UNIVERSITÉ DE GENÈVE

Archive ouverte UNIGE
https://archive-ouverte.unige.ch

Thèse 2023

Open Access

This version of the publication is provided by the author(s) and made available in accordance with the copyright holder(s).

Exploring the Extended mRNA Cap Structure: Roles of CMTR1 and CMTR2 in Mammalian Development

Dohnalkova, Michaëla

How to cite

DOHNALKOVA, Michaëla. Exploring the Extended mRNA Cap Structure: Roles of CMTR1 and CMTR2 in Mammalian Development. Doctoral Thesis, 2023. doi: 10.13097/archive-ouverte/unige:178769

This publication URL: https://archive-ouverte.unige.ch/unige:178769
Publication DOI: 10.13097/archive-ouverte/unige:178769

© This document is protected by copyright. Please refer to copyright holder(s) for terms of use.

UNIVERSITÉ DE GENÈVE FACULTÉ DES SCIENCES

Section de biologie

Département de Biologie moléculaire & cellulaire Professeur Ramesh S. Pillai

Exploring the Extended mRNA Cap Structure: Roles of CMTR1 and CMTR2 in Mammalian Development

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

Michaela DOHNÁLKOVÁ

de

Ústí nad Orlicí (Czechia)

Thèse N° 237

GENÈVE

2023

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 Mme Michaela DOHNÁLKOVÁ

intitulée :

« Exploring the Extended mRNA Cap Structure: Roles of CMTR1 and CMTR2 in Mammalian Development »

Les Facultés de médecine et des sciences, sur le préavis de Monsieur Ramesh PILLAI, Professeur ordinaire et directeur de thèse (Département de Biologie Moléculaire et Cellulaire), Monsieur Florian STEINER, Professeur associé (Département de Biologie Moléculaire et Cellulaire), Madame Hana CAHOVÁ, Professeure (Institute of Organic Chemistry and Biochemistry of the CAS) et Monsieur Jean-Yves ROIGNANT, Professeur (UNIL, Faculty of Biology and Medicine) autorisent l'impression de la présente thèse, sans exprimer d'opinion sur les propositions qui y sont énoncées.

Genève, le 13 septembre 2023

Thèse - 237 -

Le Doyen
Faculté de médecine

La Doyenne
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 content

Abbreviations

2′-dU	2′-Deoxyuridine
2-5A	2'-5'-oligoadenylate
2D TLC	Two-Dimensional Thin Layer Chromatography
3′ ss	3′ Splice Site
3′-UTR	3′ Untranslated Region
5′TOP	5′ Terminal Oligopyrimidine Tract
5′-UTR	5′ Untranslated Region
5mC	*N5*-methylcytosine (in DNA)
6mA	*N6*-methyladenosine (in DNA)
A	Adenosine
ac^{4}C	*N4*-acetylcytidine
ADAR	Adenosine Deaminase Acting on RNA
AdoMet	S-adenosyl methionine
AGO	Argonaute protein
AGS	Aicardi-Goutières Syndrome
ALKBH	AlkB Homolog, Alpha-Ketoglutarate-Dependent Dioxygenase
A-to-I	Adenosine-to-Inosine (editing in RNA)
C	Cytosine

Cap0	mRNA cap m^7Gppp-RNA
Cap1	mRNA cap m^7Gppp-N_m-RNA
Cap2	mRNA cap m^7Gppp-N_m-N_mRNA
CARDs	Caspase recruitment domains
CBC	Cap binding complex
CBD	Cap Binding Domain
CBP20	Cap-binding protein 20
CBP80	Cap-binding protein 80
CD4	Cluster of Differentiation 4
cDNA	Complementary Deoxyribonucleic Acid
cKO	conditional Knockout
Cm	2'-O-methylcytidine
CMTR	Cap methyltransferase 1 or 2
Cre	Cre recombinase
CTD	Carboxy terminal domain
DCP	Decapping protein
DdRp	DNA-dependent RNA polymerase
DENV	Dengue virus
DHX15	DEAH-box helicase 15
DIV2	Day in vitro 2

dNTP	deoxy-ribonucleoside triphosphate
dsRNA	Double-Stranded RNA
DUSP11	Dual specificity protein phosphatase 11
DXO	Decapping exonuclease
E. coli	Escherichia coli
E	Embryonic day refers to the age of a developing embryo.
eIF2	Eukaryotic Initiation Factor 2
eIF3	Eukaryotic Initiation Factor 3
eIF4A	Eukaryotic Initiation Factor 4A
eIF4E	Eukaryotic initiation factor 4E
eIF4F	Eukaryotic Initiation Factor 4F
eIF4G	Eukaryotic Initiation Factor 4G
EJC	Exon Junction Complex
ELISA	Enzyme-Linked Immunosorbent Assay
Emx1	Empty spiracles homeobox 1
EN	Endonuclease
FTSJ3	FtsJ RNA 2′-O-methyltransferase 3
G	Guanosine
GO terms	Gene Ontology terms
GTase	Guanylyltransferase

GW182	Glycine-Tryptophan protein of 182 kDa
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HEK	Human embryonic kidney 293 cells
HepG2A	human liver cancer cell line
HIV	Human Immunodeficiency Virus
hm5C	*N5*-hydroxymethylcytosine
HPLC	High-Performance Liquid Chromatography
HSV1	Herpes Simplex Virus Type 1
I	Inosine
I^5U	*N5*-Iodouridine
IFIT1	Interferon-induced protein with tetratricopeptide repeats 1
IFN	Interferon
IFNAR	Interferon Alpha/Beta Receptor
IRES	Internal Ribosome Entry Site
ISG	Interferon-stimulated genes
ISG95	Interferon-stimulated gene 95
IUGR	Intrauterine growth restriction
KEGG	Kyoto Encyclopedia of Genes and Genomes
KO	Knockout

L	Large Protein
LARP1	La ribonucleoprotein domain family member 1
LC-MS/MS	Liquid Chromatography-Tandem Mass Spectrometry
LGP2	Laboratory of genetics and physiology 2
LIF	Leukemia Inhibitory Factor
LPS	Lipopolysaccharide
LSM	Like Sm
m^1A	*N1*-methyladenosine
$m^1\Psi$	*N1*-methylpseudouridine
$m^{2,2,7}G$	*N2,2,7*-trimethylguanosine
m^5C	*N5*-methylcytosine
m^5dC	*N5*- methyl-*2′*-deoxycytidine
m^6A	*N6*-Methyladenosine
m^6A_m	*N6*,*2′-O*-dimethyladenosine
m^6dA	*N6*-methyl-deoxyadenosine in DNA
m^7G	*N7*-methylguanosine
MAVS	Mitochondrial antiviral-signalling protein
MDA5	Melanoma Differentiation-Associated Protein 5
MEF	Mouse Embryonic Fibroblasts -
mESC	Mouse Embryonic Stem Cells

METTL	Methyltransferase Like 3, 14 or 16
miRNA	Micro RNA
mRNA	Messenger RNA
mRNP	Messenger Ribonucleoprotein
MS	Mass Spectrometry
mTORC1	Mammalian target of rapamycin complex 1
MvhCre	Mouse vasa homolog-Cre
NCBP1	Nuclear cap-binding protein 1
NCBP2	Nuclear cap-binding protein 2
ncRNA	Non-coding Ribonucleic Acid
NLS	Nuclear localisation signal
Nm	*2'-O*-methyl nucleotide
NMD	Nonsense-mediated mRNA Decay
Nm Seq	*2'-O*-methylation sequencing method
NPC	Nuclear Pore Complex
NSP	Non-structural Protein
NSUN	NOP2/Sun RNA Methyltransferase Family Member 2 or 6
nt	Nucleotide
NUDT2	Nudix Hydrolase 2
OAS	Oligoadenylate Synthase-like Proteins

P	Postnatal day
PABPN1	Poly(A) Binding Protein Nuclear 1
PAMPs	Pathogen Associated Molecular Patterns
PAN	Poly(A) Nuclease 2 or 3
PARN	Poly(A)-Specific Ribonuclease
PB	Polymerase Basic protein (Influenza virus protein)
PBs	Processing bodies
PCIF1	Phosphorylated CTD Interacting Factor 1
PCR	Polymerase Chain Reaction
PD1	Programmed cell death protein 1
PDE	Phosphodiesterase
PGCs	Primordial Germ Cells
piRNA	Piwi-Interacting RNA
PKR	Protein Kinase R
Prm1Cre	Cre recombinase expressed under *Prm1* promoter
PRRs	Pattern Recognition Receptors
PS-G	α-Phosphorothioguanosine
PUS	Pseudouridine Synthases
RADAR	Restriction by an adenosine deaminase acting on RNA
RAM	RNMT-activating mini-protein

RdrA	ATPase part of the RADAR system
RdrB	Adenosine deaminase part of RADAR system
RdRp	RNA-dependent RNA polymerase
RFM	RNA recognition motif
RIG-I	Retinoic Acid-Inducible Gene I
RISC	RNA-Induced Silencing Complex
RLRs	RIG-I-like receptors
RNA pol	RNA polymerase
RNA	Ribonucleic acid
RNase L	Ribonuclease L
RNase T2	Ribonuclease T2
RNGTT	RNA guanylyltransferase and 5′-phosphatase
RNMT	RNA (guanine-7-)methyltransferase
RNu	Small nuclear RNAs U11 or U12
RRACH	purine, purine, adenosine, cytosine, any except guanosine
rRNA	Ribosomal RNA
RT	Reverse transcriptase
RTL-P	Reverse Transcription at Low dNTP concentrations followed by PCR
S. cerevisiae	Saccharomyces cerevisiae
s^2U	*N2*-Thiouridine

s^4U	*N4*-Thiouridine
SAH	S-adenosylhomocysteine
SAM	S-adenosyl methionine
siRNA	Small interfering RNA
SMG6	Suppressor with Morphogenetic Effect on Genitalia 6
SMG7	Suppressor with Morphogenetic effect on Genitalia 7
Snhgene	snoRNA host gene
snoRNAs	Small nucleolar RNAs
snRNAs	Small nuclear RNAs
ssRNA	Single-stranded RNA
STAT	Signal Transducer and Activator of Transcription
Stra8Cre	Cre recombinase expressed under *Stra8* promoter
TAP/NXF1	Tip-associated Protein/Nuclear Export Factor 1
THO	Hpr1p, Tho2p, Mft1p, and Thp2p proteins Complex
TLC	Thin Layer Chromatography
TLR	Toll-Like Receptor
TOP	Terminal Oligopyrimidine
TREX	Transcription/export complex
TRIF	TIR-Domain-Containing Adapter-Inducing Interferon-β
TRMT	tRNA methyltransferases

tRNA	Transfer RNA
TSS	Transcription Start Site
U	Uridine
U1 snRNP	U1 small nuclear ribonucleoprotein
U2 snoRNA	U2 Small Nuclear RNA, a component of the major spliceosome
U2AF(65)	U2 Small Nuclear RNA Auxiliary Factor 1
Um	*2′-O*-methyluridine
UTR	Untranslated region
VasaCre	Cre recombinase expressed under *Vasa* promoter
VP	Viral Protein
VSV	Vesicular Stomatitis Virus
WT	Wild type
WW	Tryptophan residues domain
XIST	X-inactive specific transcript
XRN1	5′ to 3′ Exoribonuclease 1
YSPTSPS	tyrosine, serine, proline, threonine, serine, proline, serine
ZIKV	Zika virus
Zp3Cre	Cre recombinase expressed under *Zp3Cre* promoter
Ψ	Pseudouridine

Abstract (EN)

Messenger RNA (mRNA) is transcribed from the coding DNA sequences by the RNA polymerase II. The nascently synthesised mRNA, called pre-mRNA, undergoes a series of maturation steps consisting of 5′ capping, removing non-coding sequences in the process of splicing and finally, 3′ end cleavage and polyadenylation. In mammals, 5′ cap structure includes template-independent terminal *N7*-methylguanosine (m^7G), connected to mRNA via an unusual 5′ to 5′ triphosphate linkage and followed by the two first transcribed nucleotides, which are further modified with *2′-O*-methyl (N_m) marks and, in the case of the first nucleotide being A, additionally *N6*-methyladenosine (m^6A). The formation of the m^7G cap is mediated by two enzymes: RNGTT and RNMT-RAM. N_m marks are added to the first and second nucleotide by the cap-specific RNA methyltransferases CMTR1 and CMTR2, forming cap1 and cap2 structures, respectively. Finally, if the first nucleotide is A, the PCIF1 enzyme deposits the m^6A mark. While virtually all mRNA transcripts have cap1 modification, only approximately 50% have cap2. In addition to protecting mRNA from degradation and enhancing translation, these marks are essential components of the innate immune system, as uncapped transcripts are detected by the cytoplasmic sensors of foreign RNA, triggering interferon response and apoptosis.

The aim of my PhD was to further characterise the physiological and molecular functions of the mRNA cap1 and cap2 structure. Using a combination of mouse models and cell culture models, we show that both CMTR1 and CMT2 are vital for embryogenesis, with both mutant embryos dying around embryonic day E7.5, before the organogenesis stage. Sequencing of E6.5 and E7.5 mutant embryos showed that CMTR1 and CMTR2 regulate non-overlapping subset of genes indicating their separate functions. Interestingly, for both mutants, *Cmtr1* and *Cmtr2*, this occurs without any activation of the innate immune response, suggesting that the functions of N_m extend beyond simply marking cellular RNAs as self. Conditional depletion of CMTR1 in male germ cells results in complete infertility, while only some females are infertile. Depletion of CMTR1 in fully developed organs, like the liver, leads to chronic activation of interferon expression. Interestingly, among commonly dysregulated genes in CMTR1 Knocked-out (KO) backgrounds are ribosomal genes overlapping with 5′ *TOP* transcripts and snoRNA host genes.

This investigation offers a comprehensive examination of the roles of CMTRs in development and in selected organs and illuminates their complex functions behind protecting from triggering autoimmune reactions.

Abstract (FR)

ARN messager (ARNm) est transcrit à partir des séquences d'ADN codantes par l'ARN polymérase II. Le nouvellement synthétisé ARNm, appelé pré-ARNm, subit une série d'étapes de maturation comprenant la coiffe 5′ (5′ capping), l'élimination des séquences non codantes lors du processus d'épissage, et enfin la clivage en 3′ et la polyadénylation. Chez les mammifères, la structure de la coiffe 5' comprend une *N7*-méthylguanosine (m^7G) indépendante du modèle, reliée à l'ARNm via une liaison triphosphate inhabituelle 5′ à 5′, suivie des deux premiers nucléotides transcrits, qui sont ensuite modifiés par des marques 2′-O-méthyl (Nm) et, dans le cas où le premier nucléotide est A, par l'ajout de la *N6*-méthyladénosine (m^6A). La formation de la coiffe m^7G est médiée par deux enzymes : RNGTT et RNMT-RAM. Les marques Nm sont ajoutées au premier et au deuxième nucléotide par les ARN méthyltransférases spécifiques de la coiffe, CMTR1 et CMTR2, formant respectivement les structures cap1 et cap2. Enfin, si le premier nucléotide est A, l'enzyme PCIF1 dépose la marque m^6A. Alors que pratiquement tous les transcrits d'ARNm ont la modification cap1, seul environ 50 % ont la cap2. En plus de protéger l'ARNm de la dégradation et d'améliorer la traduction, ces marques sont des composants essentiels du système immunitaire inné, car les transcrits non coiffés sont détectés par les capteurs cytoplasmiques de l'ARN étranger, déclenchant la réponse de l'interféron et l'apoptose.

L'objectif de ma thèse de doctorat était de caractériser davantage les fonctions physiologiques et moléculaires de la structure cap1 et cap2 de l'ARNm. En utilisant une combinaison de modèles de souris et de cultures cellulaires, nous montrons que CMTR1 et CMTR2 sont tous deux essentiels pour l'embryogenèse, car les embryons mutants meurent tous deux vers le jour embryonnaire E7.5, avant le stade de l'organogenèse. Le séquençage des embryons mutants E6.5 et E7.5 a montré que CMTR1 et CMTR2 régulent un sous-ensemble de gènes non chevauchants, indiquant leurs fonctions distinctes. Fait intéressant, pour les deux mutants Cmtr1 et Cmtr2, cela se produit sans aucune activation de la réponse immunitaire innée, suggérant que les fonctions de Nm vont au-delà de simplement marquer les ARN cellulaires comme étant du soi. La déplétion conditionnelle de CMTR1 dans les cellules germinales mâles entraîne une infertilité totale, tandis que seules certaines femelles sont infertiles. La déplétion de CMTR1 dans des organes entièrement développés, comme le foie, entraîne une activation chronique de l'expression d'interféron. Fait intéressant, parmi les gènes couramment dysrégulés dans les contextes de déficience de CMTR1 se trouvent des gènes ribosomaux chevauchant avec les transcrits 5' TOP et les gènes hôtes des snoARN.

