). Interestingly, the levels of U6 snRNA were two-fold increased in the *mett-10* KO worms (Chapter II, Fig. S1), while they were not affected in *S. pombe* (Ishigami et al., 2021). However, we can't explain why the loss of methylation would lead to increased U6 snRNA levels as the m^6A methylated region is not involved in U6 processing.

The reason why we did not see splicing abnormalities similar to *S. pombe* Δ*mtl16* is not apparent. Either, because of focusing on global splicing defects, we missed a subset of affected transcripts, or the 5′SS is recognized differently in worms and mice compared to yeast. While in yeast (*S. cerevisiae*), the U6 snRNA ACAGA motif strongly pairs with the intron sequence, a recent structure

of the human C* spliceosome complex shows that the m^{6}A methylated adenosine on U6 snRNA (U6-A43) does not pair with the intron, suggesting that it might not be essential for splicing in humans (Bertram et al., 2017). Early biochemical experiments seem to confirm that in humans, A43 is not essential for *in vitro* splicing as A to G mutation can still splice (although with lower efficiency). However, it is important to note that other mutations (A to C or A to U) completely inhibit splicing (Datta and Weiner, 1993). This is not the case in *S. cerevisiae*, where any mutation of AC**A**GA adenosine leads to lethality (Madhani et al., 1990).

8.5. Regulation of U6 snRNA m^{6}A methylation

Although we have not seen any significant splicing defect in *Mettl16* and *mett-10* KO, U6 snRNA m^{6}A may play a role in some specific developmental or environmental conditions. Thus, another interesting question is whether U6 snRNA m^{6}A methylation is deposited constitutively on all U6 snRNA molecules or whether the level of this modification can be modulated. Previous reports suggest that the former is true, with *S. pombe* (Ishigami et al., 2021) and HeLa cells (Liu and Pan, 2015) U6 snRNA ACAGA motif being completely m^{6}A methylated. However, Western Blot of METTL16 distribution in mouse tissues shows that its protein levels differ significantly between tissues, with METTL16 absent in several tissues, including kidney, liver, muscle and bone marrow (Chapter II, Fig. 6B). That raises an intriguing possibility that U6 snRNA m^{6}A methylation might be regulated in order to dynamically modify 5′SS usage. In addition, a demonstration that U6 snRNA is not equally m^{6}A methylated would not be the first example of dynamic changes in snRNA modifications. In 2019, FTO was proposed to regulate m^{6}A methylation of snRNA cap, being important for alternative splicing regulation (Mauer et al., 2019). Another study showed that methylation of certain 2′-O-methyl sites is affected in Jurkat cells (leukaemia model) compared to the activated T cells (Krogh et al., 2017). These results suggest that snRNA modifications might be more dynamic than commonly believed and involved in alternative splicing regulation.

Another interesting question is the level of U6 snRNA methylation in other organisms where METTL16 is conserved. *Drosophila melanogaster* is especially interesting as flies, compared to other species, have a very short METTL16 homologue. Fly protein is 305 aa long, while the worm is 479 aa and human is 562 aa. Although the methyltransferase domain is very well conserved between species, the C-terminal end is missing in flies. The region was shown to be essential for U6 snRNA binding and methylation by METTL16 (Aoyama et al., 2020), so its lack in flies could potentially impair U6 snRNA methylation efficiency.

8.6. The structural features of METTL16

To understand how METTL16 can recognize its targets and why it requires both the long consensus motif (UACAGAGAA in mammals) and the hairpin structure, we crystalised and solved the structure of the METTL16 methyltransferase domain (MTD) (Chapter I). Two constructs were selected based on a limited proteolysis analysis: full-length methyltransferase domain (1 – 291 aa) and ΔN construct (40 – 291 aa). We failed to crystalize constructs with the RNA substrate (hp1 from *Mat2a* mRNA); however, we obtained crystal structures of both constructs bound to SAH (Chapter I, Figure 1). Based on the structure, we defined two regions to be essential for METTL16 activity: METTL16 specific N-terminal region (1 – 78 aa) and a flexible loop region (F187 – G223), located next to the catalytic residues (NPPF, N184 – F187), which was disordered in our structure. Using various biochemical assays, we showed that the N-terminal region is rich in negatively charged amino acids and is important for RNA binding. At the same time, the flexible loop is needed for substrate methylation, but not binding, suggesting that it might be needed for substrate positioning. Interestingly, the loop region in *C. elegans* is longer (48 aa 35 aa) and has poor sequence conservation, suggesting that it might have different substrate specificity. Indeed, while human METTL16 methylates the worm *sams* region, the opposite is not true (Chapter II, Figure 3F and S3B).

Two other structures of METTL16 MTD were published, one describing MTD alone (Ruszkowska et al., 2018) and the second having MTD bound to its RNA substrate (*Mat2a* hp1) (Doxtader et al., 2018). While the structure of MTD alone shows a high degree of overlap with our study, MTD bound to hp1 RNA shows interesting differences. First of all, the loop region is stabilized upon the RNA binding and forms a "clamp" that holds the RNA and positions the substrate for methylation in the catalytic pocket. Mutation of a single residue in this region (R200Q), found in certain types of cancer, leads to upregulation in methylation activity (Doxtader et al., 2018). The second interesting region is the so-called K-loop (KTLLMD, 163 – 167 aa), which occludes the SAM binding site and was proposed to be an autoregulatory switch, which could modulate METTL16 dwell time on targets (Doxtader et al., 2018). The presence of this region might potentially improve SAM level sensing and allow for switching between splicing enhancement/methylation modes. An exciting, but not yet confirmed possibility, would be the deposition of post-translational modifications on this region to modulate the function of METTL16.

Additionally, a comparison of our structure and the one from Doxtader et al. explains why we didn't obtain the crystals of MTD bound to the RNA. While the boundaries of our construct were 1 – 291 aa, the other group used 1 to 310 amino acids. Analysis of the sequence between 291 and 310 amino acids shows several positively charged amino acids (DDVTVPSPPS**KRRK**LE**K**P**RK,** Uniprot: Q86W50 (Bateman et al., 2021)), which very likely are needed for stabilisation of the METTL16-protein complex. Additionally, while we used an endogenous hp1 sequence, Yunsun Nam used a mutated hp1 hairpin (called hp1x), where the stem region was mutated to obtain stronger pairing and more stable stem (Doxtader et al., 2018). Although METTL16 does not contact the stem region, stronger pairing might have stabilised the RNA substrate, improving crystallization conditions.

Finally, although the studies broaden understanding of METTL16 structure and substrate recognition, information about the C-terminus of the protein, containing so-called VCR regions, was still missing.

8.6.1. What is the role of the VCR domain?

One of the controversial questions is the role of METTL16 C-terminal domains called vertebrate conserved regions (VCR). Initially, it was proposed that while METTL16 MTD binds to the hairpin structures in *MAT2A* pre-mRNA, VCR domains are essential for *MAT2A* pre-mRNA splicing. The model was supported by tethering assays, where VCR domains alone could stimulate the splicing of a construct containing the last intron of *MAT2A* pre-mRNA (Pendleton et al., 2017). Although a precise mechanism of splicing regulation remained unknown, the authors suggested that additional co-factors will be necessary for splicing induction (Pendleton et al., 2017).

Recently, a follow-up study identified CFI_m25 (CPSF5, NUDT21), a component of the cleavage factor Im (CFI_m) complex, as being important for *MAT2A* intron splicing (Scarborough et al., 2021). CFI_m25 tethering was enough to stimulate *MAT2A* pre-mRNA splicing in *METTL16* siRNA KD, suggesting that METTL16 is upstream of CFI_m25 and might recruit it (Scarborough et al., 2021). However, the study failed to show any interaction between METTL16 and CFI_m25, implying that VCR do not interact directly with CFI_m25. What is then a function of the VCR domains? How does it recruit CFI_m25? The authors hypothesised that VCR would not bind CFI_m25 directly, but rather stimulate CFI_m25 binding by rearranging the RNA secondary structure to unravel the CFI_m25 binding motif.

A recent study seems to support the role of VCR in RNA binding and potential structural rearrangements. They show that VCR is an RNA binding domain needed for U6 snRNA, but not *MAT2A* hairpin, binding (Aoyama et al., 2020). The binding affinity of METTL16 MTD alone is

over 100x weaker for U6 snRNA compared to the full-length protein, while only 3x weaker in the case of *Mat2a* hairpin. The group crystalized METTL16 VCR domains and showed that their structure is similar to the kinase-associated 1 (KA1) domain of terminal uridylyltransferase 1 (TUT1), a known U6 snRNA binding protein. In addition, they mapped VCR binding to the internal stem-loop (ISL) region of U6 snRNA and proposed that it introduces structural changes in U6 snRNA to facilitate methylation by METTL16 (Aoyama et al., 2020). It is worth noting that already before, the VCR region was shown to be important for binding to *MALAT1* (Ruszkowska et al., 2018).

In summary, our understanding of the VCR domain role has evolved from a platform for splicing factors recruitment to an RNA binding domain, required for *MALAT1* and U6 snRNA binding and potentially introducing structural changes to these targets. At the same time, the VCR domain was proposed to be conserved only in vertebrates, so how U6 snRNA is recognized and methylated in the other species, especially in *C. elegans*? Is U6 snRNA recognised and bound differently, or is the function of the C-terminal region also conserved in non-vertebrates? METTL16 homologues in many species, involving *C. elegans* METT-10, have a long C-terminal tail. Analysis of METT-10 region 200 – 479 aa using Phyre2 database (Kelley et al., 2015) finds high similarity to the vertebrate VCR domains. I decided to explore it further using AlphaFold2, a recently released machine learning algorithm, which offers significantly improved *in silico* structure prediction (Jumper et al., 2021). Comparison of the full-length structures of human METTL16 and *C. elegans* METT-10 shows high similarity in the C-terminal region, suggesting that the VCR region might not be limited to vertebrates only (Figure 6). However, further structural and biochemical studies are needed to confirm this prediction.

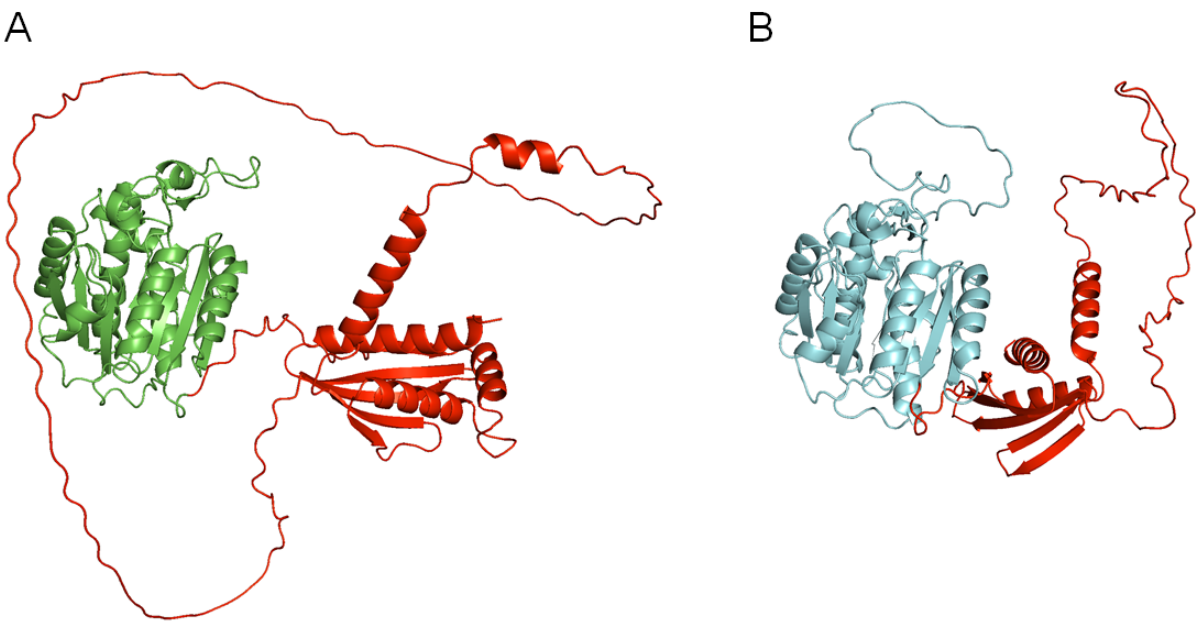


Figure 6. Comparison of METTL16 (A) and METT-10 (B) full-length structures predicted using AlphaFold2 algorithm (Jumper et al., 2021). C-terminal region (marked in red) shows a high similarity between human and worm protein, suggesting that VCR regions might be conserved in non-verebrates.

8.6.2. RNA substrate requirements of METTL16

While the level of m^{6}A methylation of SAM synthetase transcripts is dynamically changing in different nutrient conditions in *C. elegans* (Chapter II, Fig. 4), methylation of U6 snRNA seems to remain stable both in low and high nutrient conditions (Chapter II, Fig. S1). Either both targets are methylated at the same rate, and lack of changes in U6 snRNA m^{6}A methylation levels is due to its long half-life (U6 is so stable that it will finally be methylated, while the turnover of *sams* transcripts is much faster), or perhaps METT-10 (and METTL16) can distinguish between different substrates and adjust its mode of action. Ours and other groups results seem to suggest that METTL16/METT-10 is able to sense differences in RNA substrates, and even small changes might result in significant changes in methylation efficiency (Doxtader et al., 2018; Mendel et al., 2018).

Our analysis of different *Mat2a* hp1 substrate mutants showed that METTL16 recognizes and methylates substrates with consensus motif present in a bulge of a hairpin structure. Both consensus

motif and secondary structure are essential for the methylation, although mutations of the stem region showed that the stem sequence is not essential as long as the pairing, leading to hairpin formation, is preserved (Chapter I, Fig. 3). These results were explained by the structure of METTL16 bound to RNA, where the stem region is not contacted by METTL16 (Doxtader et al., 2018). More importantly, the same study showed that both the size of the bulge region as well as mutations in the hp1 consensus motif can modulate methylation activity. A single mutation (UACAGA**G**AA into UACAGA**A**AA) present in the consensus motif of *Mat2a* hp5 hairpin increases methylation activity 5-fold (Doxtader et al., 2018)

Finally, it is important to mention that methylation assays indicate that METTL16 and METT-10 have different target requirements, with METT-10 having a much more limited set of substrates. While human METTL16 can efficiently methylate all tested *Mat2a* hairpins (mouse, worm, silkworm, and fly) (Chapter II, Figure 3F), worm METT-10 does not recognize mouse *Mat2a* hairpin and methylates efficiently only the worm sequence (Chapter II, Figure S3). The difference might be caused by different loop regions or discrepancies in the C-terminal region.

8.7. Conclusions

In conclusion, during my PhD, I gained important insights about METTL16 function and mode of action in mice and the role of METTL16 homologue in *C. elegans*, METT-10. In addition, I described a new mechanism of m^6A-driven splicing regulation, which is conserved between worms and mammals.

The most important points are:

- In mice, METTL16 is an essential RNA m^6A methyltransferase required for both early embryonic development and male germline differentiation;
- METTL16 structure analysis identified two crucial regions: the N-terminal region required for RNA binding and the loop region necessary for RNA methylation;
- METTL16 homologue in *C. elegans*, METT-10, is mRNA m^6A methyltransferase, methylating U6 snRNA and SAM synthetase transcripts;
- Methylation of SAM synthetase transcripts by METT-10 is essential for SAM levels homeostasis in *C. elegans*;
- m^6A deposited on the 3′SS directly inhibits splice site usage through blocking U2AF35 splicing factor binding, with the mechanism of splicing inhibition conserved between nematodes and mammals.

9. Bibliography

Akichika, S., Hirano, S., Shichino, Y., Suzuki, T., Nishimasu, H., Ishitani, R., Sugita, A., Hirose, Y., Iwasaki, S., Nureki, O., and Suzuki, T. (2019). Cap-specific terminal N 6 -methylation of RNA by an RNA polymerase II–associated methyltransferase. Science (80-.). *363*, eaav0080.

Allen, F.W. (1941). The Biochemistry of the Nucleic Acids, Purines, and Pyrimidines. Annu. Rev. Biochem. *10*, 221–244.

Aoyama, T., Yamashita, S., and Tomita, K. (2020). Mechanistic insights into m6A modification of U6 snRNA by human METTL16. Nucleic Acids Res. *48*, 5157–5168.

Arango, D., Sturgill, D., Alhusaini, N., Dillman, A.A., Sweet, T.J., Hanson, G., Hosogane, M., Sinclair, W.R., Nanan, K.K., Mandler, M.D., Fox, S.D., Zengeya, T.T., Andresson, T., Meier, J.L., Coller, J., and Oberdoerffer, S. (2018). Acetylation of Cytidine in mRNA Promotes Translation Efficiency. Cell *175*, 1872-1886.e24.

Avery, O.T., MacLeod, C.M., and McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. J. Exp. Med. *79*, 137–158.

Bailey, A.S., Batista, P.J., Gold, R.S., Grace Chen, Y., de Rooij, D.G., Chang, H.Y., and Fuller, M.T. (2017). The conserved RNA helicase YTHDC2 regulates the transition from proliferation to differentiation in the germline. Elife *6*, 1–29.

Barbieri, I., and Kouzarides, T. (2020). Role of RNA modifications in cancer. Nat. Rev. Cancer *20*, 303–322.

Barbosa-Morais, N.L., Irimia, M., Pan, Q., Xiong, H.Y., Gueroussov, S., Lee, L.J., Slobodeniuc, V., Kutter, C., Watt, S., Çolak, R., Kim, T.H., Misquitta-Ali, C.M., Wilson, M.D., Kim, P.M., Odom, D.T., Frey, B.J., and Blencowe, B.J. (2012). The evolutionary landscape of alternative splicing in vertebrate species. Science (80-.). *338*, 1587–1593.