Cette étude offre un examen complet des rôles des CMTRs dans le développement et dans certains organes sélectionnés, et éclaire leurs fonctions complexes de protection contre le déclenchement de réactions auto-immunes.

Acknowledgement

I am really thankful for the fact that I was born in Europe at the end of the 20th century. Just this allowed me to use my potential regardless of my family′s economic status, or whether I was coming from the city or village. This fact of growing up in Czechia offered me the possibility to study whatever I wanted and develop into the person who I am without any burdens, such as how to pay for my studies later or whether my education would be financially possible or salary later good enough to feed the whole family.

Science, often unnoticed, shapes our everyday lives in ways we often take for granted. I am profoundly grateful for the relentless efforts of the many scientists whose discoveries have made our lives better. Thanks to these advancements, I was able to recover from a bacterial infection and "survive" three different of pneumonia, among other illnesses. Vaccines protected me from numerous serious diseases, and the advancements in healthcare, together with scientific discoveries, gifted me an extra cherished year with my grandmother. These experiences, among countless others, have highlighted the fundamental role of science in not only maintaining our physical but also mental well-being and also supporting our overall resilience and adaptability in life′s challenging times.

I need to thank my high school teachers, who showed me how interesting biology and chemistry might be. My curiosity led me to the field of genetics, and later to the lab of Štěpánka Vaňáčová. As a naive 2nd-year bachelor′s student, I entered the lab, only to emerge as an individual with a deep-rooted passion for science. I am profoundly grateful for the stimulating environment that lab provided, significantly shaping my subsequent journey. The lab taught me how to speak English, not to be shy, and started my path on how to present data, but mainly though me how to ask questions, how to pipette and how to design and use proper controls. None of those was perfect when I started PhD, but they gave me a very strong background on how to independently start my PhD that I did not need to feel alone and lost in my beginnings.

My four-years, or 1,466-days PhD journey, which represents 14% of my life to date, was transformative. While the lab of Štěpánka Vaňáčová unveiled the intricate world of RNA and ignited my passion for science, Ramesh Pillai′s lab was instrumental in facilitating my self-discovery. The experience taught me how to work under pressure, manage my time efficiently, make presentations, and gave me an experience how to write scientific proposals. The lab taught me how important it is to work in a nice environment with enthusiastic people, support each other throughout the process.

It also honed my ability to write scientific proposals. I am immensely proud to have been awarded the Swiss Excellence Fellowship as part of this training. This fellowship not only provided funding but also presented me with the opportunity to network with individuals from diverse fields and nationalities while discovering the beauty of Switzerland.

A compulsory part of PhD at UNIGE was to pass the Thesis Advisory Committee meeting. My meeting was motivating and I really enjoyed discussing my project with you. Thank you, Mary Anne O´Connell, Françoise Stutz, and Ueli Schibler.

Thank you to all my colleagues and friends from Pillai lab for their valuable guidance, support, and collaboration. It was fun to spend time with you outside of the lab, visiting Switzerland together, sharing experiences and tastes from various cuisines and traditions and having coffee breaks together. I am really sorry to all of you that my lactose intolerance was transmitted to most of you! Despite this, I hope you continue to enjoy our shared moments, albeit without milk!

I am also thankful to all of you who invested your time in reading all my abstracts, CVs, motivation letters, and this thesis. All your criticisms are highly appreciated!

I am grateful for the chance to gain external perspectives on science through various meetings and conferences. My heartfelt thanks to the organisers, speakers, and attendees of events such as NCCR RNA & Disease, Aegean Conference The Long and the Short of Non-coding RNAs, and Gene Regulation Workshop. Each encounter at these events fuelled my enthusiasm, spurred creativity, and expanded my network of scientific enthusiasts. Nevertheless, among mentioned, I need to emphasise the importance of NCCR RNA & Disease. All events organised were just perfect, and those events not only allowed me to enrich my scientific network but also showed me how important are scientific collaboration and building communities.

As the network showed me how big an impact organising things can bring, I joined the PhD Scimed organising team of 2023. Unfortunately, as it was towards the end of my PhD journey, those were really hard times to find time and passion for meetings and tasks outside of finishing my paper and PhD. Nevertheless, it allowed me to give back to the scientific community. Though it was a challenging period, the experience was highly enriching, and I am thankful for it. And thank you all, committee guys; it was a really great event!

As a nod to my heritage, I acknowledge the importance of my native language in shaping my identity. I am grateful to the organisers of the Slovak and Czech Meetup in Zurich in 2021, which provided an avenue for me to interact in my native language. I was privileged

to join the organising team after the event, helping to cultivate a thriving networking platform for Czechs and Slovaks in Switzerland. I was always looking forward to meeting my friends and interacting with other Czechs. As much as the events we organised were cheering me up, I am sure they energised someone else as well. Thank you CometX guys, all the time spent with you organising the event, being at an event together and later after the event and having fun after were completely worth it and kept me going in the hard times!

As other activities supporting my mental health during PhD were my bouldering and board games sessions. Bouldering helped my back to stretch after long days, and board games fuelled my creativity and competitiveness. I am really grateful that I joined those two groups! Over time, you became my friends!

Last but not least, I would like to extend my heartfelt thanks to everyone, my friends, and colleagues from the lab and outside who provided me with a safe space to express my frustrations. I hope that I managed to give back at least a bit of the support that I received. Thank you for being there, always ready to discuss and listen to both life's and lab's challenges. Your patience and understanding served as an invaluable support system, constantly pushing me forward. Please remember, you are always welcome to approach me whenever you need.

In conclusion, I extend my heartfelt appreciation for the countless experiences, valuable lessons, and friendships made along my academic journey. This journey was not solely about science, but also about forming connections and giving back to the community, which has made it all the more enriching.

Thank you all and I hope we will keep in touch!

Introduction

Regulation of gene expression

One of the fundamental principles of molecular biology is the central dogma theory, first articulated by Francis Crick in his lecture in 1958 (Cobb, 2017; Crick, 1970). The original model of the central dogma depicted a straightforward flow of genetic information: a single gene, encoded in DNA, is transcribed into one messenger RNA (mRNA) molecule, which is subsequently translated into a single protein (**Figure 1**). However, the landscape of molecular biology was dramatically reshaped with the discovery of mRNA splicing in the late 1970s (Berget et al., 1977; Chow & Broker, 1978). This revelation illustrated that a single gene could give rise to multiple distinct proteins, thereby adding a layer of complexity to the central dogma. This phenomenon, known as alternative splicing, enables a single gene to produce a diverse set of mRNA molecules, each of which can be translated into different proteins (**Figure 1**).

The regulation of gene expression controls the timing, quantity, and location of effecting molecule, protein or RNA presence. This complex process involves various mechanisms operating at the DNA, RNA, and protein levels (**Figure 1**)

DNA, composed of four canonical nucleotides (Adenine, Thymine, Cytosine, and Guanine), is not "naked" within cells but instead forms a compacted structure known as chromatin through interactions with proteins and RNA molecules. Histones, a primary component of DNA condensation, play a crucial role in this process. Additionally, DNA can undergo modifications, such as cytosine methylation, resulting in *5*-methylcytosine (^{5}mC), which serves as an epigenetic mark associated with gene silencing.

The regulation of gene expression extends beyond DNA modifications. After their synthesis, proteins can undergo post-translational modifications, which involve adding or removing chemical groups to specific amino acid residues. These modifications, such as phosphorylation, methylation, or acetylation, can profoundly impact protein stability, activity, and localisation with other molecules. Another type of modification, ubiquitination marks proteins for degradation. Interestingly, histones themselves can undergo various post-translational modifications that modulate gene expression. These histone marks include acetylation, methylation, phosphorylation, and ubiquitination. For instance, histone acetylation is associated with gene activation, as it relaxes the chromatin structure, allowing easier access of transcriptional machinery to the DNA. On the other hand, histone methylation can have

diverse effects depending on the specific site and degree of methylation. For example, trimethylation of histone H3 at lysine 4 (H3K4me3) is generally associated with gene activation, while trimethylation at lysine 9 (H3K9me3) is linked to gene repression.

By integrating DNA and protein modifications, the field of epigenetics explores how these intricate molecular mechanisms collectively regulate gene expression at both the chromatin and protein levels. The dynamic interplay between DNA modifications, histone marks, and protein post-translational modifications orchestrates a complex gene regulatory network in cells. Lastly, regulation of gene expression can be modulated by changes in RNA molecules, encompassing nucleotide editing and RNA modifications. These alterations can affect alternative splicing, localisation, translation, and stability of RNA molecules or directly

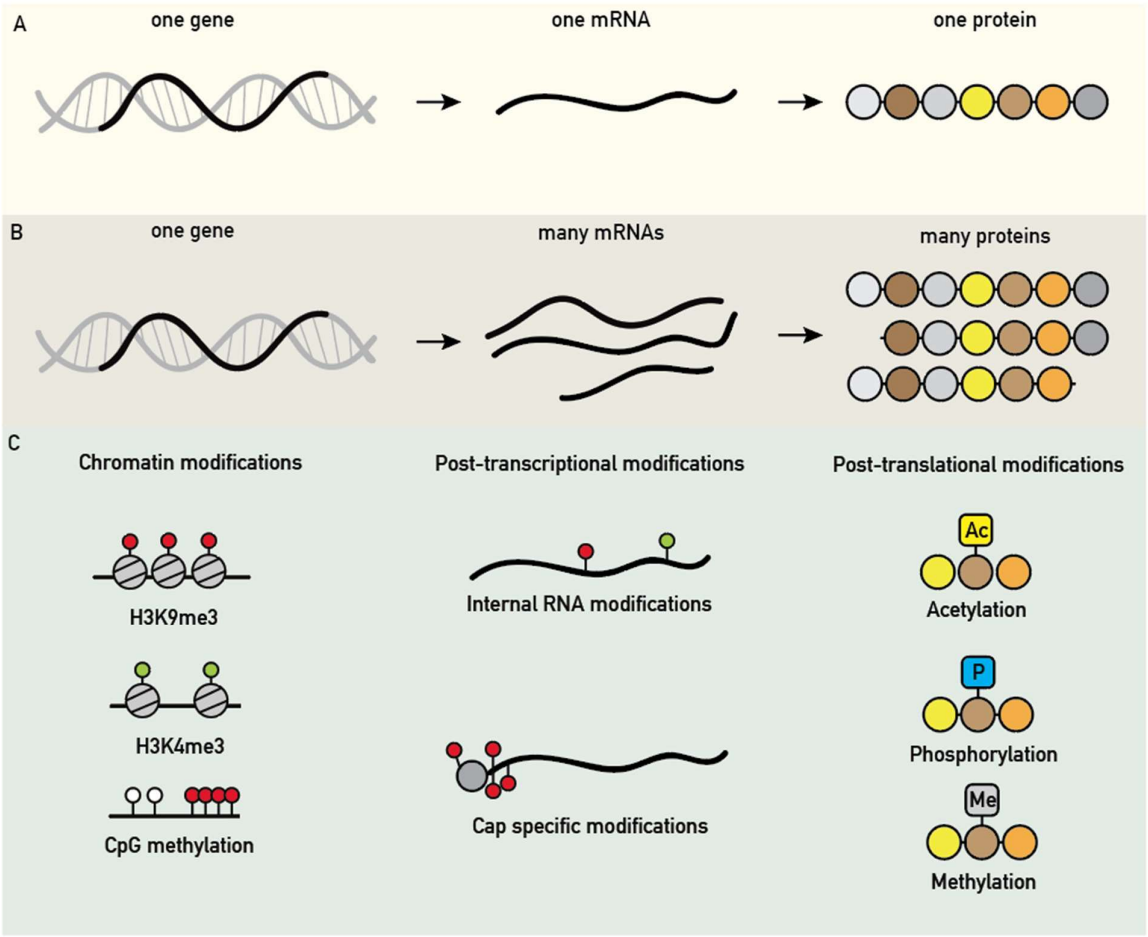


Figure 1 - Central dogma of molecular biology. A) One gene encodes for one mRNA and one protein. B) Many mRNAs can originate from one gene for example, as a result of alternative splicing, and thus one gene can result in many proteins. C) DNA, RNA and protein molecules can be modified. AC – acetylation, P – phosphorylation, Me – methylation, H3K9me3 – tri-methylation of histone three at lysine

9, H3K4me3 – trimethylation of histone 3 at lysine 4, CpG methylation – methylation of cytosine at CpG islands

influence transcription. The field study relating to RNA modifications is called epitranscriptomics. Interestingly, while mutations in DNA are typically irreversible, modifications that lead to alternative gene expressions are reversible. This flexibility allows for an extensive range of adaptability in response to internal and external stimuli.

mRNA processing

The life of mRNA begins with its transcription and concludes with its degradation. The so

mRNA transcription and transcription termination

The process of mRNA transcription is the first step in gene expression, during which the genetic information stored in DNA is converted into a single-stranded mRNA molecule. In case of mRNA, this process is carried out by an enzyme called RNA polymerase II (RNA pol II), which binds to a specific DNA sequence known as the promoter region. Upon binding, the RNA polymerase unwinds the DNA double helix to allow access to other transcriptional machinery, a process often facilitated by chromatin and nucleosome remodelling. As the RNA polymerase moves along the DNA, it synthesises a complementary mRNA strand using the DNA as a template. It adds ribonucleotides that are complementary to the DNA bases, pairing adenine (A) in DNA with uracil (U) in RNA, thymine (T) in DNA with adenine (A) in RNA, cytosine (C) in DNA with guanine (G) in RNA, and guanine (G) in DNA with cytosine (C) in RNA.

One of the RNA pol II regulations is through its carboxy-terminal domain (CTD) localised at its biggest subunit. CTD consist of repetitions of the heptapeptide of YSPTSPS. This sequence is repeated multiple times, usually around 52 times in humans, resulting in a long unstructured tail. The CTD plays a critical role in coordinating various stages of transcription, including initiation, elongation, and termination. It serves as a platform for recruiting and assembling various factors involved in RNA processing and transcriptional regulation. The CTD undergoes dynamic phosphorylation and dephosphorylation events of its heptapeptide repeats at specific residues, primarily at serine 2 and serine 5, during different stages of transcription. For example, genome-wide occupancy profiles for all CTD phosphorylation marks showed that serine 5 and serine 7 are phosphorylated during transcription initiation, whereas tyrosine 1, serine 2, and threonine 4 phosphorylation signals

increase during the transcription towards the 3′ end and polyadenylation site (Heidemann et al., 2013).

Once the RNA polymerase reaches a cleavage and polyadenylation signal, often represented by the canonical AAUAAA sequence in humans, the polymerase slows down but continues the transcription. Cleavage and Polyadenylation Specificity Factor (CPSF) recognises the transcribed polyadenylation signal on newly-synthesised mRNA, which attracts other factors necessary for polyadenylation. The nascent transcript is cleaved downstream of the polyadenylation signal, and poly(A) polymerase adds a series of adenine (A) nucleotides to the 3′ end of the mRNA at the cleavage site. During this process, a poly(A) tail consisting of dozens of adenine nucleotides, with a median of 50–100 nt (Chang et al., 2014), is added to the 3′ end of the mRNA molecule, which helps enhance stability, promote nuclear export, and facilitate translation initiation. The dissociation of RNA pol II from the DNA is facilitated by the exonuclease Xrn2, which degrades the remaining RNA still attached to the polymerase.