Bartosovic, M., Molares, H.C., Gregorova, P., Hrossova, D., Kudla, G., and Vanacova, S. (2017). N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3′-end processing. Nucleic Acids Res. *45*, 11356–11370.

Bateman, A., Martin, M.J., Orchard, S., Magrane, M., Agivetova, R., Ahmad, S., Alpi, E., Bowler-Barnett, E.H., Britto, R., Bursteinas, B., Bye-A-Jee, H., Coetzee, R., … Zhang, J. (2021). UniProt: The universal protein knowledgebase in 2021. Nucleic Acids Res. *49*, D480–D489.

Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., Carter, A.C., Flynn, R.A., Zhou, C., Lim, K.-S., Dedon, P., Wernig, M., Mullen, A.C., Xing, Y., Giallourakis, C.C., and Chang, H.Y. (2014). m6A RNA Modification Controls Cell Fate Transition in Mammalian Embryonic Stem Cells. Cell Stem Cell *15*, 707–719.

Bawankar, P., Lence, T., Paolantoni, C., Haussmann, I.U., Kazlauskiene, M., Jacob, D., Heidelberger, J.B., Richter, F.M., Nallasivan, M.P., Morin, V., Kreim, N., Beli, P., Helm, M., Jinek, M., Soller, M., and Roignant, J.Y. (2021). Hakai is required for stabilization of core components of the m6A mRNA methylation machinery. Nat. Commun. *12*, 1–15.

Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5’ terminus of adenovirus 2 late mRNA. Proc. Natl. Acad. Sci. U. S. A. *74*, 3171–3175.

Berglund, J.A., Chua, K., Abovich, N., Reed, R., and Rosbash, M. (1997). The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. Cell *89*, 781–787.

Bertram, K., Agafonov, D.E., Liu, W.T., Dybkov, O., Will, C.L., Hartmuth, K., Urlaub, H., Kastner, B., Stark, H., and Lührmann, R. (2017). Cryo-EM structure of a human spliceosome activated for step 2 of

splicing. Nature *542*, 318–323.

Blumenthal, T. (2012). Trans-splicing and operons in C. elegans. WormBook 1–11.

Blumenthal, T., and Thomas, J. (1988). Cis and trans mRNA splicing in C. elegans. Trends Genet. *4*, 305–308.

Boccaletto, P., Machnicka, M.A., Purta, E., Piątkowski, P., Bagiński, B., Wirecki, T.K., de Crécy-Lagard, V., Ross, R., Limbach, P.A., Kotter, A., Helm, M., and Bujnicki, J.M. (2018). MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res. *46*, D303–D307.

Bokar, J.A., Rath-Shambaugh, M.E., Ludwiczak, R., Narayan, P., and Rottman, F. (1994). Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. J. Biol. Chem. *269*, 17697–17704.

Bokar, J.A., Shambaugh, M.E., Polayes, D., Matera, A.G., and Rottman, F.M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA *3*, 1233–1247.

Brenner, S., Jacob, F., and Meselson, M. (1961). An Unstable Intermediate Carrying Information from Genes to Ribosomes for Protein Synthesis. Nature *190*, 576–581.

Bringmann, P., and Lührmann, R. (1987). Antibodies specific for N 6 -methyladenosine react with intact snRNPs U2 and U4/U6. FEBS Lett. *213*, 309–315.

Brody, E., and Abelson, J. (1985). The “spliceosome”: yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. Science (80-.). *228*, 963–967.

Brown, J.A., Kinzig, C.G., DeGregorio, S.J., and Steitz, J.A. (2016). Methyltransferase-like protein 16 binds the 3′-terminal triple helix of MALAT1 long noncoding RNA. Proc. Natl. Acad. Sci. 201614759.

Busch, A., and Hertel, K.J. (2012). Evolution of SR protein and hnRNP splicing regulatory factors. Wiley Interdiscip. Rev. RNA *3*, 1–12.

Carroll, S.M., Narayan, P., and Rottman, F.M. (1990). N6-methyladenosine residues in an intron-specific region of prolactin pre-mRNA. Mol. Cell. Biol. *10*, 4456–4465.

Chargaff, E., Lipshitz, R., and Green, C. (1952). Composition of the desoxypentose nucleic acids of four genera of sea-urchin. J. Biol. Chem. *195*, 155–160.

Chen, C., Zhao, X., Kierzek, R., and Yu, Y.-T. (2010). A Flexible RNA Backbone within the Polypyrimidine Tract Is Required for U2AF65 Binding and Pre-mRNA Splicing InVivo. Mol. Cell. Biol. *30*, 4108–4119.

Chen, D., Wu, C., Zhao, S., Geng, Q., Gao, Y., Li, X., Zhang, Y., and Wang, Z. (2014). Three RNA Binding Proteins Form a Complex to Promote Differentiation of Germline Stem Cell Lineage in Drosophila. PLoS Genet. *10*.

Chen, H., Gu, L., Orellana, E.A., Wang, Y., Guo, J., Liu, Q., Wang, L., Shen, Z., Wu, H., Gregory, R.I., Xing, Y., and Shi, Y. (2020). METTL4 is an snRNA m6Am methyltransferase that regulates RNA splicing. Cell Res. *30*, 544–547.

Chen, M., Urs, M.J., Sánchez-González, I., Olayioye, M.A., Herde, M., and Witte, C.-P.P. (2018). m6A RNA Degradation Products Are Catabolized by an Evolutionarily Conserved N 6 -Methyl-AMP Deaminase in Plant and Mammalian Cells. Plant Cell *30*, 1511–1522.

Choe, J., Lin, S., Zhang, W., Liu, Q., Wang, L., Ramirez-Moya, J., Du, P., Kim, W., Tang, S., Sliz, P., Santisteban, P., George, R.E., Richards, W.G., Wong, K.K., Locker, N., Slack, F.J., and Gregory, R.I. (2018). mRNA circularization by METTL3–eIF3h enhances translation and promotes oncogenesis. Nature *561*, 556–560.

Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977). An amazing sequence arrangement at the

5′ ends of adenovirus 2 messenger RNA. Cell *12*, 1–8.

Clancy, M.J., Shambaugh, M.E., Timpte, C.S., and Bokar, J.A. (2002). Induction of sporulation in Saccharomyces cerevisiae leads to the formation of N6-methyladenosine in mRNA: A potential mechanism for the activity of the IME4 gene. Nucleic Acids Res. *30*, 4509–4518.

Conrad, N.K., Shu, M. Di, Uyhazi, K.E., and Steitz, J.A. (2007). Mutational analysis of a viral RNA element that counteracts rapid RNA decay by interaction with the polyadenylate tail. Proc. Natl. Acad. Sci. U. S. A. *104*, 10412–10417.

Crick, F.H. (1958). On protein synthesis. Symp. Soc. Exp. Biol. *12*, 138–163.

Dahm, R. (2005). Friedrich Miescher and the discovery of DNA. Dev. Biol. *278*, 274–288.

Darnell, J.E., Wall, R., and Tushinski, R.J. (1971). An adenylic acid-rich sequence in messenger RNA of HeLa cells and its possible relationship to reiterated sites in DNA. Proc. Natl. Acad. Sci. U. S. A. *68*, 1321–1325.

Datta, B., and Weiner, A.M. (1993). The phylogenetically invariant ACAGAGA and AGC sequences of U6 small nuclear RNA are more tolerant of mutation in human cells than in Saccharomyces cerevisiae. Mol. Cell. Biol. *13*, 5377–5382.

Davidson, J.N., and Waymouth, C. (1943). Ribonucleic Acids in Animal Tissues. Nature *152*, 47–48.

Davis, F.F., and Allen, F.W. (1957). Ribonucleic acids from yeast which contain a fifth nucleotide. J. Biol. Chem. *227*, 907–915.

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

van Delft, P., Akay, A., Huber, S.M., Bueschl, C., Rudolph, K.L.M., Di Domenico, T., Schuhmacher, R., Miska, E.A., and Balasubramanian, S. (2017). The Profile and Dynamics of RNA Modifications in Animals. ChemBioChem *18*, 979–984.

Deng, X., Chen, K., Luo, G.Z., Weng, X., Ji, Q., Zhou, T., and He, C. (2015). Widespread occurrence of N6-methyladenosine in bacterial mRNA. Nucleic Acids Res. *43*, 6557–6567.

Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of Methylated Nucleosides in Messenger RNA from Novikoff Hepatoma Cells. Proc. Natl. Acad. Sci. U. S. A. *71*, 3971–3975.

Devarkar, S.C., Wang, C., Miller, M.T., Ramanathan, A., Jiang, F., Khan, A.G., Patel, S.S., and Marcotrigiano, J. (2016). Structural basis for m7G recognition and 2′-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. Proc. Natl. Acad. Sci. U. S. A. *113*, 596–601.

Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan, T.F., Weninger, W.J., Westerberg, H., Adissu, H., Baker, C.N., Bower, L., … Murakami, A. (2016). High-throughput discovery of novel developmental phenotypes. Nature *537*, 508–514.