The majority of human genes can also contain alternative polyadenylation signals. The choice of polyadenylation site determines the length of the mRNA and affects its stability, localisation, and translational efficiency. Alternative polyadenylation can produce mRNA isoforms with different 3′ untranslated regions (3′UTRs), which can contain different regulatory elements, such as microRNA binding sites, leading to the differential expression of genes. Regulation of alternative polyadenylation can occur in response to changes in developmental stage, environmental conditions, and disease states and can affect gene expression and cellular function. Dysregulation of alternative polyadenylation has been linked to several human diseases, such as cancer and neurodegenerative disorders, highlighting the importance of this process in maintaining normal cellular function.

Finally, during transcription, the mRNA is processed in a series of molecular events that occur simultaneously with mRNA transcription (co-transcriptionally), shaping the maturation and functionality of the nascent mRNA molecule. These processes include mRNA capping, splicing, and RNA editing, where specific nucleotides within the mRNA are modified or replaced. Alternative splicing, a co-transcriptional process that generates multiple mRNA variants from a single gene by splicing together different combinations of exons, enhances protein diversity and complexity within the cell.

mRNA splicing and alternative splicing

mRNA splicing is a crucial process in gene expression that involves the removal of the non-coding intronic regions of pre-mRNA and the joining the exonic coding regions together by the spliceosome. Splicing is a highly dynamic process, which can be regulated to result in multiple mRNA isoforms in alternative splicing.

The spliceosome is a dynamic macromolecular complex composed of five small nuclear ribonucleoproteins (snRNPs), protein complexes and small nuclear RNAs (snRNAs). The snRNAs, including U1, U2, U4, U5, and U6, play pivotal roles in splicing. They recognise the intron-exon boundaries and catalyse the intron removal and exon ligation reactions, ensuring the correct splicing of pre-mRNA (Matera & Wang, 2014; Will & Lührmann, 2011).

The Exon Junction Complex (EJC), another important player in the mRNA splicing process, is a multi-protein complex that is deposited onto mRNA at the exon-exon junctions during the splicing process. It participates in multiple facets of mRNA metabolism, including nuclear export, surveillance, localisation, and translation. The EJC works in concert with the spliceosome to ensure accurate splicing, thereby maintaining the integrity of the genetic message (Asthana et al., 2022; H. Martin et al., 2022; Schlautmann & Gehring, 2020).

Alternative splicing widely increases the diversity of the proteome. One example of a differentially spliced gene is Titin. Titin is a giant protein found in muscle cells that functions as a molecular spring, contributing to muscle elasticity and passive tension. The TTN gene encoding Titin undergoes extensive alternative splicing, giving rise to multiple isoforms that differ in size and have particular mechanical properties. Different Titin isoforms are expressed in various muscle types and during muscle development, with longer, more compliant isoforms found in cardiac muscle and shorter, stiffer isoforms in skeletal muscle (Labeit & Kolmerer, 1995).

Another, layer of complexity to our understanding of genes is added by the alternative splicing by expanding the coding capacity of our genomes. Even though humans have approximately 20,000 coding genes, we have around 180,000 different mRNAs (Kersey et al., 2018). It seems that only 80,000 mRNA isoforms can produce sequence-distinct protein isoforms (Ezkurdia et al., 2015). Nevertheless, it is still unclear how many genes have only one isoform, called a single major isoform. The estimates vary from 28% (Tress et al., 2017), 30% (Ezkurdia et al., 2015) to less than 50% (Tapial et al., 2017) of all human genes having a single major. Interestingly, gene expression patterns between the same organs in different organisms seem to be conserved, but their alternative splicing differs (Merkin et al., 2012).

mRNA export

Fully processed mRNAs, capped, spliced, and polyadenylated, are exported from the nucleus to the cytoplasm, where they can be translated. In order to avoid exporting faulty mRNA, this process is highly regulated and involves various protein factors and complexes. One such complex is the exon junction complex (EJC), deposited on spliced mRNAs at exon-exon junctions during splicing. The EJC plays several roles in mRNA metabolism, including mRNA export, translation, and nonsense-mediated mRNA decay (NMD) (Asthana et al., 2022; H. Martin et al., 2022; Schlautmann & Gehring, 2020). The EJC recruits the TREX (transcription-export) complex for mRNA export, which is responsible for coupling transcription with mRNA export. TREX's more extensive components, such as the THO complex, Aly/REF, and the mRNA export receptor TAP/NXF1, help to facilitate mRNA export through the nuclear pore complex (NPC). The EJC, along with other factors and complexes, ensures that only mature, properly processed mRNAs are exported to the cytoplasm, thus maintaining the fidelity and efficiency of gene expression (Asthana et al., 2022; Gromadzka et al., 2016; H. Martin et al., 2022).

mRNA translation

Once mRNA reaches the cytoplasm, the information encoded in mRNA is converted into a protein through a process called translation. Translation occurs on ribosomes, large molecular machines composed of ribosomal RNA (rRNA) and ribosomal proteins. The ribosomes provide the platform for the synthesis of proteins by facilitating the interaction between mRNA and transfer RNA (tRNA) (Dever et al., 2018; Moore & Steitz, 2011; Simonović & Steitz, 2009; Steitz, 2008).

tRNA molecules play a crucial role in translation by bringing the appropriate amino acids to the ribosome. Each tRNA molecule has a specific sequence of nucleotides corresponding to a particular amino acid. The tRNA recognises codon, a three-nucleotide sequence, on the mRNA through its complementary anticodon sequence. The complementary base pairing between the tRNA's anticodon and the mRNA's codon ensures the amino acid's accurate placement in the growing protein chain. The translation is divided into three main stages: initiation, elongation, and termination (Orellana et al., 2022; Phizicky & Hopper, 2010).

During initiation, the small ribosomal subunit binds to the mRNA, typically at the 5′ cap structure, and scans along the mRNA until it encounters the start codon (usually AUG). In

eukaryotes, this start codon is often part of a larger sequence known as the Kozak consensus sequence (Kozak, 1981). The sequence aids in accurately identifying the RNA's 5′ cap structure and positioning the start codon for the small ribosomal subunit.

The pairing of the start codon with the anticodon of the initiator tRNA, which carries the amino acid methionine, marks the next step in initiation. The large ribosomal subunit then joins this assembly, completing the formation of the translation initiation complex. With the ribosome correctly positioned at the start codon - an accuracy significantly aided by the Kozak sequence - the translation process is primed to transition into the elongation phase, where the nascent protein chain begins synthesising.

As the elongation phase starts, distinct aminoacyl-tRNAs, each carrying a specific amino acid, come into play. These tRNAs complement the mRNA codons that the ribosome is processing. The ribosome acts as a catalyst, forming peptide bonds that connect these amino acids, thus building up the emerging polypeptide chain (Dever et al., 2018).

Termination occurs when the ribosome encounters a stop codon (UAA, UAG, or UGA) on the mRNA. These codons do not code for any amino acids and are recognised by release factors, which facilitate the release of the completed polypeptide chain from the ribosome.

As the nascently synthesised peptide emerges from the ribosome, it begins to fold into its native structure co-translationally, often with the assistance of molecular chaperones. After the ribosomal subunits dissociate from the mRNA, the protein might continue to fold post-translationally, sometimes needing additional chaperone assistance or undergoing specific modifications to reach its final functional form (Ellgaard et al., 2016). The ribosomal subunits then dissociate from the mRNA, and the newly synthesised protein can undergo post-translation-folding and post-translational modifications to become functional (Dobson, 2003; Keenan et al., 2021).

Overall, mRNA translation is a highly regulated and complex process essential for producing proteins and properly functioning cells. Dysregulation of translation as well as protein miss folding can lead to numerous diseases.

Translation initiation

In eukaryotes, there are two main types of translation initiation: cap-dependent and cap-independent. Cap-dependent initiation, the most prevalent mechanism, begins with the interaction between the mRNA and specific protein complexes.

Immediately after the mRNA is transcribed in the nucleus, the Cap Binding Complex (CBC) binds to the 5′ cap structure. This binding protects the mRNA from degradation, facilitates its export from the nucleus to the cytoplasm, and can even aid in the early stages of translation initiation (Gonatopoulos-Pournatzis & Cowling, 2014).

Once the mRNA is in the cytoplasm, there is a transition during which the eukaryotic initiation factor 4F (eIF4F) complex replaces the CBC at the 5′ cap structure. The eIF4F complex, which includes the proteins eIF4E, eIF4G, and eIF4A, is responsible for recruiting the small ribosomal subunit to the mRNA, marking the onset of the translation process (Gonatopoulos-Pournatzis & Cowling, 2014).

In contrast to cap-dependent initiation, cap-independent initiation does not rely on mRNA's 5′ cap structure. Instead, it makes use of specific RNA elements located within the mRNA.

The most common mRNA element located within the 5′UTR is internal ribosome entry sites (IRES). IRES involves the direct recruitment of ribosomes to an internal region of the mRNA, bypassing the need for a 5′ cap. The IRES-mediated translation is typical for some viruses, but also some of cellular mRNAs carry their IRES elements which allows them to initiate translation in a cap-independent manner (Hoshi et al., 1984; Y. Yang & Wang, 2019).

Notably, there is evidence that *N6*-methyladenosine modification in the 5′ UTR might also be involved in cap-independent translation. m^6A in the 5′ UTR is recognised by a protein called eIF3 which can further recruit the 43S complex to initiate translation (Meyer et al., 2015).

mRNA degradation

mRNA degradation is a crucial process that regulates gene expression by controlling the abundance and stability of mRNAs. The degradation mechanisms differ between the nucleus and the cytoplasm and between the mRNA molecule's 5′ and 3′ ends. In the nucleus, mRNA degradation primarily occurs through the exosome complex in the 3′ to 5′ direction, degrading aberrant mRNAs that fail to process properly, such as capping, splicing, or polyadenylation. This quality control mechanism ensures that only mature mRNAs are exported to the cytoplasm.

In the cytoplasm, mRNA degradation typically involves two main pathways: deadenylation-dependent and endonucleolytic cleavage. Deadenylation-dependent decay starts with removing the protective poly(A) tail at the 3′ end of the mRNA by deadenylase enzymes.

This is followed by decapping the 5′ end, exposing the mRNA to exonucleases that degrade the molecule from both ends. Alternatively, endonucleolytic cleavage can occur, in which specific endonucleases cleave the mRNA internally, generating fragments that are subsequently degraded by exonucleases (Beelman & Parker, 1995; Garneau et al., 2007).

These mRNA degradation pathways ensure proper gene expression by maintaining the appropriate levels of mRNAs in the cell and removing any defective or unnecessary transcripts. This fine-tunes protein synthesis and enables rapid cellular responses to changes in environmental conditions or cellular signals.

An example of adaptation to changes via mRNA degradation is modulation of iron uptake in the cells. Under low iron conditions, IRP1 binds to iron-responsive elements (IREs) in the mRNA of Transferrin encoded by TFR1, stabilising the mRNA and promoting the production of TFR1 protein, which increases iron uptake into the cell. Conversely, when iron levels are high, IRP1 loses its ability to bind to IREs due to the incorporation of iron, leading to the degradation of TFR1 mRNA and a decrease in iron uptake (Casey et al., 1989; Müllner & Kühn, 1988; Theil, 1994).

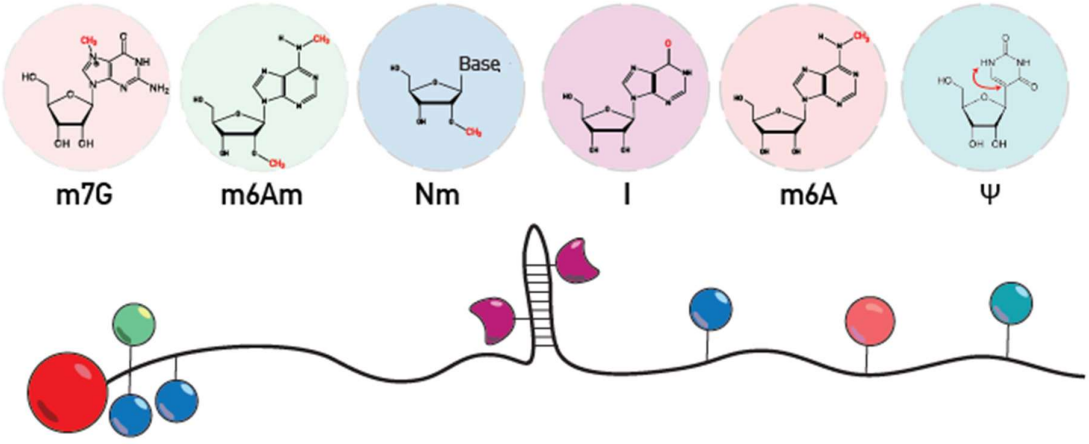


Figure 2 - mRNA modifications in mammals. *N*^{7}-methylguanosine (m^{7}G), *2′-O*-methylation (N$_m$), *N6,2′-O*-dimethyladenosine (m^{6}A$_m$) are specific for the 5′end of the mRNA, whereas *N6*-methyladenosine (m^{6}A), 2′-O-methylation (N$_m$), inosine (I) and Pseudouridine (Ψ) are deposited in its body.

RNA modifications and editing

RNA consists of four canonical nucleotides – adenosine (A), cytosine (C), guanosine (G) and uracil (U) which are incorporated into the RNA molecules during transcription by the RNA polymerases. Those can be further modified by adding or removing some chemical groups. In the 50s, some RNA modifications could be identified using thin-layer chromatography (Davis & Allen, 1957; Dunn, 1959; Littlefield & Dunn, 1958). Up to today, over 170 different modifications have been identified, mostly on non-coding RNAs like tRNAs or rRNAs (Boccaletto et al., 2018). In addition to non-coding RNAs, also mRNA can be extensively chemically modified.

Identification of mRNA modifications was possible only after the establishment of purification of polyadenylated RNAs (Gros et al., 1959; S Brenner & Meselson, 1957). The most common modifications are cap-specific *N*^{7}-methylguanosine (m^{7}G), *2′-O*-methylation (N$_m$), *N6,2′*-O-dimethyladenosine (m^{6}A$_m$) (Gonatopoulos-Pournatzis & Cowling, 2014) and internal modifications *N6*-methyladenosine (m^{6}A), *2′-O*-methylation (N$_m$), inosine (I), Pseudouridine (Ψ), *5*-hydroxymethylcytosine (hm^{5}C) (Lan et al., 2020), *N4*-acetylcytidine (ac^{4}C) (**Figure 2**) (Thalalla Gamage et al., 2021). The comprehensive study and understanding of these RNA modifications and their biological implications fall under a rapidly evolving field known as epitranscriptomics.

Methods to detect RNA modifications

Accurately identifying and mapping RNA modifications has long been a challenge in epitranscriptomics. This is due to several technical difficulties, such as the low abundance of some modifications, their presence only in a subset of mRNAs, and their low chemical reactivity, among other factors. Over the years, several methods have been established to overcome these challenges, playing a crucial role in expanding the field. These methods can be broadly classified into multiple categories: either as direct and indirect or global and site-specific.

Direct and indirect RNA modifications methods

Direct methods involve the identification of modified nucleotides without the use of intermediate steps or markers. These techniques often involve high-resolution mass spectrometry or next-generation sequencing methods. Direct methods include techniques like

high-performance liquid chromatography (HPLC), 2D thin-layer chromatography (TLC) with radioactive isotopes and Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS). which can detect modifications directly on the RNA molecule.

On the other hand, indirect methods rely on the use of chemical treatment, antibodies or enzymes to recognise specific modifications. After the treatment, the modifications can detected by techniques such as immunoblotting or next-generation sequencing.

Global RNA modifications Identification

For global identification of modifications, techniques like HPLC, LC-MS/MS, and TLC are used, which give an overview of all modifications in a given RNA sample. In contrast, site-specific identification techniques, such as third-generation sequencing and reverse transcription at single nucleotide resolution with the use of antibodies or after chemical treatment of the sample, can provide information about the exact location of modifications on the RNA molecule.

Historically, the first method used to identify modified nucleotides was two-dimensional thin-layer chromatography, which utilised radioactive isotopes. This method facilitates the separation of nucleotides according to their distinct chemical properties. By comparing the migration patterns of these nucleotides with established standards, it becomes possible to identify modified nucleotides through their unique migration patterns.

High-Performance Liquid Chromatography (HPLC) is an analytical technique used to separate and purify components of a mixture based on their chemical properties. In HPLC, a mixture is passed through a column packed with a stationary phase. This solid or liquid material interacts with the sample components differently based on their chemical properties, such as polarity, size, and charge (Nees et al., 2014).