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

Dorsett, M., and Schedl, T. (2009). A role for dynein in the inhibition of germ cell proliferative fate. Mol. Cell. Biol. *29*, 6128–6139.

Dorsett, M., Westlund, B., and Schedl, T. (2009). METT-10, a putative methyltransferase, inhibits germ cell proliferative fate in Caenorhabditis elegans. Genetics *183*, 233–247.

Doucet, A.J., Droc, G., Siol, O., Audoux, J., and Gilbert, N. (2015). U6 snRNA pseudogenes: Markers of retrotransposition dynamics in mammals. Mol. Biol. Evol. *32*, 1815–1832.

Doxtader, K.A., Wang, P., Scarborough, A.M., Seo, D., Conrad, N.K., and Nam, Y. (2018). Structural Basis for Regulation of METTL16, an S-Adenosylmethionine Homeostasis Factor. Mol. Cell *71*, 1001-1011.e4.

Du, H., Zhao, Y., He, J., Zhang, Y., Xi, H., Liu, M., Ma, J., and Wu, L. (2016). YTHDF2 destabilizes m6A-containing RNA through direct recruitment of the CCR4–NOT deadenylase complex. Nat. Commun. *7*, 12626.

Edmonds, M., Vaughan, M.H., and Nakazato, H. (1971). Polyadenylic acid sequences in the heterogeneous nuclear RNA and rapidly-labeled polyribosomal RNA of HeLa cells: possible evidence for a precursor relationship. Proc. Natl. Acad. Sci. U. S. A. *68*, 1336–1340.

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

Fedeles, B.I., Singh, V., Delaney, J.C., Li, D., and Essigmann, J.M. (2015). The AlkB family of Fe(II)/α-ketoglutarate-dependent dioxygenases: Repairing nucleic acid alkylation damage and beyond. J. Biol. Chem. *290*, 20734–20742.

Finkelstein, J.D. (1990). Methionine metabolism in mammals. J. Nutr. Biochem. *1*, 228–237.

Fortes, P., Inada, T., Preiss, T., Hentze, M.W., Mattaj, I.W., and Sachs, A.B. (2000). The yeast nuclear cap binding complex can interact with translation factor eIF4G and mediate translation initiation. Mol. Cell *6*, 191–196.

Furlanis, E., and Scheiffele, P. (2018). Regulation of Neuronal Differentiation, Function, and Plasticity by Alternative Splicing. Annu. Rev. Cell Dev. Biol. *34*, 451–469.

Fury, M.G., and Zieve, G.W. (1996). U6 snRNA maturation and stability. Exp. Cell Res. *228*, 160–163.

Fustin, J.-M.M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., Isagawa, T., Morioka, M.S., Kakeya, H., Manabe, I., and Okamura, H. (2013). RNA-Methylation-Dependent RNA Processing Controls the Speed of the Circadian Clock. Cell *155*, 793–806.

Fustin, J.M., Kojima, R., Itoh, K., Chang, H.Y., Ye, S., Zhuang, B., Oji, A., Gibo, S., Narasimamurthy, R., Virshup, D., Kurosawa, G., Doi, M., Manabe, I., Ishihama, Y., Ikawa, M., and Okamura, H. (2018). Two Ck1δ transcripts regulated by m6A methylation code for two antagonistic kinases in the control of the circadian clock. Proc. Natl. Acad. Sci. *115*, E6386–E6386.

Gerken, T., Girard, C.A., Tung, Y.C.L., Webby, C.J., Saudek, V., Hewitson, K.S., Yeo, G.S.H., McDonough, M.A., Cunliffe, S., McNeill, L.A., Galvanovskis, J., Rorsman, P., Robins, P., Prieur, X., Coll, A.P., Ma, M., Jovanovic, Z., Farooqi, I.S., Sedgwick, B., Barroso, I., Lindahl, T., Ponting, C.P., Ashcroft, F.M., O'Rahilly, S., and Schofield, C.J. (2007). The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science (80-.). *318*, 1469–1472.

Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., Ben-Haim, M.S., Eyal, E., Yunger, S., Pinto, Y., Jaitin, D.A., Viukov, S., Rais, Y., Krupalnik, V., Chomsky, E., Zerbib, M., Maza, I., Rechavi, Y., Massarwa, R., Hanna, S., Amit, I., Levanon, E.Y., Amariglio, N., Stern-Ginossar, N., Novershtern, N., Rechavi, G., and Hanna, J.H. (2015). m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. Science (80-.). *347*, 1002–1006.

Goh, Y.T., Koh, C.W.Q., Sim, D.Y., Roca, X., and Goh, W.S.S. (2020). METTL4 catalyzes m6Am methylation in U2 snRNA to regulate pre-mRNA splicing. Nucleic Acids Res. *48*, 9250–9261.

Graveley, B.R. (2001). Alternative splicing: Increasing diversity in the proteomic world. Trends Genet. *17*, 100–107.

Greer, E.L., Blanco, M.A., Gu, L., Sendinc, E., Liu, J., Aristizábal-Corrales, D., Hsu, C.H., Aravind, L., He, C., and Shi, Y. (2015). DNA methylation on N6-adenine in C. elegans. Cell *161*, 868–878.

Gros, F., Hiatt, H., Gilbert, W., Kurland, C.G., Risebrough, R.W., and Watson, J.D. (1961). Unstable Ribonucleic Acid Revealed by Pulse Labelling of Escherichia Coli. Nature *190*, 581–585.

Grozhik, A. V., Olarerin-George, A.O., Sindelar, M., Li, X., Gross, S.S., and Jaffrey, S.R. (2019). Antibody cross-reactivity accounts for widespread appearance of m1A in 5'UTRs. Nat. Commun. *10*, 1–13.

Hall, K., and Sankaran, N. (2021). DNA translated: Friedrich Miescher's discovery of nuclein in its original context. Br. J. Hist. Sci. *54*, 99–107.

Harris, T.W., Arnaboldi, V., Cain, S., Chan, J., Chen, W.J., Cho, J., Davis, P., Gao, S., Grove, C.A., Kishore, R., Lee, R.Y.N., Muller, H.M., Nakamura, C., Nuin, P., Paulini, M., Raciti, D., Rodgers, F.H., Russell, M., Schindelman, G., Auken, K. V., Wang, Q., Williams, G., Wright, A.J., Yook, K., Howe, K.L., Schedl, T., Stein, L., and Sternberg, P.W. (2020). WormBase: A modern Model Organism Information Resource. Nucleic Acids Res. *48*, D762–D767.

Hartmann, A.M., Nayler, O., Schwaiger, F.W., Obermeier, A., and Stamm, S. (1999). The interaction and colocalization of Sam68 with the splicing-associated factor YT521-B in nuclear dots is regulated by the Src family kinase p59(fyn). Mol. Biol. Cell *10*, 3909–3926.

Haussmann, I.U., Bodi, Z., Sanchez-Moran, E., Mongan, N.P., Archer, N., Fray, R.G., and Soller, M. (2016). m6A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. Nature *540*, 301–304.

Henry, D. (2006). Aristotle on the mechanism of inheritance. J. Hist. Biol. *39*, 425–455.

Hess, M.E., Hess, S., Meyer, K.D., Verhagen, L.A.W., Koch, L., Brönneke, H.S., Dietrich, M.O., Jordan, S.D., Saletore, Y., Elemento, O., Belgardt, B.F., Franz, T., Horvath, T.L., Rüther, U., Jaffrey, S.R., Kloppenburg, P., and Brüning, J.C. (2013). The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. Nat. Neurosci. *16*, 1042–1048.

Hocine, S., Singer, R.H., and Grünwald, D. (2010). RNA processing and export. Cold Spring Harb. Perspect. Biol. *2*, 1–21.

Hollins, C., Zorio, D.A.R., MacMorris, M., and Blumenthal, T. (2005). U2AF binding selects for the high conservation of the C. elegans 3' splice site. RNA *11*, 248–253.

Horiuchi, K., Kawamura, T., Iwanari, H., Ohashi, R., Naito, M., Kodama, T., and Hamakubo, T. (2013). Identification of Wilms' tumor 1-associating protein complex and its role in alternative splicing and the cell cycle. J. Biol. Chem. *288*, 33292–33302.

Horowitz, S., Horowitz, A., and Nilsen, T.W. (1984). Mapping of N6-methyladenosine residues in bovine prolactin mRNA. Proc. Natl. Acad. Sci. U. S. A. *81*, 5667–5671.

Howe, K.L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Ridwan Amode, M., Armean, I.M., Azov, A.G., Bennett, R., Bhai, J., Billis, K., Boddu, S., … Flicek, P. (2021). Ensembl 2021. Nucleic Acids Res. *49*, D884–D891.

Hsu, P.J., Zhu, Y., Ma, H., Guo, Y., Shi, X., Liu, Y., Qi, M., Lu, Z., Shi, H., Wang, J., Cheng, Y., Luo, G., Dai, Q., Liu, M., Guo, X., Sha, J., Shen, B., and He, C. (2017). Ythdc2 is an N6-methyladenosine binding protein that regulates mammalian spermatogenesis. Cell Res. *27*, 1115–1127.

Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., Zhao, B.S., Mesquita, A., Liu, C., Yuan, C.L., Hu, Y.C., Hüttelmaier, S., Skibbe, J.R., Su, R., Deng, X., Dong, L., Sun, M., Li, C., Nachtergaele, S., Wang, Y., Hu, C., Ferchen, K., Greis, K.D., Jiang, X., Wei, M., Qu, L., Guan, J.L., He, C., Yang, J., and Chen, J. (2018). Recognition of RNA N6-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat. Cell Biol. *20*, 285–295.

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., Calzada-Wack, J., Becker, L., Marschall, S., Kraiger, M., Garrett, L., Seisenberger, C., Hölter, S.M., Borland, K., de Logt, E. Van, Jansen, P.W.T.C., Baltissen, M.P., Valenta, M., Vermeulen, M., Wurst, W., Gailus-Durner, V., Fuchs, H., de Angelis, M.H., Rando, O.J., Kellner, S.M., Bultmann, S., and Schneider, R. (2020). The rRNA m6A methyltransferase METTL5 is involved in pluripotency and developmental programs. Genes Dev. *34*, 715–729.

Ikeda, S., Kawahara-Miki, R., Iwata, H., Sugimoto, M., and Kume, S. (2017). Role of methionine adenosyltransferase 2A in bovine preimplantation development and its associated genomic regions. Sci. Rep. *7*, 1–10.

Ishigami, Y., Ohira, T., Isokawa, Y., Suzuki, Y., and Suzuki, T. (2021). A single m6A modification in U6 snRNA diversifies exon sequence at the 5' splice site. Nat. Commun. *12*.

Ivanova, I., Much, C., Di Giacomo, M., Azzi, C., Morgan, M., Moreira, P.N., Monahan, J., Carrieri, C., Enright, A.J., and O'Carroll, D. (2017). The RNA m6A Reader YTHDF2 Is Essential for the Post-transcriptional Regulation of the Maternal Transcriptome and Oocyte Competence. Mol. Cell *67*, 1059-1067.e4.

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

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Olaf, R., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., and Reiman, D. (2021). Highly accurate protein structure prediction.

Kandels-Lewis, S., and Séraphin, B. (1993). Involvement of U6 snRNA in 5' splice site selection. Science *262*, 2035–2039.

Kane, S.E., and Beemon, K. (1985). Precise localization of m6A in Rous sarcoma virus RNA reveals clustering of methylation sites: implications for RNA processing. Mol. Cell. Biol. *5*, 2298–2306.

Kasowitz, S.D., Ma, J., Anderson, S.J., Leu, N.A., Xu, Y., Gregory, B.D., Schultz, R.M., and Wang, P.J. (2018). Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. PLoS Genet. *14*, 1–28.

Ke, S., Alemu, E.A., Mertens, C., Gantman, E.C., Fak, J.J., Mele, A., Haripal, B., Zucker-Scharff, I., Moore, M.J., Park, C.Y., Vågbø, C.B., Kusśnierczyk, A., Klungland, A., Darnell, J.E., and Darnell, R.B. (2015). A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes Dev. *29*, 2037–2053.

Ke, S., Pandya-Jones, A., Saito, Y., Fak, J.J., Vågbø, C.B., Geula, S., Hanna, J.H., Black, D.L., Jr, J.E.D., Darnell, R.B., Darnell, J.E., Darnell, R.B., Jr, J.E.D., Darnell, R.B., Darnell, J.E., and Darnell, R.B. (2017). m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. Genes Dev. *31*, 990–1006.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. Nat. Protoc. *10*, 845–858.

Keren, H., Lev-Maor, G., and Ast, G. (2010). Alternative splicing and evolution: Diversification, exon definition and function. Nat. Rev. Genet. *11*, 345–355.

Kim, J., Kim, Y., Yeom, M., Kim, J.H., and Nam, H.G. (2008). FIONA1 is essential for regulating period length in the Arabidopsis circadian clock. Plant Cell *20*, 307–319.

Knuckles, P., Lence, T., Haussmann, I.U., Jacob, D., Kreim, N., Carl, S.H., Masiello, I., Hares, T., Villaseñor, R., Hess, D., Andrade-Navarro, M.A., Biggiogera, M., Helm, M., Soller, M., Bühler, M., and Roignant, J.Y. (2018). Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor RbM15/spenito to the m6 a machinery component Wtap/Fl(2)d. Genes Dev. *32*, 415–429.

Konarska, M.M., Padgett, R.A., and Sharp, P.A. (1984). Recognition of cap structure in splicing in vitro of mRNA precursors. Cell *38*, 731–736.

Kramer, A., and Utans, U. (1991). Three protein factors (SF1, SF3 and U2AF) function in pre-splicing complex formation in addition to snRNPs. EMBO J. *10*, 1503–1509.

Kretschmer, J., Rao, H., Hackert, P., Sloan, K.E., Höbartner, C., and Bohnsack, M.T. (2018). The m6A

reader protein YTHDC2 interacts with the small ribosomal subunit and the 5′-3′ exoribonuclease XRN1. Rna *24*, 1339–1350.

Krogh, N., Kongsbak-Wismann, M., Geisler, C., and Nielsen, H. (2017). Substoichiometric ribose methylations in spliceosomal snRNAs. Org. Biomol. Chem. *15*, 8872–8876.

Krug, R.M., Morgan, M.A., and Shatkin, A.J. (1976). Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. J. Virol. *20*, 45–53.

Lasman, L., Krupalnik, V., Viukov, S., Mor, N., Aguilera-Castrejon, A., Schneir, D., Bayerl, J., Mizrahi, O., Peles, S., Tawil, S., Sathe, S., Nachshon, A., Shani, T., Zerbib, M., Kilimnik, I., Aigner, S., Shankar, A., Mueller, J.R., Schwartz, S., Stern-Ginossar, N., Yeo, G.W., Geula, S., Novershtern, N., and Hanna, J.H. (2020). Context-dependent compensation between functional Ythdf m6A reader proteins. Genes Dev. *34*, 1373–1391.

Lence, T., Akhtar, J., Bayer, M., Schmid, K., Spindler, L., Ho, C.H., Kreim, N., Andrade-Navarro, M.A., Poeck, B., Helm, M., and Roignant, J.-Y. (2016). m6A modulates neuronal functions and sex determination in Drosophila. Nature *540*, 242–247.

Lesser, C.F., and Guthrie, C. (1993). Mutations in U6 snRNA that alter splice site specificity: implications for the active site. Science *262*, 1982–1988.

Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., Mao, Y., Lv, J., Yi, D., Chen, X.-W., Wang, C., Qian, S.-B., and Yi, C. (2017). Base-Resolution Mapping Reveals Distinct m1A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. Mol. Cell *68*, 993-1005.e9.

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., and Greer, E.L. (2020). N6-adenosine methylation of ribosomal RNA affects lipid oxidation and stress resistance. Sci. Adv. *6*, eaaz4370.

Liu, Y. (2008). A new perspective on Darwin's Pangenesis. Biol. Rev. *83*, 141–149.

Liu, N., and Pan, T. (2015). Probing RNA Modification Status at Single-Nucleotide Resolution in Total RNA. In Methods in Enzymology, (Elsevier Inc.), pp. 149–159.

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

Liu, J., Li, K., Cai, J., Zhang, M., Zhang, X., Xiong, X., Meng, H., Xu, X., Huang, Z., Peng, J., Fan, J., and Yi, C. (2020). Landscape and Regulation of m6A and m6Am Methylome across Human and Mouse Tissues. Mol. Cell *77*, 426-440.e6.

Liu, N., Parisien, M., Dai, Q., Zheng, G., He, C., and Pan, T. (2013). Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA *19*, 1848–1856.

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.

Louloupi, A., Ntini, E., Conrad, T., and Ørom, U.A.V. (2018). Transient N-6-Methyladenosine Transcriptome Sequencing Reveals a Regulatory Role of m6A in Splicing Efficiency. Cell Rep. *23*, 3429–3437.

Lu, S.C., and Mato, J.M. (2012). S-adenosylmethionine in liver health, injury, and cancer. Physiol. Rev. *92*, 1515–1542.

Lu, S.C., Alvarez, L., Huang, Z.Z., Chen, L., An, W., Corrales, F.J., Avila, M.A., Kanel, G., and Mato, J.M. (2001). Methionine adenosyltransferase 1A knockout mice are predisposed to liver injury and exhibit increased expression of genes involved in proliferation. Proc. Natl. Acad. Sci. U. S. A. *98*, 5560–5565.

Luo, S., and Tong, L. (2014). Molecular basis for the recognition of methylated adenines in RNA by the

eukaryotic YTH domain. Proc. Natl. Acad. Sci. U. S. A. *111*, 13834–13839.