Mass spectrometry (MS) is a powerful analytical technique for quantifying molecules based on their mass-to-charge ratio (m/z). The technique involves ionising molecules, separating the resulting ions based on their m/z, and detecting the ions using a detector. Interestingly, we can also combine MS with HPLC to increase its performance (Meng, 2006).

Direct global methods, such as HPLC, HPLC-MS, LC-MS/MS, and 2D thin-layer chromatography, all share a standard limitation: they lack information about the specific position of modifications. This becomes particularly problematic when dealing with modifications that are of low abundance or are present only in specific types of RNA, such as mRNAs. In such cases, the sample being analysed must be meticulously purified to eliminate any contaminants.

Take the N_m modification as an example. N_m has been identified in various types of RNA, including rRNA, tRNA, mRNA, and small RNAs. Since N_m is found in high abundance in rRNA, detecting it in mRNAs—constituting only 1-5% of total RNA—requires a sample free of rRNA for precise analysis. Additionally, the use of proper controls is essential to determine the presence of any contaminants.

Given the challenges and limitations of direct global methods, developing site-specific identification techniques for modified nucleotides has emerged as a strategic approach to overcome these issues.

Site-specific RNA modifications detection methods

An even more critical aspect of epitranscriptomics is understanding the relationship between various modifications, their location and their function. This can be achieved only by applying methods which allow site-specific detection. Such methods depend on antibodies, gene engineering of modification reader proteins, reverse transcriptase properties, chemical modifications or direct sequencing (Motorin & Marchand, 2021).

One of the first mapped mRNA modifications was inosine, identified due to their unique chemical structure. Inosine is a deaminated adenosine, which is recognised as cytosine by reverse transcriptase. Consequently, the resulting cDNA contains guanosines in place of thymines. By sequencing this cDNA and comparing it with a reference genome or transcriptome, the positions of inosine can be inferred from the locations where the cDNA sequence contains a guanine in place of the expected adenine. This method offers valuable insight into the transcriptome-wide distribution and abundance of inosine-containing RNA molecules. It is important to note, however, that this technique does not directly detect inosine, but rather implies its presence based on the characteristics of reverse transcription (Bazak et al., 2014).

Antibody-based approaches are one way to identify the modified nucleotide position and sequence in which it is deposited. Such antibodies have a high affinity to modified nucleotides, and by pulling down and sequencing such fragments, we can identify their sequence. Antibodies were developed to m^6A/ m^6A_m, m^1A, hm^5C, ac^4C and m^7G (Weichmann et al., 2020; Y. Zhang et al., 2022). Moreover, antibodies can be used in the determination of relative quantities by the use of ELISA or dot blot.

Nanopore sequencing is a next-generation sequencing technology that uses a nanopore-based sensor to read the sequence of DNA or RNA molecules as they pass through a tiny pore.

This technology has been developed by Oxford Nanopore Technologies, and it offers several advantages over other sequencing methods.

One of the key advantages of nanopore sequencing is its ability to produce long reads, which can span thousands of bases (Nobuaki Kono & Kazuharu Arakawa, 2019; Y. Wang et al., 2021).

In nanopore sequencing, a single-stranded DNA or RNA molecule is passed through a nanopore, which is a membrane protein to which a motor protein is attached. As the molecule passes through the nanopore, the k-mer (5 consequent nucleotides) creates a change in electrical current that can be measured and recorded. This change in current is specific to each base in the DNA or RNA sequence, allowing the sequence to be determined (Jain et al., 2016; van Dijk et al., 2018; Y. Wang et al., 2021).

The change of charge and time of going through the pore can serve to detect the modified bases, which were first applied for DNA and $m^5dC/^5mC$ or $m^6dA/^6mA$ (Rand et al., 2017; Stoiber1 et al., 2017; Y. Wang et al., 2019). Recently, nanopore was used to identify sequence motifs of various mRNA modifications such as m^6A (Jenjaroenpun et al., 2021; Parker et al., 2020; Piechotta et al., 2022; Price et al., 2020), m^5C (Acera Mateos et al., 2023), N_m (Abebe et al., 2022; Begik et al., 2021; Hassan et al., 2022; Stephenson et al., 2022) or pseudouridine (Stephenson et al., 2022).

Recent studies show how endogenous modifications in ribosomal RNAs from *E. coli* and *S. cerevisiae* can be detected using nanopore sequencing. The modifications were identified through changes in the electrical signal and dwell times of the RNA passing through the nanopore, which were caused by interactions between the modified nucleotides and the helicase motor protein. By analysing these changes, the specific types and locations of the modifications in the RNA sequences were identified (Begik et al., 2021; Hassan et al., 2022; Stephenson et al., 2022).

Methods to map *2′-O*-methylation

As mentioned earlier, *2′-O*-methylation (N_m) modification is a common modification of rRNA, tRNAs, mRNAs, and small RNAs. Interestingly, N_m modification was shown to be impactful in various cellular mechanisms such as translation. Thus, a handful of methods were developed for its mapping. Interestingly, N_m modification can be found in mRNAs at internal sites and at the 5′ end, where it only occurs at the first one or two mRNA nucleotides.

Reverse transcriptase, when transcribing at low deoxy-ribonucleoside triphosphate (dNTP), stops the transcription at the N_m nucleotide and thus, a method called Reverse Transcription at Low deoxy-ribonucleoside triphosphate (dNTP) concentrations followed by polymerase chain reaction (PCR) (RTL-P) was developed (Z.-W. Dong et al., 2012).

Another characteristic of the N_m modification is its ability to protect the phosphodiester bond between the modified nucleotide and the downstream nucleotide during alkaline cleavage. When combined with deep sequencing, this property allows for identifying sequences less susceptible to random fragmentation. A method known as RiboMethSeq has been developed based on this property (Birkedal et al., 2015; Motorin et al., 2021; Pichot et al., 2020). However, one limitation of this method is its requirement for random cleavage and high-depth sequencing, making it more suitable for longer RNAs. Conversely, RiboMethSeq has also been demonstrated to apply to mRNAs (Ringeard et al., 2019).

Other methods are N_mSeq (Dai et al., 2017) and RibOxi-seq (Zhu et al., 2017). Both methods rely on the differential reactivity of 2′-OMe versus 2′-OH nucleotides. Fragmented RNA undergoes oxidation which renders the non-methylated ends incapable of ligation to linkers used for high-throughput library construction. The resulting reads are aligned to a reference genome and analysed to determine methylation sites. However, it is worth noting that some results from the N_mSeq method, such as the identified motif, have been questioned due to the potential issue of adaptor sequence detection, possibly affecting the validity of the results.

All the methods described above are suitable only for internal N_m sites. To map the N_m sites and prevalence of single vs double N_m sites at the 5′ end of mRNA, two methods were developed: CapTag-seq and CLAM-Cap-seq (Despic & Jaffrey, 2023).

CapTag-seq is a quantitative method used to measure the levels of cap1 and cap2. To perform CapTag-seq, polyA purified RNAs undergo enzymatic decapping, followed by ligation of a 5′ adapter composed of *2′-O*-methylated nucleotides. The RNA is then subjected to RNase T2, which cleaves all phosphodiester bonds except those with N_m modification. After the RNase T2 treatment, the 5′ adaptor is left with either 1, 2, or 3 nucleotides. A single nucleotide corresponds to RNA with an unmodified 5′ end, while three nucleotides correspond to RNA with 2 N_m modifications at the 5′ end CapTag-seq provides a highly accurate method for quantitatively measuring the levels of cap1 and cap2 (Despic & Jaffrey, 2023).

CLAM-Cap-seq is a method used to identify transcripts with cap0, cap1, and cap2 modifications on their 5′ ends. The mRNA is first decapped, and then reverse transcription is

applied. The resulting complementary strand is circularised using a circular ligase, which enables it to connect with the mRNA molecule. The RNA-DNA hybrid is then subjected to RNA degradation using RNase T2 and KOH, leaving only RNA with N_m modifications in the DNA-RNA hybrid. After adaptor ligation, the library is subjected to Illumina sequencing, enabling transcriptome-wide mapping of 5′ end N_m modifications. Nevertheless, the biggest drawback of this method is requirement of 7 µg poly(A)+ RNA as starting material which might not be feasible for many tissues (Despic & Jaffrey, 2023).

Internal mRNA modifications

The intricate landscape of gene expression is significantly influenced by an array of internal mRNA modifications. Among the most well-studied internal mRNA modifications are *N6*-methyladenosine (m^6A), inosine (I), *2′-O*-methylation (N_m) and pseudouridine (Ψ). Each of these modifications has a distinct role in mRNA processing and regulation, shaping the intricate landscape of gene expression and thus contributing to the complexity of gene expression. Meanwhile, other less abundant modifications such as *N5*-methylcytosine (m^5C), *N1*-methyladenosine (m^1A), 5-hydroxymethylcytosine (hm^5C), *N7*-methyladenosine and *N4*-acetylcytidine (ac^4C) were also identified in mRNAs. However, these are less abundant, and as a result, their roles are not as well-established. The significance and potential functions of these less prevalent modifications remain topics of ongoing discussion and research in the field (S. Kumar & Mohapatra, 2021; Sun et al., 2023).

This section will focus to briefly elucidate the multifaceted nature of some internal mRNA modifications.

N6-methyladenosine (m^6A)

N6-methyladenosine (m^6A), the first discovered modification of mRNA, has been identified in a wide range of biological systems. These include rats (Desrosiers et al., 1974), human cells (Wei et al., 1976), mouse cells (Schibler et al., 1977), and even viruses (Krug et al., 1976). The high prevalence of m^6A in mRNA likely enabled its early discovery in mRNAs.

In 2012, two groups developed a method to map m^6A sites using m^6A-specific antibodies and subsequent next-generation sequencing (Dominissini et al., 2012; Meyer et al., 2012). Despite initial evidence of m^6A appearing in the 1970s, it was the novel technologies, such as mapping of m^6A, that allowed the significant expansion of the m^6A field. This might be observed by a brief search on PubMed for m^6A and RNA. Publications have grown from

approximately 94 prior to 2012 to 4,996 since then, averaging 27 new publications per week in the current year. With new approaches for mapping m^6A across the whole transcriptome calculated approximately one m^6A site per 2000 nucleotides in RNA HepG2 cells (Dominissini et al., 2012).

In humans, m^6A is deposited on mRNAs co-transcriptionally by two methyl transferases (METTL), monomeric METTL16 and heterodimeric complex METTL3/14. METTL16 recognises a loop of specific TACAGAGAA sequence (Mendel et al., 2018) and the heterodimeric complex METTL3/14 recognises a single-stranded sequence motif RRACH (IUPAC sequence standing for puRin, puRin, Adenosine, Cytosine, H - any except Guanosine). Even though the RRACH motif is uniformly distributed among themRNA transcripts, m^6A is enriched around the STOP codon (Dominissini et al., 2012; Meyer et al., 2012). Recently three publications showed that the exon-junction complex shapes the m^6A distribution by abrogating the METTL3/14 complex's binding to mRNA (P. C. He et al., 2023; Uzonyi et al., 2023; X. Yang et al., 2022).

Studies of those two separate complexes – METTL16 and METT3/14 – showed that the same modification can lead to different consequences based on it position. When deposited at 3′ splice site it can inhibit splicing (Mendel et al., 2021), when deposited at different loci it can affect export (indirectly) (Lesbirel et al., 2018; Zheng et al., 2013), localisation (Flamand & Meyer 2022), translation (Coots et al., 2017; Mao et al., 2019), polyadenylation (Wu et al., 2023) and mRNA stability (Lasman et al., 2020).

Furthermore, m^6A affects not only molecular functions but also physiological functions. It plays crucial role in regulation of S-adenosine methionine- the methyl group donor (Mendel et al., 2018; Pendleton et al., 2017), embryogenesis in mammals (Batista et al., 2014; Geula et al., 2015) and plants (Zhong et al., 2008), meiosis in yeast (Clancy et al., 2002; Schwartz et al., 2013), sex determination in fruit flies (Haussmann et al., 2016) and X-chromosome inactivation *via* XIST (Patil et al., 2016) or circadian clock control (J. M. Fustin et al., 2013; J.-M. Fustin et al., 2018).

In summary, m^6A modification in mRNA has emerged as a crucial regulatory mechanism, influencing all aspects of the mRNA life cycle and physiological functions across diverse biological systems.

Inosine (I)

Inosine is a modified nucleoside that is created through the deamination of adenosine in RNA. Deamination of adenosines is performed by enzymes of the ADAR (Adenosine Deaminase Acting on RNA) family of genes (Melcher et al., 1996). In humans, three ADAR genes have been identified: ADAR1, ADAR2, and ADAR3. All of them possess the double-stranded RNA (dsRNA) binding domain, which explains the position of inosines in dsRNA regions.

ADAR1 has three isoforms that differ in their N-terminal sequences and subcellular localisation (Galipon et al., 2017; Pestal et al., 2015). ADAR1 is involved in the editing of both coding and non-coding RNA and plays important roles in innate immunity (Liddicoat et al., 2015a; Mannion et al., 2014; Niescierowicz et al., 2022) and RNA interference (W. Yang et al., 2005). ADAR1 marks cellular dsRNA as self-molecules to avoid triggering autoimmunity (Rice et al., 2012). The involvement of ADAR1 and inosine in innate immunity is discussed in the next chapter.

ADAR2 is primarily expressed in the brain. ADAR2 mutant mice have been shown to experience seizures and die within the first three weeks after birth. Unlike ADAR1, it has been found that their early death is not due to the activation of the innate immune system. Rather, the mutant mice exhibit a phenotype that results from a single under-edited position in the transcript of one subunit, one Glutamate receptor (GluA2). Unedited mRNA is translated with a single mutation leading to a dysfunctional protein (Brusa et al., 1995). The phenotype of the ADAR2 mutant is rescued by the knock-in of a single point mutation in GluA2 (Higuchi et al., 2000; Seeburg et al., 1998).

ADAR3 is an inactive Adenosine deaminase but has retained its ability to bind dsRNA. ADAR3 mutant mice have increased anxiety levels and deficits in hippocampus-dependent short- and long-term memory formation (Miles et al., 2018).

The presence of inosine in dsRNA molecules is a feature conserved across metazoans. The significance of inosine has been demonstrated in the fruit fly, *Drosophila melanogaster*. Genetic knockout of the only ADAR gene in flies leads to several abnormalities, including defects in motor control, mating, and flight, as seen in ADAR null mutants - flies with no functional ADAR gene (Palladino et al., 2000). These null mutants also exhibit increased sleep due to synaptic dysfunction in glutamatergic neurons (J. E. Robinson et al., 2015). Moreover, both catalytically dead and null ADAR knockouts were found to express immune-induced molecules, highlighting ADAR's essential role in maintaining a normal immune response (Deng et al., 2020a).

N5-methylcytosine (m^5C)

m5C, one of the first RNA modifications identified, initially stirred debate about its existence in mRNA. This was mainly because it was commonly found in tRNA but seemed rare in mRNA. However, thanks to recent advancements in mapping technologies, the presence of m^5C in mRNAs was confirmed. (Huang et al., 2019; Selmi et al., 2021). NSUN2 and NSUN6 are known m^5C methyltransferases that act on distinct sets of mRNAs and motifs (Huang et al., 2019; Schumann et al., 2020; Selmi et al., 2021; Trixl & Lusser, 2019). In general, m^5C is found preferentially near the start codon (J. Liu et al., 2022; Schumann et al., 2020; X. Yang et al., 2017), and m^5C-containing transcripts are less efficiently translated (Schumann et al., 2020). However, NSUN6 sites are located near the stop codon, and their presence correlates with higher translation rates (Selmi et al., 2021). Recent studies have shown that m^5C is enriched on maternal mRNAs in various species, including fruit flies, frogs, zebrafish, mice, and humans, and preferentially near the 3′ UTR (J. Liu et al., 2022). Another proposed function of m^5C is its involvement in mRNA export (X. Yang et al., 2017).

Pseudouridine (Ψ)

Pseudouridine (Ψ) is a type of RNA modification where the uridine (U) base is isomerised to form a different base structure. Ψ was one of the first identified RNA modifications (Cohn & Volkin, 1951; Davis & Allen, 1957). Ψ is installed by the family of enzymes called pseudouridine synthases (PUS) and found in almost all types of RNA, including rRNA, tRNA, mRNA, and non-coding RNA (Martinez et al., 2022). Ψ enhances stabilises the RNA structure of tRNA and rRNA (Arnez & Steitz, 1994; Davis & Poulter, 1991; Newby & Greenbaum, 2002)by stabilisation of structure in duplexes between Ψ -A, Ψ -G, Ψ -U and Ψ -C pairs (Kierzek et al., 2014)

Ψ in mRNAs is present in pseudouridine/U ratio in about 0.2–0.6% (X. Li et al., 2015). It′s deposited co-transcriptionally as well as post-transcriptionally. Presence of Ψ in the polypyrimidine tract in introns leading to the splicing failure due to the disability of U2AF(65) to recognise the pseudouridylated polypyrimidine tract (C. Chen et al., 2010). In contrast, intronic pseudouridine upstream of the 3′ ss enhances splicing (Martinez et al., 2022). Moreover, the depletion of pseudouridine synthetases, PUS1, PUS7 and RPUSD4, leads to widespread effects on alternative splicing (Martinez et al., 2022).

2′-O-methylation (N_m)

Internal *2′-O*-methylation (N_m) refers to methylation occurring on the ribose sugar of any nucleotide, regardless of the nucleotide's base forming A_m, U_m, G_m or C_m. This modification is common in rRNA and tRNA molecules. In humans, the process of N_m of RNA is primarily catalysed by a class of enzymes known as methyltransferases. Specifically, the 2′O-methylation of rRNA is performed by a complex known as small nucleolar ribonucleoprotein particles (snoRNPs) consisting of two key components, snoRNA and Fibrillarinn, where snoRNA guides the site-specificic methylation. At the same time, a different set of enzymes conducts the N_m of transfer RNAs (tRNAs) in humans. A family of tRNA methyltransferase (TRMT) is mainly responsible for these modifications. Within this family, several enzymes, including TRMT61A and TRMT6/61B, help add a methyl group at different positions in the tRNA molecule.

Interestingly, N_m also appears as a modification within mRNA molecules. Two enzymes, FTSJ3, belonging to TRMT and Fibrillarin methylating primarily rRNAs, have been shown to deposit internal N_m sites on mRNA. A method called Reverse Transcription at Low dNTP concentrations followed by PCR (RTL-P) (Z.-W. Dong et al., 2012) has revealed that Fibrillarin, along with snoRNAs, modifies the mRNA of a gene called Pxdn with N_m. This modification increases the stability of Pxdn mRNA but reduces Pxdn protein levels, which suggests that N_m may hinder translation. (Elliott et al., 2019). This probably due to the N_m in mRNA codons that can significantly impair protein translation by causing a high rejection rate of the corresponding tRNA (J. Choi et al., 2018).

The role of internal N_m modification in mRNA has been particularly investigated in the context of viral infections. The N_m modification in the mRNA cap structure serve as a mark of cellular RNAs whereas the internal N_m modification can stabilise the RNA. Its role in innate immunity will be further discussed in the next chapter.

N1-methyladenosine (m^1A)

Studies have revealed that the deposition of m^1A is primarily limited to tRNA-like structures within a select number of gene coding transcripts, mediated by the TRMT6/TRMT61A complex (Dominissini et al., 2016; X. Li et al., 2016; Safra et al., 2017). It is believed that this

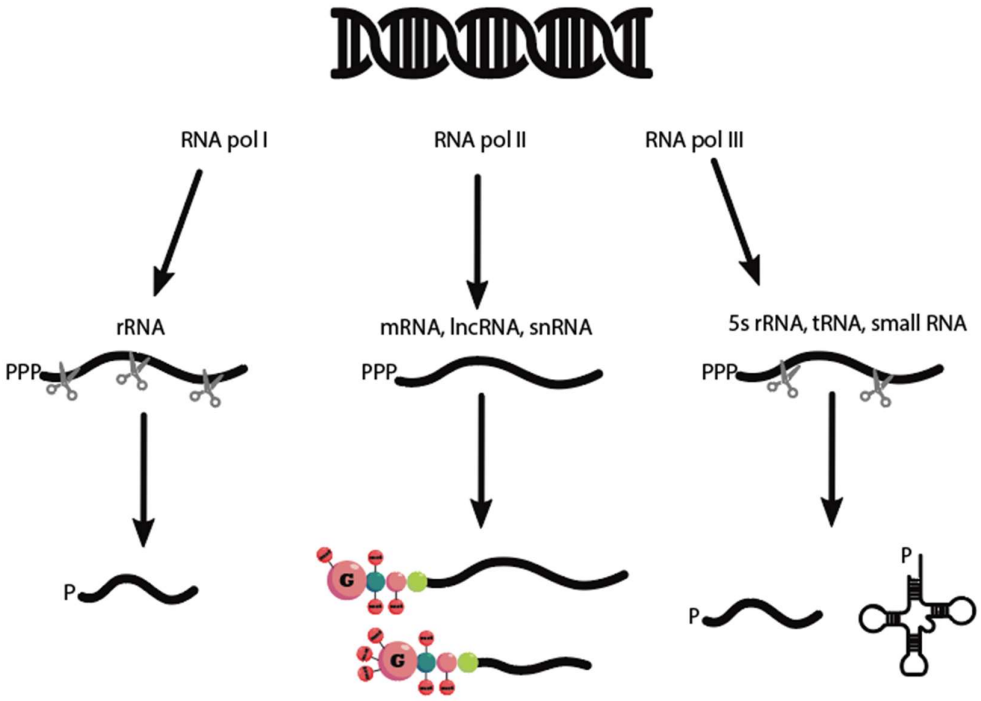


Figure 3. Processing of 5′end termini in different cell types. P – terminal phosphate, rRNA – ribosomal RNA, lncRNA – long non-coding RNA, snRNA – small nuclear RNA, tRNA – transfer RNA

modification disrupts the A-T base pairing, thereby altering the secondary structure of the RNA (L. Lu et al., 2010). However, the demethylation of m1A by ALKBH3 can reverse this modification (X. Li et al., 2016). Interestingly, mRNAs that contain m^1A are inefficiently translated (Safra et al., 2017).

Nevertheless, the reanalysis of data proposed that many of the formerly identified m^1A sites were misannotated and proposed that cytosolic mRNAs m^1A is a rare internal modification at very low stoichiometries and at very low number of sites (Schwartz, 2018; Wiener & Schwartz, 2021)

mRNA cap modifications

RNA polymerases use the nucleotide triphosphates as the building blocks to synthesise RNA molecules. This results in 5′ end triphosphates on all RNA molecules. Nevertheless, such 5′ end termini are unfavoured and are further processed. Most RNA classes possess the 5′

monophosphate, which is secured by endonucleolytic cleavage or by removal of diphosphate by the DUSP11 (Burke & Sullivan, 2017). In contrast, the transcripts of RNA polymerase II undergo a process called mRNA capping. Moreover, all RNA classes need to be protected from degradation either by making secondary structures or by 5′ end binding proteins (**Chyba! Nenalezen zdroj odkazů.**).

Initially, mRNA transcribed by RNA polymerase II has a triphosphate at its 5′ end. When the transcript reaches about 25 nt in length, the triphosphate is converted to the 5′-5′ triphosphate bridge-linked *N7*-methylguanosine (m^7G) cap in the process of capping (**Chyba! Nenalezen zdroj odkazů.**) (Coppola et al., 1983; Shatkin & Manley, 2000).

In higher eukaryotes, RNA capping is carried by two enzymes performing three activities: RNGTT (triphosphatase and guanylyltransferase) removes the terminal phosphate and adds GMP forming Gppp-RNA. At the same time, (RNA guanine-7-methyltransferase) monomethylates the added guanine base on the seventh position, forming m^7Gppp-RNA. This mRNA cap structure (m^7GpppNN) is termed the cap0.

Moreover, in higher organisms, mRNA cap structure can be further modified by *2′-O*-methylation on the first transcribed nucleotide (cap1) or the first two transcribed nucleotides (cap2). The transcription start site (TSS) nucleotide is always modified by the CMTR1 enzyme forming cap1 (m^7Gppp N_mpNp) structure (Bélanger et al., 2010). Another mammalian ribose methylatransferase, CMTR2, methylates the second transcribed nucleotide creating cap2 (m^7Gppp N_mpN_mp) structure which is present in 50% of transcripts of polyadenylated RNAs in human HeLa cells (Despic & Jaffrey, 2023; Furuichi et al., 1975; Wei Cha-Mer and Gershowitz, 1975). Additionally, if the first transcribed nucleotide is adenosine, it can be further modified at the N^6 position of the adenine base to generate *N6,2′-O*-dimethyl adenosine (m^6A_m) (Akichika et al., 2018; Boulias et al., 2019; Sendinc et al., 2019; Sun et al., 2019; Wei Cha-Merand Gershowitz, 1975). Interestingly, mRNA capping happens co-transcriptionally (Akichika et al., 2018; Galloway et al., 2021; Inesta-Vaquera et al., 2018; Pandey et al., 2020) except *2′-O*-methylation at the second transcribed nucleotide which happens only in the cytoplasm where CMTR2 localises (**Figure *4***) (Despic & Jaffrey, 2023; Werner et al., 2011).

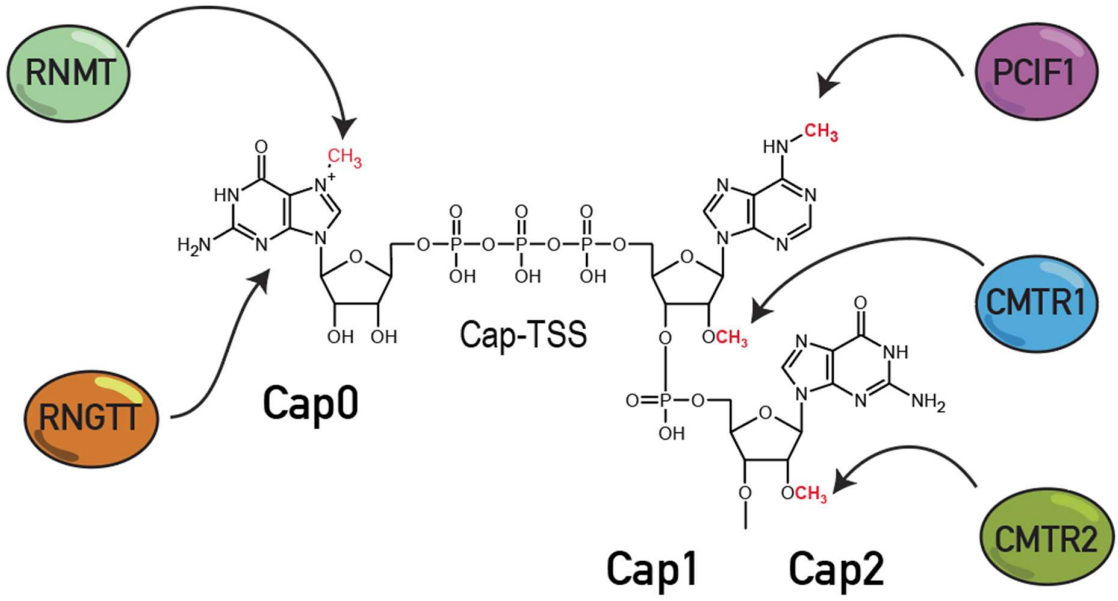


Figure 4. Mammalian mRNA cap structure. Mammalian cap structure depends on the activities of 5 enzymes.

Cap0 function

Cap0 was implicated in multiple different functions in mRNA, from protecting from RNA degradation by blocking XRN1 exonuclease through mRNA export and stimulation mRNA splicing to being essential for translation initiation (Gonatopoulos-Pournatzis & Cowling, 2014).

Cap0 protect the mRNA from degradation. Triphosphorylated mRNAs are recognised by the NUDT2, which catalyses only the first step of mRNA degradation. The monophosphorylated mRNAs are then recognised and cleaved by 5′- 3′- end exoribonucleases – XRN1 and XRN2 (Laudenbach et al., 2021).

Cap0 is recognised by the cap-binding complex (CBC), which coordinates various processes involved in mRNA processing, including pre-mRNA splicing, 3′-end processing, nonsense-mediated decay, nuclear–cytoplasmic transport, recruitment of translation factors in the cytoplasm as well as the nuclear export of snRNAs (Elisa Izaurralde et al., 1995). The CBC is composed of a heterodimer consisting of nuclear cap-binding protein 2 (NCBP2 or Cap-binding protein 20 - CBP20) and NCBP1 (also known as CBP80) (Elisa Izaurralde et al., 1995). NCBP2 directly binds to the mRNA cap, while NCBP1 acts as a stabilising factor and an adaptor for other processing factors (Gebhardt et al., 2015; Gonatopoulos-Pournatzis & Cowling, 2014; Topisirovic et al., 2011).

In *in vitro* splicing studies was shown that the mRNA cap is unnecessary for splicing, but it increases the splicing efficiency in HeLa nuclear extracts. Nevertheless, the addition of a cap analogue to the reaction inhibits the splicing which pointed out to its essential role in splicing (Edery & Sonenberg, 1985; Konarska et al., 1984; Patzelt et al., 1987). CBC facilitates the association of U1 snRNP with the cap proximal 5′splice site during the formation of the E (early) complex showing its importance in splicing of the first intron. However, in a pre-mRNA containing two introns, CBC is not required for splicing of the cap distal intron. (Lewis et al., 1996). CBC remains associated with the pre-mRNA throughout the whole splicing cycle (Lewis et al., 1996).

Notably, only appropriately processed mRNAs - those that are capped, spliced, and possess a polyA tail - are exported to the cytoplasm (Carmody & Wente, 2009). This understanding was reinforced by experimental data showing that uncapped RNAs when microinjected into a Xenopus nucleus, failed to make this transition to the cytoplasm (Cheng et al., 2006). An exciting aspect of this process is that some export factors, such as ALYREF, directly interact with the EJC, ensuring that only spliced mRNAs are transported from the nucleus to the cytoplasm (Gromadzka et al., 2016). Furthermore, the cap on mRNA is recognised by CBP80, which in turn interacts with ALYREF. This interaction promotes further association with the TREX complex, setting the stage for nuclear export (Cheng et al., 2006). Moreover, inhibition of polyadenylation leads to mRNA retention in the nucleus (Apponi et al., 2010; Z. Chen et al., 1999; Tudek et al., 2018). Interestingly, the Cap Binding Complex (CBC) remains consistently associated with the mRNA cap throughout the entirety of the nuclear export process (Visa et al., 1977).

When mRNA is exported from the nucleus to the cytoplasm, the mRNA cap guides another process. Cap-dependent mRNA translation initiation. The first few rounds of translation are called the pioneer round of translation, which is CBP dependent. During this phase, CBC directly interacts with eIF4G (McKendrick et al., 2001). Later on, the translation initiation is dependent on eIF4e. The transition from the pioneer round of translation to the standard mode of translation depends on the exchange of CBC for eIF4E at m^7G, which is regulated by importins (Sato & Maquat, 2009). The importance of eIF4 and cap7 on translation was shown in various experimental setups. *In vitro*, Brome mosaic virus RNAs lacking m^7G cap leads to the reduction but not complete abolition of translation (Shih et al., 1976). *In vivo* in tobacco protoplasts, Chinese hamsters, ovary cells, and yeast following delivery by electroporation, only capped mRNAs are translated (Gallie, 1991).

Interestingly, there was only limited evidence involving cap-independent translation initiation. First are viral-specific loops in the 5′ UTRs called Internal ribosomal entry. Second, were m^6A sites in 5′UTR (Hoshi et al., 1984; Y. Yang & Wang, 2019). Nevertheless, for the internal ribosomal entry site, there are still no known endogenous ones, and second, for m^6A-dependent translation initiation, only limited evidence was shown (Meyer et al., 2015).

Cap1 function

The first transcribed nucleotide of mRNA is always modified with a *2'O*-methylation (Furuichi et al., 1975; Wei Cha-Mer and Gershowitz, 1975) by the enzyme CMTR1 forming $m^7Gppp\ N_m$ -RNA (cap1) (Bélanger et al., 2010). Cap1 mRNA structure was shown to be important for mRNA stability, translation and recognition of self- and non-self- RNA molecules. Probably none if these is caused by the CBC complex, as cap1 does not affect the affinity of the cap-binding protein (CBP) to m^7G cap analogues (Worch et al., 2005).