Ma, H., Wang, X., Cai, J., Dai, Q., Natchiar, S.K., Lv, R., Chen, K., Lu, Z., Chen, H., Shi, Y.G., Lan, F., Fan, J., Klaholz, B.P., Pan, T., Shi, Y., and He, C. (2019). N6-Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. Nat. Chem. Biol. *15*, 88–94.

Madhani, H.D., Bordonne, R., and Guthrie, C. (1990). Multiple roles for U6 snRNA in the splicing pathway. Genes Dev. *4*, 2264–2277.

Martinez, N.M., Su, A., Nussbacher, J.K., Burns, M.C., Schaening, C., Sathe, S., Yeo, G.W., and Gilbert, W. V (2020). Pseudouridine synthases modify human pre-mRNA co-transcriptionally and affect splicing. BioRxiv.

Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D.P., Linder, B., Pickering, B.F., Vasseur, J.-J., Chen, Q., Gross, S.S., Elemento, O., Debart, F., Kiledjian, M., and Jaffrey, S.R. (2017). Reversible methylation of m6Am in the 5' cap controls mRNA stability. Nature *541*, 371–375.

Mauer, J., Sindelar, M., Despic, V., Guez, T., Hawley, B.R., Vasseur, J.J., Rentmeister, A., Gross, S.S., Pellizzoni, L., Debart, F., Goodarzi, H., and Jaffrey, S.R. (2019). FTO controls reversible m 6 Am RNA methylation during snRNA biogenesis. Nat. Chem. Biol. *15*, 340–347.

Mendel, M., Chen, K.-M., Homolka, D., Gos, P., Pandey, R.R., McCarthy, A.A., and Pillai, R.S. (2018). Methylation of Structured RNA by the m6A Writer METTL16 Is Essential for Mouse Embryonic Development. Mol. Cell *71*, 986-1000.e11.

Meneely, P.M., Dahlberg, C.L., and Rose, J.K. (2019). Working with Worms: Caenorhabditis elegans as a Model Organism. Curr. Protoc. Essent. Lab. Tech. *19*, 1–35.

Meyer, K.D., and Jaffrey, S.R. (2017). Rethinking m 6 A Readers, Writers, and Erasers. Annu. Rev. Cell Dev. Biol. *33*, 319–342.

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 m6A Promotes Cap-Independent Translation. Cell *163*, 999–1010.

Montemayor, E.J., Curran, E.C., Liao, H.H., Andrews, K.L., Treba, C.N., Butcher, S.E., and Brow, D.A. (2014). Core structure of the U6 small nuclear ribonucleoprotein at 1.7-Å resolution. Nat. Struct. Mol. Biol. *21*, 544–551.

Morais, P., Adachi, H., and Yu, Y.-T. (2021). Spliceosomal snRNA Epitranscriptomics. Front. Genet. *12*, 1–16.

Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J.A. (1983). The U1 small nuclear RNA-protein complex selectively binds a 5′ splice site in vitro. Cell *33*, 509–518.

Munns, T.W., Liszewski, M.K., and Sims, H.F. (1977). Characterization of Antibodies Specific for N6-Methyladenosine and for 7-Methylguanosine. Biochemistry *16*, 2163–2168.

Nakagawa, S., Ip, J.Y., Shioi, G., Tripathi, V., Zong, X., Hirose, T., and Prasanth, K. V. (2012). Malat1 is not an essential component of nuclear speckles in mice. Rna *18*, 1487–1499.

Nojima, T., Hirose, T., Kimura, H., and Hagiwara, M. (2007). The interaction between cap-binding complex and RNA export factor is required for intronless mRNA export. J. Biol. Chem. *282*, 15645–15651.

O'Brown, Z.K., Boulias, K., Wang, J., Wang, S.Y., O'Brown, N.M., Hao, Z., Shibuya, H., Fady, P.E., Shi, Y., He, C., Megason, S.G., Liu, T., and Greer, E.L. (2019). Sources of artifact in measurements of 6mA and 4mC abundance in eukaryotic genomic DNA. BMC Genomics *20*, 1–15.

Pan, T. (2018). Modifications and functional genomics of human transfer RNA. Cell Res. *28*, 395–404.

Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing

complexity in the human transcriptome by high-throughput sequencing. Nat. Genet. *40*, 1413–1415.

Patil, D.P., Chen, C.-K., Pickering, B.F., Chow, A., Jackson, C., Guttman, M., and Jaffrey, S.R. (2016). m6A RNA methylation promotes XIST-mediated transcriptional repression. Nature *537*, 369–373.

Patil, D.P., Pickering, B.F., and Jaffrey, S.R. (2018). Reading m6A in the Transcriptome: m6A-Binding Proteins. Trends Cell Biol. *28*, 113–127.

Pendleton, K.E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B.P., and Conrad, N.K. (2017). The U6 snRNA m6A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. Cell *169*, 824-835.e14.

Perry, R.P., and Kelley, D.E. (1974). Existence of methylated messenger RNA in mouse L cells. Cell *1*, 37–42.

Perry, R.P., Kelley, D.E., Friderici, K., and Rottman, F. (1975). The methylated constituents of L cell messenger RNA: Evidence for an unusual cluster at the 5′ terminus. Cell *4*, 387–394.

Ping, X.L., Sun, B.F., Wang, L., Xiao, W., Yang, X., Wang, W.J., Adhikari, S., Shi, Y., Lv, Y., Chen, Y.S., Zhao, X., Li, A., Yang, Y., Dahal, U., Lou, X.M., Liu, X., Huang, J., Yuan, W.P., Zhu, X.F., Cheng, T., Zhao, Y.L., Wang, X., Danielsen, J.M.R., Liu, F., and Yang, Y.G. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. *24*, 177–189.

Plaschka, C., Newman, A.J., and Nagai, K. (2019). Structural Basis of Nuclear pre-mRNA Splicing: Lessons from Yeast. Cold Spring Harb. Perspect. Biol. *11*, a032391.

Reddy, R., Henning, D., Das, G., Harless, M., and Wright, D. (1987). The capped U6 small nuclear RNA is transcribed by RNA polymerase III. J. Biol. Chem. *262*, 75–81.

Reik, W., Dean, W., and Walter, J. (2001). Epigenetic reprogramming in mammalian development. Science (80-.). *293*, 1089–1093.

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.

Rong, B., Zhang, Q., Wan, J., Xing, S., Dai, R., Li, Y., Cai, J., Xie, J., Song, Y., Chen, J., Zhang, L., Yan, G., Zhang, W., Gao, H., Han, J.D.J., Qu, Q., Ma, H., Tian, Y., and Lan, F. (2020). Ribosome 18S m6A Methyltransferase METTL5 Promotes Translation Initiation and Breast Cancer Cell Growth. Cell Rep. *33*, 108544.

Roost, C., Lynch, S.R., Batista, P.J., Qu, K., Chang, H.Y., and Kool, E.T. (2015). Structure and thermodynamics of N6-methyladenosine in RNA: A spring-loaded base modification. J. Am. Chem. Soc. *137*, 2107–2115.

Roundtree, I.A., Luo, G.Z., Zhang, Z., Wang, X., Zhou, T., Cui, Y., Sha, J., Huang, X., Guerrero, L., Xie, P., He, E., Shen, B., and He, C. (2017). YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. Elife *6*, 1–28.

Ruskin, B., Zamore, P.D., and Green, M.R. (1988). A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. Cell *52*, 207–219.

Ruszkowska, A., Ruszkowski, M., Dauter, Z., and Brown, J.A. (2018). Structural insights into the RNA methyltransferase domain of METTL16. Sci. Rep. 1–13.

Safra, M., Sas-Chen, A., Nir, R., Winkler, R., Nachshon, A., Bar-Yaacov, D., Erlacher, M., Rossmanith, W., Stern-Ginossar, N., and Schwartz, S. (2017). The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature *551*, 251–255.

Sas-Chen, A., Thomas, J.M., Matzov, D., Taoka, M., Nance, K.D., Nir, R., Bryson, K.M., Shachar, R., Liman, G.L.S., Burkhart, B.W., Gamage, S.T., Nobe, Y., Briney, C.A., Levy, M.J., Fuchs, R.T., Robb, G.B., Hartmann, J., Sharma, S., Lin, Q., Florens, L., Washburn, M.P., Isobe, T., Santangelo, T.J., Shalev-Benami,

M., Meier, J.L., and Schwartz, S. (2020). Dynamic RNA acetylation revealed by quantitative cross-evolutionary mapping. Nature *583*, 638–643.

Scarborough, A.M., Flaherty, J.N., Hunter, O. V., Liu, K., Kumar, A., Xing, C., Tu, B.P., and Conrad, N.K. (2021). SAM homeostasis is regulated by CFIm-mediated splicing of MAT2A. Elife *10*, 1–38.

Scherrer, K., Latham, H., and Darnell, J.E. (1963). Demonstration of an unstable RNA and of a precursor to ribosomal RNA in HeLa cells. Proc. Natl. Acad. Sci. U. S. A. *49*, 240–248.

Schibler, U., Kelley, D.E., and Perry, R.P. (1977). Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. J. Mol. Biol. *115*, 695–714.

Schöller, E., Weichmann, F., Treiber, T., Ringle, S., Treiber, N., Flatley, A., Feederle, R., Bruckmann, A., and Meister, G. (2018). Interactions, localization, and phosphorylation of the m 6 A generating METTL3–METTL14–WTAP complex. RNA *24*, 499–512.

Sendinc, E., Valle-Garcia, D., Dhall, A., Chen, H., Henriques, T., Navarrete-Perea, J., Sheng, W., Gygi, S.P., Adelman, K., and Shi, Y. (2019). PCIF1 Catalyzes m6Am mRNA Methylation to Regulate Gene Expression. Mol. Cell *75*, 620-630.e9.

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

Sergiev, P. V., Serebryakova, M. V., Bogdanov, A.A., and Dontsova, O.A. (2008). The ybiN Gene of Escherichia coli Encodes Adenine-N6 Methyltransferase Specific for Modification of A1618 of 23 S Ribosomal RNA, a Methylated Residue Located Close to the Ribosomal Exit Tunnel. J. Mol. Biol. *375*, 291–300.

Shatkin, A.J. (1976). Capping of eucaryotic mRNAs. Cell *9*, 645–653.

Shatkin, A.J., and Manley, J.L. (2000). The ends of the affair: Capping and polyadenylation. Nat. Struct. Biol. *7*, 838–842.

Shen, C., Sheng, Y., Zhu, A.C., Robinson, S., Jiang, X., Dong, L., Chen, H., Su, R., Yin, Z., Li, W., Deng, X., Chen, Y., Hu, Y.C., Weng, H., Huang, H., Prince, E., Cogle, C.R., Sun, M., Zhang, B., Chen, C.W., Marcucci, G., He, C., Qian, Z., and Chen, J. (2020). RNA Demethylase ALKBH5 Selectively Promotes Tumorigenesis and Cancer Stem Cell Self-Renewal in Acute Myeloid Leukemia. Cell Stem Cell *27*, 64-80.e9.

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

Shi, H., Zhang, X., Weng, Y.L., Lu, Z., Liu, Y., Lu, Z., Li, J., Hao, P., Zhang, Y., Zhang, F., Wu, Y., Delgado, J.Y., Su, Y., Patel, M.J., Cao, X., Shen, B., Huang, X., Ming, G. li, Zhuang, X., Song, H., He, C., and Zhou, T. (2018). m6A facilitates hippocampus-dependent learning and memory through YTHDF1. Nature *563*, 249–253.

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 N 6 -Adenosine Methylation and mRNA Degradation Involving METTL16 and YTHDC1. Cell Rep. *21*, 3354–3363.

Shimba, S., Bokar, J.A., Rottman, F., and Reddy, R. (1995). Accurate and efficient N -6-adenosine methylation in spliceosomal U6 small Nucelar RNA by HeLa cell extract in vitro. Nucleic Acids Res. *23*, 2421–2426.

Shimotohno, K., Kodama, Y., Hashimoto, J., and Miura, K.I. (1977). Importance of 5'-terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis. Proc. Natl. Acad. Sci. U. S. A. *74*, 2734–2738.

Shiraki, N., Shiraki, Y., Tsuyama, T., Obata, F., Miura, M., Nagae, G., Aburatani, H., Kume, K., Endo, F., and Kume, S. (2014). Methionine metabolism regulates maintenance and differentiation of human pluripotent stem cells. Cell Metab. *19*, 780–794.

Singh, R., and Reddy, R. (1989). Gamma-monomethyl phosphate: a cap structure in spliceosomal U6 small nuclear RNA. Proc. Natl. Acad. Sci. *86*, 8280–8283.

Śledź, P., and Jinek, M. (2016). Structural insights into the molecular mechanism of the m6A writer complex. Elife *5*, 1–16.

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

Soeiro, E., Birnboim, H.C., and Darnell, J.E. (1966). Rapidly labeled HeLa cell nuclear RNA: II. Base composition and cellular localization of a heterogeneous RNA fraction. J. Mol. Biol. *19*, 362–372.

Soh, Y.Q.S., Mikedis, M.M., Kojima, M., Godfrey, A.K., de Rooij, D.G., and Page, D.C. (2017). Meioc maintains an extended meiotic prophase I in mice. PLOS Genet. *13*, e1006704.

Stoilov, P., Rafalska, I., and Stamm, S. (2002). YTH: A new domain in nuclear proteins. Trends Biochem. Sci. *27*, 495–497.

Stoltzfus, C.M., and Dane, R.W. (1982). Accumulation of Spliced Avian Retrovirus mRNA Is Inhibited in S-Adenosylmethionine-Depleted Chicken Embryo Fibroblasts. J. Virol. *42*, 918–931.

Tang, C., Klukovich, R., Peng, H., Wang, Z., Yu, T., Zhang, Y., Zheng, H., Klungland, A., and Yan, W. (2017). ALKBH5-dependent m6A demethylation controls splicing and stability of long 3'-UTR mRNAs in male germ cells. Proc. Natl. Acad. Sci. U. S. A. *115*, E325–E333.

Tapial, J., Ha, K.C.H., Sterne-Weiler, T., Gohr, A., Braunschweig, U., Hermoso-Pulido, A., Quesnel-Vallières, M., Permanyer, J., Sodaei, R., Marquez, Y., Cozzuto, L., Wang, X., Gómez-Velázquez, M., Rayon, T., Manzanares, M., Ponomarenko, J., Blencowe, B.J., and Irimia, M. (2017). An atlas of alternative splicing profiles and functional associations reveals new regulatory programs and genes that simultaneously express multiple major isoforms. Genome Res. *27*, 1759–1768.

Theler, D., Dominguez, C., Blatter, M., Boudet, J., and Allain, F.H.T. (2014). Solution structure of the YTH domain in complex with N6-methyladenosine RNA: A reader of methylated RNA. Nucleic Acids Res. *42*, 13911–13919.

Vankan, P., McGuigan, C., and Mattaj, I.W. (1990). Domains of U4 and U6 snRNAs required for snRNP assembly and splicing complementation in Xenopus oocytes. EMBO J. *9*, 3397–3404.

Vu, L.P., Pickering, B.F., Cheng, Y., Zaccara, S., Nguyen, D., Minuesa, G., Chou, T., Chow, A., Saletore, Y., Mackay, M., Schulman, J., Famulare, C., Patel, M., Klimek, V.M., Garrett-Bakelman, F.E., Melnick, A., Carroll, M., Mason, C.E., Jaffrey, S.R., and Kharas, M.G. (2017). The N 6 -methyladenosine (m 6 A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. Nat. Med. *23*, 1369–1376.

Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015). N6-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell *161*, 1388–1399.

Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., Gong, Z., Wang, Q., Huang, J., Tang, C., Zou, T., and Yin, P. (2016a). Structural basis of N6-adenosine methylation by the METTL3–METTL14 complex. Nature *534*, 1–15.

Wang, X., Feng, J., Xue, Y., Guan, Z., Zhang, D., Liu, Z., Gong, Z., Wang, Q., Huang, J., Tang, C., Zou, T., and Yin, P. (2016b). Structural basis of N6-adenosine methylation by the METTL3-METTL14 complex. Nature *534*, 575–578.

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

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 N 6 -methyladenosine (m 6 A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep. *18*, 2004–2014.

Watabe, E., Togo-Ohno, M., Ishigami, Y., Wani, S., Hirota, K., Kimura-Asami, M., Hasan, S., Takei, S., Fukamizu, A., Suzuki, Y., Suzuki, T., and Kuroyanagi, H. (2021). m6A-mediated alternative splicing coupled with nonsense-mediated mRNA decay regulates SAM synthetase homeostasis. EMBO J. *40*, 1–19.

Watson, J.D., and Crick, F.H.C. (1953). Genetical implications of the structure of deoxyribonucleic acid. Nature *171*, 964–967.

Wei, C.M., Gershowitz, A., and Moss, B. (1975). Methylated nucleotides block 5′ terminus of HeLa cell messenger RNA. Cell *4*, 379–386.

Wei, C.M., Gershowitz, A., and Moss, B. (1976). 5′-Terminal and Internal Methylated Nucleotide Sequences in Hela Cell mRNA. Biochemistry *15*, 397–401.

Wei, G., Almeida, M., Pintacuda, G., Coker, H., Bowness, J.S., Ule, J., and Brockdorff, N. (2021). Acute depletion of METTL3 implicates N 6 -methyladenosine in alternative intron/exon inclusion in the nascent transcriptome. Genome Res. *31*, 1395–1408.

Wei, J., Liu, F., Lu, Z., Fei, Q., Ai, Y., He, P.C., Shi, H., Cui, X., Su, R., Klungland, A., Jia, G., Chen, J., and He, C. (2018). Differential m 6 A, m 6 A m , and m 1 A Demethylation Mediated by FTO in the Cell Nucleus and Cytoplasm. Mol. Cell *71*, 973-985.e5.