Cap1 protects the mRNA against recognition by the nuclear decapping enzyme DXO which recognises mRNA without N_m modification at the first position ($m^7GpppRNAs$), hydrolyses it and promotes its degradation (Picard-Jean et al., 2018). Its nuclear localisation prevents the maturation of unmethylated endogenous RNAs and their export to the cytoplasm.

Interestingly, the cap1 mRNA structure was never tested for splicing. Nevertheless, cap1, as well as cap2 on U2 snRNA, were tested. Both cap1 and cap2 were shown to be essential for splicing in U2 snRNA. Chimeric U2 snRNA lacking N_m modification at the first and second transcribed nucleotide showed that N_m modifications are crucial for efficient splicing and E-complex formation (Dönmez et al., 2004).

In *in vitro* conditions, both cap0 and cap1 are translated in wheat germ and reticulocyte lysate systems. However, high input concentrations were m^7GpppA_m enriched up to 2-fold in the ribosome-bound fraction (Muthukrishnan et al., 1978). In *Xenopus* oocytes, the maturation of oocytes by progesterone promotes cap1 and cap2 methylation and polyadenylation. Inhibition of methylation doesn't affect mRNA polyadenylation but does impact mRNA translation. In the oocytes, cap0 mRNAs are not translated (Kuge et al., 1998). In differentiated cells, cap0 mRNA molecules are translationally inhibited by the cap-binding protein IFIT1 during viral infection. IFIT1 specifically binds the cap0 structure, preventing the initiation of its translation. Furthermore, the presence of cap1 is critical to avoid the activation of the innate immune system, a topic that will be explored in more detail in the next chapter.

Cap2 function

The structure of Cap2 has been observed to vary across different organisms, tissues and human cell lines. The presence of cap2 in humans is estimated to be around 50% (Despic & Jaffrey, 2023; Furuichi et al., 1975), with a bit slightly higher occupancy in A starting transcripts (Despic & Jaffrey, 2023). More than sequence specificity, cap2 is present on longer-living mRNAs. Nevertheless, cap2 does not affect half-life of mRNAs neither their (Despic & Jaffrey, 2023). Interestingly, cap2 was also shown to play a role in marking cellular mRNAs as self (Despic & Jaffrey, 2023; Schuberth-Wagner et al., 2015; Wang Yanli et al., 2010).

m^6A_m cap function

mRNAs always possess a cap1 structure. When the TSS nucleotide is adenosine, it can be further modified at the *N6* position to form the m^7Gppp m^6A_m-RNA by PCIF1 (Akichika et al., 2018; Boulias et al., 2019; Sendinc et al., 2019; Sun et al., 2019). m^6A_m is present in 92% of A starting transcripts (Akichika et al., 2018). Investigation of m^6A_m transcripts in *PCIF1* KO background does not support either mRNA stability or mRNA translation effect or rather points out that it might have different outcomes in different cell types (Akichika et al., 2018; Boulias et al., 2019; Pandey et al., 2020; Sendinc et al., 2019).

Cap modifying enzymes

mRNA cap structure depends on the enzymatic activity of 5 enzymes, RNGTT, RNMT, CMTR1, CMTR2 and PCIF1.

RNMT

RNMT (RNA guanine-7-methyltransferase) monomethylates the Gppp-RNA to m^7Gppp-RNA (cap0). Although RNMT has an active methyltransferase domain, RNMT needs to be activated by RAM (Liang et al., 2023). RAM increases the recruitment of the methyl donor, AdoMet (S-adenosyl methionine, SAM), to RNMT (Varshney et al., 2016).

Interestingly, naive CD4 T cells do not express RNMT either RAM; only their activation drives their expression. To study the effect of RNMT on CD4 cells differentiation, a conditional knockout mouse (cKO) model mouse was established, which specific deletion of RNMT in CD4 T cells *via* Tg(Cd4-cre)1Cwi. The cKO T cells were confirmed to possess a Gppp-RNA cap structure. Induction of T cell activation leads to defects in the cell cycle and apoptosis. The characterisation of naïve T cells by sequencing showed over 1500 dysregulated

genes. The analysis of the transcriptome of cKO CD4 cells showed that RNMT (cap0) is essential for the expression of some snoRNAs and terminal polypyrimidine tract (TOP) mRNAs (Galloway et al., 2021). Moreover, the lack of RNMT leads to decreased ribosome synthesis, reduced translation rates and proliferation failure (Galloway et al., 2021). The fact that RNMT is not expressed in naïve T cells and the cKO T cells fail to activate indicates that the presence of cap0 is more crucial in activated cells.

TOP transcripts

TOP transcripts, also known as TOP mRNAs, refer to a specific group of messenger RNAs that contain a characteristic motif known as the "TOP motif" in their 5′ UTR. The term "TOP" stands for "terminal oligopyrimidine tract." These transcripts play a crucial role in regulating protein synthesis in response to various cellular conditions, including stress (Avni et al., 1997).

TOP mRNAs are highly sensitive to stress signals and exhibit unique regulatory properties. They are particularly responsive to stress conditions affecting cellular energy, such as nutrient deprivation, hypoxia, and amino acid limitation. Under normal conditions, TOP mRNAs are efficiently translated, contributing to synthesising ribosomal proteins and other components of the translational machinery (Cockman et al., 2020).

TOP mRNAs are targets of cap0 binding protein LARP1 which can be regulated by mTORC1 (Philippe et al., 2020). In the context of cellular stress, LARP1 plays a unique role by anchoring 5′ TOP transcripts within stress granules (SGs) and processing bodies (PBs) (Wilbertz et al., 2019). While this relocation to granules is a stress-induced phenomenon, it does not appear to affect the translation or decay of these transcripts during recovery (Wilbertz et al., 2019). This indicates a nuanced, context-dependent regulation of 5′TOP mRNAs.

CMTR1

CMTR1 is a multidomain protein with nuclear localisation signal (NLS), G-patch domain, Rossman-fold methyltransferase (RFM), GTase-like domain and WW domain (Werner et al., 2011). CMTR1 localises to the nucleus (Werner et al., 2011) and interacts with RNA pol II *via* its WW domain (Inesta-Vaquera et al., 2018). WW domain is well-known for interacting with RNA pol II C-terminal domain (CTD). CMTR1 interacts with RNA pol II only when CTD is phosphorylated at Ser5 (Inesta-Vaquera et al., 2018). This interaction enables the methylation of pre-mRNAs during the early elongation stages of transcription when the mRNA is m^7G capped (Heidemann et al., 2013).

Another protein interacting factor of CMTR1 is DHX15 (Inesta-Vaquera et al., 2018; Toczydlowska-Socha et al., 2018). DHX15 was shown as splicing factor under the gene name Prp43 in yeasts. CMTR1 can methylate only ssRNA, and as DHX15 is a helicase, one proposed function of this interaction is enabling CMTR1 to methylate structured mRNAs by DHX15 unwinding the dsRNA regions. Indeed such a case was confirmed *in vitro,* where CMTR1 could methylate dsRNA only when incubated with DHX15 (Toczydlowska-Socha et al., 2018). Nevertheless, another study showed that methylation activity on ssRNA of CMTR1 decreases when in complex with DHX15 (Inesta-Vaquera et al., 2018). Moreover, the interaction of CMTR1 with DHX15 and RNA pol II is mutually exclusive with distinct localisations indicating that DHX15-CMTR1 complex might have distinct, yet unknown fucntion (Inesta-Vaquera et al., 2018).

The physiological function of CMTR1

Differentiation of Mouse embryonic stem cells (mESC) upon LIF (Leukemia Inhibitory Factor) withdrawal showed that levels of CMTR1 increase during the differentiation. The lack of CMTR1 by siRNA knockdown (KD) leads to the failure of mESC cells to differentiate upon LIF withdrawal as well as during neural differentiation (Y. L. Lee et al., 2020; Liang et al., 2022). Knockdown of *Cmtr1* leads to downregulation of histone and ribosomal genes (Liang et al., 2022). Reduction of histones leads to DNA damage (Hogan & Foltz, 2021), also confirmed in the *Cmtr1* KD cells (Liang et al., 2022). Moreover, *Cmtr1* KO mice are embryonically lethal (Y. L. Lee et al., 2020).

siRNA-mediated *Cmtr1* KD in DIV2 (day *in vitro* neurons) rat neurons impairs dendritic development. Moreover, *Cmtr1*-cKO Emx1 (cKO in EMX1-expressing neuronal progenitors) mice show reduced cortical size and abnormal dendritic morphology (Y. L. Lee et al., 2020).

In addition, CMTR1 was also implicated to have a role in cancer progression. Upregulation of CMTR1 is correlated with poor prognosis in colorectal cancer, while CMTR1 downregulation by siRNA-mediated KD led to suppressed cell proliferation and tumorigenicity (You et al., 2023).

CMTR2

The second transcribe nucleotide is modified in approximately 50% (Despic & Jaffrey, 2023; Furuichi et al., 1975) cases by the action of CMTR2 enzyme forming $m^7GpppN_mN_m$-RNA (cap2) (Werner et al., 2011). CMTR2 is less studied compared to CMTR1.

CMTR2 localises primarily to the cytoplasm (Werner et al., 2011). CMTR2 consists of two Rossman-fold MTase domains, with only the first domain possessing catalytic activity (Smietanski et al., 2014). *In vitro* methylation studies have shown that CMTR2 can methylate both m^7G and $m^{2,2,7}G$ caps in mRNA and snRNA (Werner et al., 2011).

The physiological function of CMTR2

A genetic fly model of CMTr2 exhibited no defects. Immunofluorescence analysis revealed that CMTr2 localises to the polytene chromosomes in flies (Haussmann et al., 2022), while in mammals, it predominantly localises to the cytoplasm (Werner et al., 2011). Moreover, dCMTR2 in flies shows an affinity for methylating the first transcribed nucleotide to form the cap1 structure. This indicates the possibility of redundant functions between Drosophila melanogaster CMTr1 and CMTr2. Fly double mutants lacking both CMTr1 and CMTr2 are viable but exhibit defects in reward learning (Haussmann et al., 2022).

As well as cap1, it has been shown that cap2 plays a role in recognising self and non-self-RNA molecules (Despic & Jaffrey, 2023; Schuberth-Wagner et al., 2015).

PCIF1

Phosphorylated C-terminal domain-interacting Factor 1 (PCIF1) methylates A starting transcripts to generate m^6A_m. This enzyme is primarily localised in the nucleus and is known to interact with RNA polymerase II (RNA pol II) during the early stages of transcription elongation, facilitated by its WW domain (Akichika et al., 2018). Interestingly, the absence of PCIF1 does not affect cell viability or growth in human cell lines, highlighting its non-essential role in basic cell function (Akichika et al., 2018; Boulias et al., 2019; Sendinc et al., 2019). However, cells lacking PCIF1 exhibit increased sensitivity to oxidative stress, implying a potential role of PCIF1 in the cellular response to oxidative damage (Akichika et al., 2018). Studies on animal models revealed that mice that PCIF1 is not crucial for viability or fertility, but, he PCIF1 KO animals display decreased body weight, hinting to at an important role of PCIF1 that is not yet understood (Pandey et al., 2020).

Innate immunity

In the process of evolution, various changes occur in populations of organisms over time, leading to adaptations that increase their fitness in response to environmental pressures. These genetic changes are passed down from generation to generation, allowing for the emergence of new traits and, in some cases, the formation of new species. These changes can occur due to various mechanisms, including mutation, gene flow, genetic drift, and natural selection. One of the less obvious but key factor in the evolution of organisms is the pathogen-host interaction, where there is an ongoing arms race in which pathogens evolve new ways to infect hosts, and hosts evolve new defences to resist infection. This arms race was a key component in the evolution of the immune system.

The immune system can be divided into two main branches: innate immunity and adaptive immunity. Innate immunity is evolutionarily older, fast, and does not need the individual to meet the pathogen in advance. One possesses it from birth, whereas adaptive immunity is acquired over time.

Innate immunity uses a variety of specialised cells and molecules to sense the presence of pathogens and initiate a rapid response. These cells and molecules are collectively known as pattern recognition receptors (PRRs), and they can recognise conserved structures commonly found on many different types of pathogens. The conserved pathogens' structures are called pathogen-associated molecular patterns (PAMPs) and could be any molecular structure uncommon for the host. When PRRs recognise PAMPs, they trigger a signalling pathway that leads to the activation of the innate immune system leading to the production of molecules such as cytokines and chemokines, which attract other immune cells to the site of infection and activate their effector functions to clear the pathogen. Examples of PAMPs could be components of bacterial cell walls like lipopolysaccharide (LPS) and peptidoglycan, fungal cell wall components like beta-glucans and viral RNA and DNA.

Nucleic acids in recognition of non-self

Innate immunity to sense the non-self PAMPs is present from bacteria to humans. Even though the innate immune systems are different among different species, they still have some similarities. Notably, over 60 different innate immune systems were found in bacteria (Duncan-Lowey et al., 2023; Tesson et al., 2022). Nevertheless, only a minimum of those are similar to the human innate immune system which cover recognitions of non-self nucleic acids.

One mechanism is to recognise the viral DNAs in bacteria, where DNA modifications are essential is the restriction digestion system. In brief, restriction enzymes are endonucleases that recognise specific DNA sequences not found in their own genome, allowing cleavage of any non-self DNA sequence, and thus, degradation of the viral genome. Evolutionary, some bacteria 'improved' their restriction digestion system and modified their genome with m^5dC or m^6dA. Their restriction endonucleases were then targeting only sequences lacking those modifications. Nevertheless, some viruses overcame this phenomenon, and are able to modify their nucleic acids as well. In mammals, if DNA or non-methylated CpG motifs are found in the cytoplasm, it is sensed as non-self DNA (Krieg, 2002).

A newly discovered bacterial system known as RADAR (restriction by an adenosine deaminase acting on RNA) is comprised of two genes: an adenosine triphosphatase (RdrA) and an adenosine deaminase (RdrB). When phage infects cells, the RADAR system deaminates the pool of ATP, effectively inhibiting phage propagation (Duncan-Lowey et al., 2023). Although it was initially thought that the RADAR system might edit adenosines in RNA, subsequent studies showed that no increase in inosines in RNA was observed (Duncan-Lowey et al., 2023).

Interestingly, adenosine deamination in mRNA molecules has been found to be a critical component of the innate immune system in metazoans, where inosines in dsRNA regions serve as mark of self-RNAS. Recognition of dsRNA without any inosines is recognised as non-self (Junqueira et al., 2016). Recognising dsRNA leads to innate immune system activation in both vertebrates and invertebrates. Nevertheless, with different outcomes. Whereas in mammals, the recognition of dsRNA leads to Interferon (IFN) expression (Liddicoat et al., 2015a; Mannion et al., 2014; Niescierowicz et al., 2022), in invertebrates, no interferon genes are present, and it leads to the expression of immune genes (Deng et al., 2020a; L. Xu et al., 2013). Some viruses, such as rotavirus (Reoviridae), use dsRNA as their genome. Additionally, some ssRNA viruses have dsRNA as an intermediate during replication or produce replication byproducts such as defective viral genomes or transcription to create mRNAs. For example, the DNA virus Herpes simplex virus has dsRNA during the amplification of its transcriptome due to bidirectional transcription.

Sensors of non-self RNA in human

There are four classes of enzymes that recognise non-self RNA molecules: Toll-like receptors (TLR), oligoadenylate synthase-like proteins (OAS), retinoic acid-inducible gene I (RIG-I)-like receptors, and protein kinase R (PKR). TLR and RIG-I-like receptors are constitutively

expressed in cells, while the expression of OASes and PKR is induced by the activation of the innate immune system *via* interferon. As a result, TLR and RIG-I-like receptors serve as the first line of defence against pathogens in the cell, whereas OASes and PKR function as the second layer (Anderson et al., 2010; U. Y. Choi et al., 2015; Lemaire et al., 2008).

Toll-like receptors

First, in endosomes, viral RNA can be recognised by Toll-like receptors (TLR) – TLR3, TLR7 and TLR8. dsRNA is sensed by TLR3 whereas ssRNA is recognised by TLR7 and TLR8 (Y. G. Chen & Hur, 2022).