Werner, M., Purta, E., Kaminska, K.H., Cymerman, I.A., Campbell, D.A., Mittra, B., Zamudio, J.R., Sturm, N.R., Jaworski, J., and Bujnicki, J.M. (2011). 2′-O-ribose methylation of cap2 in human: Function and evolution in a horizontally mobile family. Nucleic Acids Res. *39*, 4756–4768.

Wiener, D., and Schwartz, S. (2020). The epitranscriptome beyond m6A. Nat. Rev. Genet. *22*, 119–131.

Wilkinson, M.E., Charenton, C., and Nagai, K. (2020). RNA Splicing by the Spliceosome. Annu. Rev. Biochem. *89*, 359–388.

Will, C.L., and Luhrmann, R. (2011). Spliceosome Structure and Function. Cold Spring Harb. Perspect. Biol. *3*, a003707.

Wojtas, M.N., Pandey, R.R., Mendel, M., Homolka, D., Sachidanandam, R., and Pillai, R.S. (2017). Regulation of m6A Transcripts by the 3′→5′ RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline. Mol. Cell *68*, 374–387.

Worch, R., Niedzwiecka, A., Stepinski, J., Mazza, C., Jankowska-Anyszka, M., Darzynkiewicz, E., Cusack, S., and Stolarski, R. (2005). Specificity of recognition of mRNA 5′ cap by human nuclear cap-binding complex. Rna *11*, 1355–1363.

Wu, S., Romfo, C.M., Nilsen, T.W., and Green, M.R. (1999). Functional recognition 3’ splice site AG by the splicing factor U2AF35. Nature *402*, 832–835.

Xiang, Y., Laurent, B., Hsu, C.-H.H., Nachtergaele, S., Lu, Z., Sheng, W., Xu, C., Chen, H., Ouyang, J., Wang, S., Ling, D., Hsu, P.-H.H., Zou, L., Jambhekar, A., He, C., and Shi, Y. (2017). RNA m6 A methylation regulates the ultraviolet-induced DNA damage response. Nature *543*, 573–576.

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., Wang, X., Ma, H.L., Huang, C.M., Yang, Y., Huang, N., Jiang, G. Bin, Wang, H.L., Zhou, Q., Wang, X.J., Zhao, Y.L., and Yang, Y.G. (2016). Nuclear m6A Reader YTHDC1 Regulates mRNA Splicing. Mol. Cell *61*, 507–519.

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.

Yean, S.L., Wuenschell, G., Termini, J., and Lin, R.J. (2000). Metal-ion coordination by U6 small nuclear RNA contributes to catalysis in the spliceosome. Nature *408*, 881–884.

Yeo, G., Holste, D., Kreiman, G., and Burge, C.B. (2004). Variation in alternative splicing across human tissues. Genome Biol. *5*, 1–15.

Yoon, K.-J., Ringeling, F.R., Vissers, C., Jacob, F., Pokrass, M., Jimenez-Cyrus, D., Su, Y., Kim, N.-S., Zhu, Y., Zheng, L., Kim, S., Wang, X., Doré, L.C., Jin, P., Regot, S., Zhuang, X., Canzar, S., He, C., Ming, G., and Song, H. (2017). Temporal Control of Mammalian Cortical Neurogenesis by m6A Methylation. Cell *171*, 877–889.

Yoshida, H., Park, S.Y., Sakashita, G., Nariai, Y., Kuwasako, K., Muto, Y., Urano, T., and Obayashi, E. (2020). Elucidation of the aberrant 3′ splice site selection by cancer-associated mutations on the U2AF1. Nat. Commun. *11*, 1–7.

Yu, T., Qi, X., Zhang, L., Ning, W., Gao, D., Xu, T., Ma, Y., Knott, J.G., Sathanawongs, A., Cao, Z., and Zhang, Y. (2021). Dynamic reprogramming and function of RNA N 6 -methyladenosine modification during porcine early embryonic development. Zygote 1–10.

Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., Cheng, T., Gao, M., Shu, X., Ma, H., Wang, F., Wang, X., Shen, B., Wang, Y., Feng, X., He, C., and Liu, J. (2018). VIRMA mediates preferential m6A mRNA methylation in 3′UTR and near stop codon and associates with alternative polyadenylation. Cell Discov. *4*, 10.

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.

Zamore, P.D., and Green, M.R. (1991). Biochemical characterization of U2 snRNP auxiliary factor: An essential pre-mRNA splicing factor with a novel intranuclear distribution. EMBO J. *10*, 207–214.

Zhang, H., and Blumenthal, T. (1996). Functional analysis of an intron 3' splice site in Caenorhabditis elegans. RNA *2*, 380–388.

Zhang, B., Arun, G., Mao, Y.S.S., Lazar, Z., Hung, G., Bhattacharjee, G., Xiao, X., Booth, C.J.J., Wu, J., Zhang, C., and Spector, D.L.L. (2012). The lncRNA malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. Cell Rep. *2*, 111–123.

Zhang, M., Zamore, P.D., Carmo-Fonseca, M., Lamond, A.I., and Green, M.R. (1992). Cloning and intracellular localization of the U2 small nuclear ribonucleoprotein auxiliary factor small subunit. Proc. Natl. Acad. Sci. U. S. A. *89*, 8769–8773.

Zhang, Y., Wang, X., Zhang, X., Wang, J., Ma, Y., Zhang, L., and Cao, X. (2019). RNA-binding protein YTHDF3 suppresses interferon-dependent antiviral responses by promoting FOXO3 translation. Proc. Natl. Acad. Sci. U. S. A. *116*, 976–981.

Zhang, Z., Theler, D., Kaminska, K.H., Hiller, M., De La Grange, P., Pudimat, R., Rafalska, I., Heinrich, B., Bujnick, J.M., Allain, F.H.T., and Stamm, S. (2010). The YTH domain is a novel RNA binding domain. J. Biol. Chem. *285*, 14701–14710.

Zhao, X., Yang, Y., Sun, B.F., Shi, Y., Yang, X., Xiao, W., Hao, Y.J., Ping, X.L., Chen, Y.S., Wang, W.J., Jin, K.X., Wang, X., Huang, C.M., Fu, Y., Ge, X.M., Song, S.H., Jeong, H.S., Yanagisawa, H., Niu, Y., Jia, G.F., Wu, W., Tong, W.M., Okamoto, A., He, C., Danielsen, J.M.R., Wang, X.J., and Yang, Y.G. (2014). FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell Res. *24*, 1403–1419.

Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.M., Li, C.J., Vågbø, C.B., Shi, Y., Wang, W.L., Song, S.H., Lu, Z., Bosmans, R.P.G., Dai, Q., Hao, Y.J., Yang, X., Zhao, W.M., Tong, W.M., Wang, X.J., Bogdan, F., Furu, K., Fu, Y., Jia, G., Zhao, X., Liu, J., Krokan, H.E., Klungland, A., Yang, Y.G., and He, C. (2013). ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. Mol. Cell *49*, 18–29.

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, J., Wan, J., Gao, X., Zhang, X., Jaffrey, S.R., and Qian, S.-B. (2015). Dynamic m(6)A mRNA methylation directs translational control of heat shock response. Nature *526*, 591–594.

Zhou, K.I., Shi, H., Lyu, R., Wylder, A.C., Matuszek, Ż., Pan, J.N., He, C., Parisien, M., and Pan, T. (2019). Regulation of Co-transcriptional Pre-mRNA Splicing by m6A through the Low-Complexity Protein hnRNPG. Mol. Cell *76*, 70-81.e9.

10. Publications from my PhD

1. **M. Mendel**, K. Delaney, R. R. Pandey, K-M. Chen, J. M. Wenda, C. B. Vågbø, F. A. Steiner, D. Homolka, R. S. Pillai "Splice site m^{6}A methylation prevents binding of U2AF35 to inhibit RNA splicing". **Cell**, 2021, 184: 3125 - 3142. https://doi.org/10.1016/j.cell.2021.03.062
2. **M. Mendel**, R. Pillai "Nxf3: a middleman with the right connections for unspliced piRNA precursor export". **Genes & Development**, 2019, 33: 1095–1097. doi.org/10.1101/gad.330530.119.
3. **M. Mendel***, K-M. Chen*, D. Homolka, P. Gos, R.R. Pandey, A.A. McCarthy, R.S. Pillai "Methylation of Structured RNA by the m^{6}A Writer METTL16 Is Essential for Mouse Embryonic Development". **Molecular Cell**, 2018; 71: 986-1000. doi.org/10.1016/j.molcel.2018.08.004
4. M.N. Wojtas*, R.R. Pandey*, **M. Mendel***, D. Homolka, R. Sachidanandam, R.S. Pillai "Regulation of m^{6}A Transcripts by the 3'-5' RNA Helicase YTHDC2 Is Essential for a Successful Meiotic Program in the Mammalian Germline". **Molecular Cell**, 2017; 68: 374-387. doi.org/10.1016/j.molcel.2017.09.021

* - equal contribution.