Toll-like receptors have limited expression patterns, mainly expressed in immune and epithelial cells (Martínez-Espinoza & Guerrero-Plata, 2022). TLR receptors form dimers which, upon the recognition of viral RNA, change the conformation leading to the recruitment of adaptor proteins and the activation of downstream signalling pathways. This results to the activation of expression of cytokines or interferons or TNF-alpha (Petes et al., 2017).

Protein kinase R

The dsRNA can be sensed by the protein kinase R (PKR). PKR is one of the Interferon stimulated genes which, upon expression, is inactivated and activated by the binding to dsRNA, leading to its dimerisation and autophosphorylation. Activated PKR phosphorylates the α-subunit of the translation initiation factor eIF2, leading to global translation inhibition. The substrate of PKR activation is double-stranded RNA of a minimum length of 30 nucleotides (Lemaire et al., 2008). Moreover, structured small RNAs can activate PKR activity (Nallagatla et al., 2007). Modifications in ssRNA can slightly decrease PKR activity (Nallagatla & Bevilacqua, 2008), but the effect of different the modifications is more prominent on dsRNAs. Moreover, the effect of RNA modification in dsRNA and ssRNA on PKR activation differs. For example, modifications such as s^2U, s^4U, 2′-dU and PS-G (α-Phosphorothioguanosine) in dsRNA decrease the PKR activity, whereas in ssRNA s^2U, I^5U and m^6A in ssRNA (Nallagatla & Bevilacqua, 2008).

Oligoadenylate synthase-like proteins

Another class of interferon stimulated genes are oligoadenylate synthase-like proteins (OAS), which consist of 4 isoforms OAS1, OAS2, OAS3 and OASL. OAS family proteins can

synthesise the 2′-5′- linked oligoadenylate, which leads to the activation of RNase L. Activated RNase L then degrades all classes of RNAs in the cytosol (U. Y. Choi et al., 2015).

RIG-I-like receptors (RLRs)

In cytoplasm are cytoplasmatic sensors of viral RNA - retinoic acid-inducible gene I (RIG-I), melanoma differentiation-associated protein 5 (MDA5) and laboratory of genetics and physiology 2 (LGP2). All are from a family of RIG-I receptors. Compared to Toll-like receptors, the cytoplasmatic sensors are ubiquitously expressed among all tissues.

All three RIG-I-like receptor members possess a central helicase and a carboxy-terminal domain (CTD). The central helical domain is composed of three helicase domains - Hel1, Hel2i, Hel2 and the pincer domain. RIG-I and MDA5 also possess N-terminal repeated caspase recruitment domains (CARDs). Helicase and CTD domains act together to detect immunostimulatory RNAs. Binding to immunostimulatory RNAs causes activation of the sensor via conformation change, allowing the CARD domains to associate with MAVS. Activated MAVS then trigger signalling, leading to cytokine expression such as IFN. The IFN, in turn, initiate the production of a large set of interferon-stimulated genes (ISGs) in the infected cell and those nearby to create an anti-pathogenic environment (Rehwinkel & Gack, n.d.; Schoggins et al., 2011).

Both RIG-I and MDA5 recognise dsRNA. MDA5 preferentially binds long dsRNA, whereas RIG-I the short dsRNA. Moreover, RIG-I binds the RNAs that are uncapped on their 5′end. As immunostimulatory RNAs are recognised dsRNAs lacking inosines or having unprocessed or not enough process 5′end such as 5′triphosphate (Hornung et al., 2006; Pichlmair et al., 2006), 5′diphosphate (Goubau et al., 2014; P. Kumar et al., 2013; Ren et al., 2019) or only lack the 2′O- methylation on the first transcribed nucleotide (Schuberth-Wagner et al., 2015; Wang Yanli et al., 2010; Züst et al., 2011).

Double-stranded RNA molecules in recognition of non-self

Most of the RNA inside our cells is single-stranded, meaning it′s made up of a single chain. Nevertheless, parts of mRNAs loop around due to self-complementary sequences, forming short double-stranded regions. On the other hand, many viruses generate long double-stranded RNA (dsRNA) during their lifecycle, making dsRNA a potential sign of viral invasion in our cells. This triggers an activation in our innate immune system.

For instance, some viruses, like rotavirus, a member of the Reoviridae family, have dsRNA as their genetic material (Uzri & Greenberg, 2013). Other viruses, even though their primary RNA is single-stranded, create dsRNA as an intermediate during the replication of their genome, as well as during transcription. Or for example, herpes simplex virus type 1 (HSV1), a DNA virus, employs bidirectional transcription, which also produces dsRNA (J. Zhao et al., 2021).

Our cells must avoid mistakenly recognising their dsRNA as viral-like, so our cellular RNAs are marked as 'self' by an enzyme called ADAR1. ADAR1 is a nuclear protein that binds to dsRNA in the nucleus and deaminates adenosine to inosine. This process is called A-to-I RNA editing, and it is critical to prevent our immune system from mistakenly responding to our own dsRNA, which could result in autoimmune conditions (Liddicoat et al., 2015b; Mannion et al., 2014).

Inosine plays a critical role in the self vs non-self distinction within metazoan innate immune systems, acting as a marker for self-molecules to evade recognition by the immune system. Consequently, as previously discussed, double-stranded RNA (dsRNA) sensors, such as RIG-I, MDA5, and LPG2, are designed to identify only those dsRNA molecules that lack inosine (Liddicoat et al., 2015b; Mannion et al., 2014; Stok et al., 2022). In this manner, these sensors maintain a fine-tuned balance in discerning foreign from self-entities within the organism.

The correct recognition of self-RNA molecules is crucial for maintaining immune homeostasis. A deficiency in editing by ADAR1 can lead to the misrecognition of self-RNA as non-self, triggering inappropriate immune responses. This situation is exemplified in autoimmune disorders such as Aicardi-Goutières syndrome (AGS). Mutations within ADAR1 that alter its expression pattern or editing efficiency can result in persistent interferon signalling, a hallmark of this disease (Rice et al., 2012).

The importance of *Adar1'*s role in innate immunity was demonstrated in mice null mutants. Mice *Adar1* mutants die by embryonic day E12.5 due to the overexpression of interferon, as detected by the presence of interferon-stimulated genes (Mannion et al., 2014; Liddicoat et al., 2015). The phenotype can be partially rescued by the lack of *MDA5*(Liddicoat et al., 2015a) or *Mavs* (Mannion et al., 2014) receptor. The role of *Adar1* in the innate immune response is also conserved in Zebrafish (Niescierowicz et al., 2022).

As Interferon signalling is known to be part of innate immunity only in vertebrates, it is interesting that A-to-I editing has a conserved role in being a mark of self-RNA molecules

in flies as well. Both catalytically dead and null ADAR knockouts were found to express immune-induced molecules, highlighting the essential role of ADAR in maintaining a normal immune response(Deng et al., 2020b). It recognises non-self-RNA molecules in flies by Dicer-2 (Deng et al., 2020b), an enzyme generally involved in processing small RNAs. The role of Dicer-2 in sensing on-self molecules in flies is further emphasised during viral infections, where the lack of Dicer-2 leads to heightened production of immune-induced molecules (Deddouche et al., 2008)

In conclusion, the A-to-I editing represents another critical layer in distinguishing self-RNA in dsRNAs which are conserved from fly to humans. This distinction prevents an autoimmune response, illustrating ADAR's vital role in maintaining the immune system's balance.

RNA modifications in recognition of non-self

Up to today, we know about 160 different RNA modifications (Boccaletto et al., 2018). While those have many functions, mainly in gene expression, some also function in the recognition of self and non-self-RNA molecules.

In humans, many RNA modifications contribute to distinguishing self- vs non-self-RNA molecules. Modifications abolishing the activation of cytoplasmic RIG-I family receptors are mRNA cap modifications (Hornung et al., 2006; P. Kumar et al., 2013; Ren et al., 2019; Schuberth-Wagner et al., 2015), inosine in dsRNAs (Hartner et al., 2008; Junqueira et al., 2016; Liddicoat et al., 2015a; Mannion et al., 2014; Niescierowicz et al., 2022), $M^1\psi$, ψ, s^2U, m^5U, m^5C, hm^5C, m^6A (Durbin et al., 2016). On the other hand, m^5C, m^6A, m^5U, s^2U, ψ (Karikó et al., 2005), Internal N_m modifications but C_m ablates activity of TLR receptors (Cekaite et al., 2007; Eberle et al., 2008; Robbins et al., 2001).

Internal *2'-O*-methylation in innate immunity

The internal N_m modification in mRNA has been investigated in the context of viral infection. This modification has been identified in several viruses, including Human immunodeficiency virus (HIV1) (Ringeard et al., 2019), Sudan ebolavirus (B. Martin et al., 2018), as well as Flaviviruses West Nile and Dengue virus (H. Dong et al., 2012). In the Ebola virus, the L

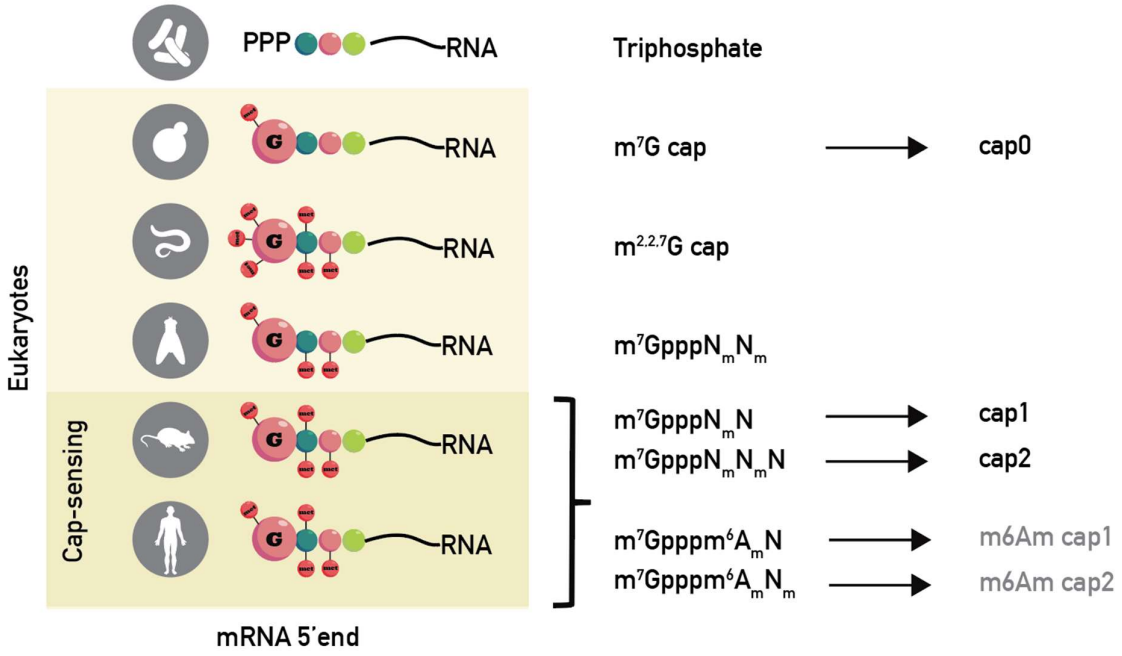


Figure 5. mRNA Cap evolution. The methylation status between various organism differ.

protein is responsible for internal N_m modifications. Similarly, in Dengue and West Nile laviviruses, the NS5 protein conducts these modifications. Both enzymes methylate not only HIV1 virus transcripts are N_m modified by the cellular enzyme FTSJ3, which helps avoid detection by MDA5 (Ringeard et al., 2019) exonuclease activity of ISG20, thereby enhancing mRNA stability upon Interferon expression (Kazzi et al., 2023).

mRNA cap modifications in recognition of non-self-RNA molecules

*A distinguishing feature of our mRNAs is the inclusion of an mRNA cap structure (***Figure 4***). Given that this specific cap structure is found in all our processed mRNAs, evolutionary differs between organisms (*

Figure *5*), makes it a sophisticated target to identify foreign RNAs.

Immunostimulatory RNAs (**Figure *6***) are recognised as dsRNAs lacking *2'O*-methylation at the first transcribed nucleotide irrespectively of cap0 (m^7GpppN-RNA) or tri- or diphosphates (pppN-RNA and ppN-RNA) at the 5′ end (Goubau et al., 2014; Hornung et al., 2006; P. Kumar et al., 2013; Pichlmair et al., 2006; Schuberth-Wagner et al., 2015; Wang Yanli et al., 2010; Züst et al., 2011). mRNA cap structure was shown to be recognised by RIG-I (Schuberth-Wagner et al., 2015) and MDA5 (Züst et al., 2011). *2'O*-methylation at the first transcribed nucleotide is crucial to prevent RIG-I activation as it presents reduced binding and, as a consequence, interferon pathway activation, as the methyl group clashes with its conserved

histidine (H830) (Devarkar et al., 2016; Schuberth-Wagner et al., 2015; Wang Yanli et al., 2010). Interestingly, transfection of mRNAs having only a second transcribed nucleotide 2′O-methylated does not completely but significantly block IFN production as $m^7GpppNN_m$-RNA as well as $pppNN_m$-RNA compared to non-self RNA molecules (Schuberth-Wagner et al., 2015; Wang Yanli et al., 2010). Similarly, KO of CMTR2 in HEK cells showed mild elevation of Interferon stimulated genes (Despic & Jaffrey, 2023). Cap1 methyltransferase, CMTR1, belongs to the ISG genes, which is why its other gene name is ISG95 (Shaw et al., 2017; Williams et al., 2020). Interestingly, stimulation of human cells with Interferon-β leads to lower expression of ISGs upon siRNA-mediated KD of CMTR1. Moreover, KD of *CMTR1* leads to

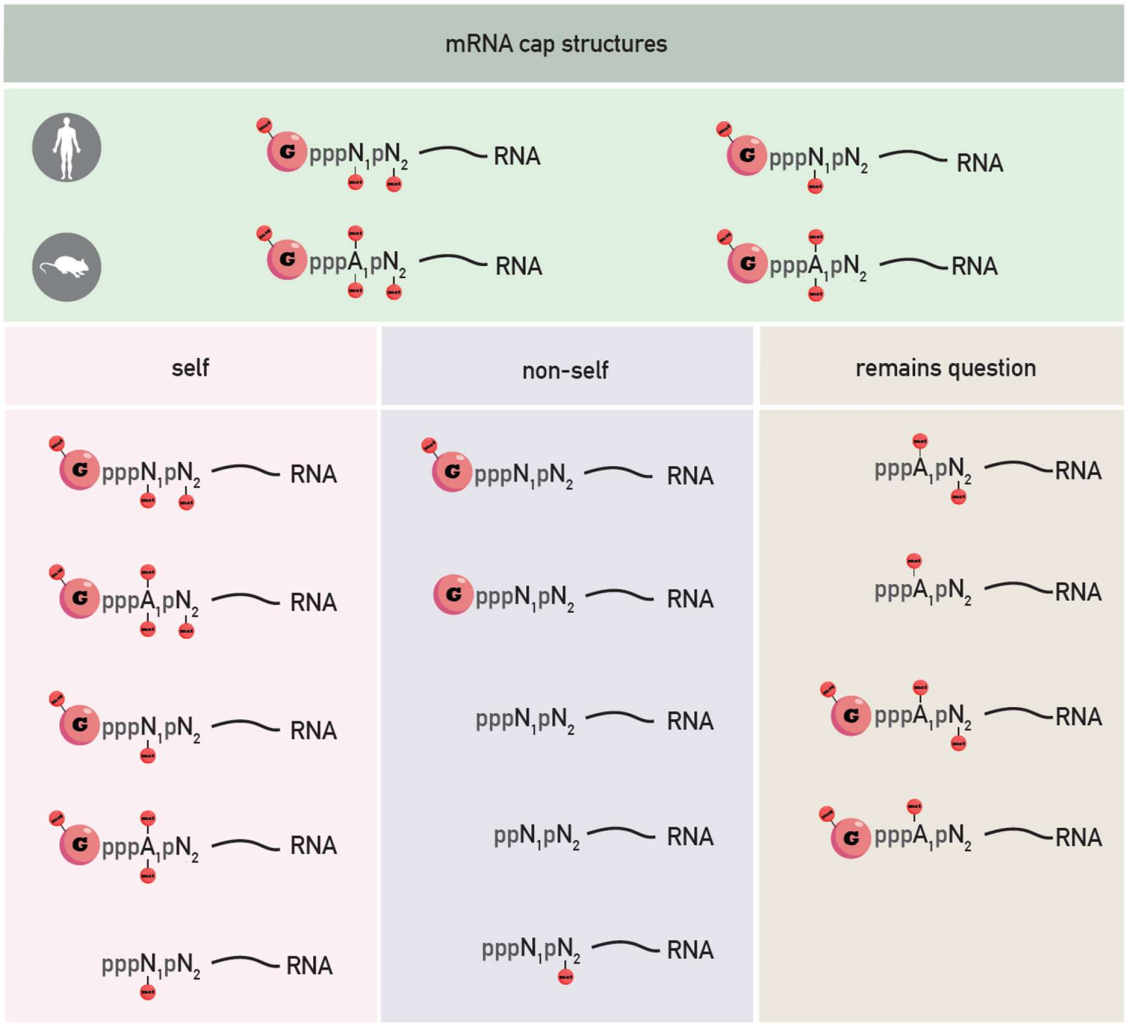


Figure 6. Sensing of the mRNA cap structure. mRNA cap structure prevents recognition of RNA molecule as non-self when 2′O-methylated (methylation mark aiming to the bottom) at first transcribed nucleotide. Interestingly, m6A (methylation mark aiming to the top) was never studied at the first transcribed nucleotide in recognition of self- and non-self-RNA molecules

higher propagation of positive-sense RNA viruses, Zika virus (ZIKV) and dengue virus (DENV) (Williams et al., 2020). Both ZIKV as well as DENV carry their own cap1 *2′O*-methyltransferase. This proposes a function of CMTR1 to increase the expression of ISGs which may slow down the replication of viruses.

Conversely, some viruses carry their own viral methyltransferase responsible for cap1 mRNA structure. Their mutations lead to decreased growth of Dengue virus (H. Dong et al., 2010), Yellow fever virus (Schuberth-Wagner et al., 2015), Coronavirus (Züst et al., 2011), West Nile virus or vaccinia virus (Daffis et al., 2010).

There are a few interferon-induced cap-binding proteins, such as IFIT1 and IFIT5 (Abbas et al., 2017; Habjan et al., 2013; Miedziak et al., 2020; Pichlmair et al., 2011). N_m at the first position was shown to increase the translation due to IFIT1, which competitively inhibits the translation machinery of unmethylated RNA (Abbas et al., 2017; Habjan et al., 2013). Interestingly, IFIT1 binds with a stronger affinity to A-starting transcripts than G-starting (Miedziak et al., 2020).

Self-RNA molecules recognised as non-self

RIG-I and MDA5 can also bind the self-RNA s with signatures of non-self. RIG-I binds the vault RNAs under the infection of Kaposi′s Sarcoma-associated herpesvirus (Y. Zhao et al., 2018). Newly transcribed vault RNAs possess the 5′-triphosphate, which is DUSP11 processes. It′s activity results in 5′ monophosphorylated RNAs. Nevertheless, during the Kaposi′s Sarcoma-associated herpesvirus infection, the vault RNAs are triphosphorylated, probably due to the decreased levels of DUSP11 (Burke & Sullivan, 2017). Interestingly, the triphosphorylated vault RNA blocks the Kaposi′s Sarcoma-associated herpesvirus lytic reactivation (Y. Zhao et al., 2018).

How viruses try to avoid recognition

Viruses utilise various mechanisms within host cells to evade recognition as pathogens invading the cell. One key factor for recognising non-self is the mRNA cap structure, and many viruses have found ways to utilise it. There are four primary mechanisms that viruses use to avoid recognition, with some viruses using more than one mechanism. Firstly, some viruses with a nuclear replication cycle utilise cellular transcription machinery, resulting in viral mRNAs being capped and modified like cellular mRNAs. Secondly, certain cytoplasmic viruses encode their own RNA-capping machinery. The third mechanism is to acquire the cap

structure from cellular mRNA via a process called cap snatching. The fourth option involves encoding proteins that covalently bind to the 5′ ends of mRNA to prevent recognition by RIG-I-like receptors (**Table *1***). Moreover, to avoid sensing, some viruses also inhibit the innate immunity pathway (Bowie & Unterholzner, 2008; H.-C. Lee et al., 2019; Simmons et al., 2013).

Nuclear transcription using cellular machinery

Many DNA viruses, including both DNA viruses and retroviruses such as Polyomavirus, Papillomavirus, Parvovirus, Hepatitis B Virus (HBV), and Human Immunodeficiency Virus (HIV), undergo a nuclear replication cycle. This allows them to utilise the host′s cellular transcription machinery for their propagation. Specifically, these viruses can employ the host′s DNA-dependent RNA polymerase. This interaction enables the modification of viral RNA in a manner similar to host RNA (Baltimore, 1971; K. H. Choi, 2012).

Encoding own capping machinery

Numerous viruses, particularly those with a cytoplasmic life cycle, encode their own RNA-modifying enzymes as part of their replication and transcription machinery. This adaptation is crucial as RNAs, without any modifications, would be swiftly recognised as foreign by host cells, leading to an immune response. To circumvent this, such viruses come equipped with their own machinery for capping RNA. In humans, the placement of cap1 modification relies on three enzymes (RNGTT, RNMT, CMTR1), whereas viruses typically utilise one or two enzymes with the necessary catalytic domains for these activities.

Certain viruses, such as poxvirus and coronaviruses, utilise two enzymes for this process. Poxvirus carries the genes D1R and D12L (Cobb, 2017; Crick, 1970; Farlow et al., 2010), while coronaviruses employ NSP14 and NSP16 (Bobrovs et al., 2021; Pan et al., 2022; Wilamowski et al., 2021) to deposit cap0 and cap1 modifications. Conversely, some viruses consolidate both cap0 and cap1 activities into a single protein. An example is the rotavirus, which uses the protein VP3 for this purpose (Ogden et al., 2014). This mechanism is considered less efficient than the cellular capping enzymes (Moreno-Contreras et al., 2022; Uzri & Greenberg, 2013).

Table 1. Key replicative characteristics, viral RNA 5′ modifications, and responsible enzymes in diverse families of mammalian viruses. The table lists selected examples of viral families that infect mammals, along with the location of their lifecycle, the detected 5′ end modifications of viral RNA, and the enzymes responsible for these modifications. Note the unique ′cap-snatching′ mechanisms employed by the Orthomyxoviridae and certain Bunyavirales species. ′RT′ stands for reverse transcriptase, and ′DdRp′ refers to DNA-dependent RNA polymerase.

GENOME		FAMILY	EXAMPLES	REPLICATION	RNA 5'END	RNA 5'END PROCESSING ENZYME
DNA	dsDNA	*Polyomaviridae*	Poliomavirus JV virus	nucleus	cap1	host DdRp
		Papillomaviridae	Human papilomavirus (HPV)	nucleus	cap1	host DdRp
		Adenoviridae	Adenovirus serotype 2	nucleus	cap1	host DdRp
		Herpesviridae	Herpes simplex virus (HSV-1)	nucleus	cap1	host DdRp
		Poxviridae	Variola virus	cytoplasm	cap1	D1R, D12L
	ssDNA	*Parvoviridae*	Canine parvovirus	nucleus	cap1	host DdRp
RNA	dsRNA	*Reoviridae*	Rotavirus	cytoplasm	cap1	VP3
	(+)ssRNA	*Picornaviridae, Caliciviridae, Hepeviridae*	Poliovirus, Norovirus, Hepatitis E virus	cytoplasm	peptide-RNA	VPg
		Togaviridae, Flaviviridae	Chikungunya virus, Yellow fever virus	cytoplasm	cap1	NS5
		Arteriviridae, Coronaviridae	SARS, MERS, SARS-CoV-2	cytoplasm	cap1	NSP14, NSP16
	(-)ssRNA	*Rhabdoviridae, (Paramyxoviridae, Bornaviridae)*	Rabies virus, Vesicular stomatitis virus, Mumps virus	cytoplasm	cap1	L protein
		Filoviridae	Ebola	cytoplasm	cap1	L protein
		Orthomyxoviridae	Influenza virus A	nucleus	cap1	RdRp complex (PA–PB1–PB2 subunits) - cap snatching
		Bunyavirales	Lassa virus	cytoplasm	cap1	L protein - cap snatching
	RNA w/ RT	*Retroviridae*	Human immunodeficiency virus (HIV)	both	cap1	host DdRp
	DNA w/ RT	*Hepadnaviridae*	Hepatitis B virus	both	cap1	host DdRp

Certain viruses, such as Flaviviruses, VSV (Rhabdoviridae), and Ebola virus (Filoviridae), utilise specific proteins that bear all the catalytic activities required for replication, transcription, and RNA capping in one single enzyme. Flaviviruses employ the NS5 (L. Liu et al., 2010). At the same time, VSV (*Rhabdoviridae*) (Ogino & Green, 2019) or Ebola virus (*Filoviridae*) (Martin et al., 2018) rely on the L protein, both proteins bearing all the catalytic activities for replication/transcription and RNA capping.

In conclusion, despite variations in the number of proteins utilised, viruses' ultimate goal is to evade recognition by the host cell. Whether it involves employing multiple enzymes or consolidating activities into a single protein, viruses adapt their replication and transcription machinery to ensure escape from host immune responses. This highlights the remarkable ability of viruses to manipulate host processes and underscores the importance of studying these mechanisms for developing effective antiviral strategies.

Cap snatching

In the mechanism known as ′cap-snatching,′ certain viruses exploit their viral proteins to cleave the 5′ terminal RNA of the host cell, which bears the cap1 structure, and use it as a primer for its own transcription. This strategy culminates in generating hybrid molecules, integrating elements from both the host and viral mRNA.

A subset of viruses, including influenza viruses and viruses from the *Bunyavirales* order, such as Lassa virus, hantaviruses, and arenaviruses, employ cap-snatching. However, notable differences exist between these two categories. Influenza viruses, for instance, localise to the nucleus, where their RNA-dependent RNA polymerase (RdRp) interacts with RNA polymerase II, cleaving the 5′ ends of small nuclear RNAs (snRNAs) or pre-mRNAs. In contrast, bunyaviruses localise to the cytoplasm, where they cap-snatch mRNA (Olschewski et al., 2020).

Moreover, influenza viruses rely on an RdRp protein complex consisting of three different proteins—PA, which contains an endonuclease (EN); PB1, serving as the RdRp; and PB2, featuring a cap-binding domain (CBD) (De Vlugt et al., 2018; Olschewski et al., 2020). In contrast, Bunyavirales use a single L protein encompassing all three domains (Olschewski et al., 2020). A fascinating example is the mosquito-transmitted Rift Valley fever virus, a bunyavirus that preferentially cap-snatches mRNA terminal oligopyrimidine (TOP) transcripts (Hopkins et al., 2015).

Interestingly, the Influenza virus, as it is dependent on CMTR1 to make cap1 RNA structure, was shown that the KD of *CMTR1* inhibits viral replication and up-regulates anti-viral genes in human fibroblasts infected with the Influenza virus (B. Li et al., 2020). Moreover, a selective CMTR1 inhibitor was shown to abrogate the capping of the Influenza virus and, thus, its replication in cells and mice (Tsukamoto et al., 2023). In summary, the cap-snatching mechanism employed by various viruses underscores the complex interplay between host and pathogen, shedding light on potential therapeutic targets such as CMTR1, whose inhibition could stifle viral replication and bolster the host's anti-viral response.

Viral cap-binding proteins

The known viruses without m^7G cap structure belong to +ssRNA viruses, such as *Picornaviridae, Caliciviridae, and Hepeviridae.* Those viruses either possess monophosphate at their 5′end or covalently attached protein to the 5′end (Nomoto et al., 1976). Such a protein in Poliovirus is called viral protein genome-linked (VPg). VPg is by its tyrosine linked by phosphodiester linkage to viral RNA (Ambros & Baltimore, 1978; Crawford & Baltimore, 1983). Protein-bound to the 5′end of mRNA diminishes recognition of such RNA by cytoplasmic RIG-I-like sensors (Furuichi, 2015).

Inhibition of the innate immune system

In addition to evading recognition by cellular machinery by modifying their RNA caps, viruses have developed strategies to inhibit crucial cellular processes, including those involved in viral nucleic acid detection and interferon response. IFNs play a pivotal role in the innate immune response by triggering antiviral defences and orchestrating the immune system′s response to viral infections. However, viruses have evolved sophisticated mechanisms to subvert these defence mechanisms, enabling them to establish successful infection and replication within host cells.

For instance, Influenza, a common respiratory virus, encodes a protein known as non-structural protein (NS1). NS1 can directly bind to and decrease the activation of RIG-I, a sensor of non-self-RNA (Pichlmair et al., 2006). This decreases the host cell′s ability to detect and respond to the virus, thereby promoting viral replication (Jureka et al., 2020).

In a similar fashion, the Rotavirus, a common cause of severe diarrhoea in young children, uses its Viral Protein 3 (VP3) to inhibit the host′s innate immune response. The VP3 protein is multifunctional; it not only acts as an mRNA capping enzyme but also plays a significant role in the phosphorylation of MAVS, another sensor of viral infections. The

phosphorylation leads to the degradation of MAVS, further inhibiting the host′s immune response. Moreover, VP3 also possesses a 2′-5′-phosphodiesterase (PDE) domain that can degrade 2′-5′-oligoadenylate (2-5A), an activator of RNase L, effectively preventing its activation. This leads to the inhibition of another level of innate immunity (Song et al., 2020; R. Zhang et al., 2013, Ding et al., 2018)

The Zika virus, particularly a Brazilian isolate, has evolved a different strategy to counter the host′s immune response. Its protein NS5 has been demonstrated to disrupt the interferon-α/β receptor (IFNAR) signalling, an integral part of the interferon response, by depleting signal transducer and activator of transcription 2 (STAT) and blocking STAT1 phosphorylation. Upon activation, IFNAR dimerises activating downstream JAK/STAT signalling and thus abrogating induction of ISGs expression (Hertzog et al., 2018).

In addition to the above examples, NS5 of the Dengue virus also demonstrates an antagonistic effect on the innate immune system. NS5 can associate with and target for degradation of STAT2, inhibiting STAT2-dependent ISGs expression (Ashour et al., 2010; Morrison & García-Sastre, 2014).

Similarly, the Hepatitis C virus (HCV) has also evolved strategies to inhibit interferon signalling and other innate immune responses. One such mechanism involves the HCV NS3/4A protease, which cleaves Toll-like receptor 3 adaptor protein TRIF and thus reduces signalling via TLR3 (K. Li et al., 2005). The second target of NS3/4A protease is MAVS. Cleavage of MAVS leads to its dissociation from the mitochondrial membrane, abrogating the activation ability of MAVS, thereby suppressing the induction of IFN. Moreover, HCV employs another strategy to disrupt innate immunity. This interaction by NS5A impedes the PKR′s activation process, usually involving its dimerisation (Gale et al., 1998). Both NS3/4A and NS5A play crucial roles in antagonising the activation and the signalling of the innate immune system, thus aiding the survival of the virus in the host cell.

The examples provided demonstrate just a fraction of the diverse strategies viruses use to dodge removal by the host cell, thereby boosting their chances of survival. This diversity across different viruses underscores the vital role of these mechanisms in the viruses′ success. Viruses exemplify the continual struggle throughout evolutionary history between pathogens and their hosts by developing these varied methods to sidestep and impair the innate immune response.

RNA modifications in the development of RNA therapy

Understanding the role of RNA molecules in protecting being recognised as non-self RNA molecules and their molecular properties showed their importance in developing RNA-based therapies. As shown in early studies of Katalyn Kariko and later her group, transfection of unmodified RNA molecules leads to the expression of interferon. In contrast, RNA with RNA modifications (m^5C, m^6A, m^5U, s^2U, or Ψ) leads to decreased to non-detectable expression of interferon via the TLR receptors (Karikó et al., 2005). Moreover, incorporating Ψ into the transfected RNA enhances translation yields (Karikó et al., 2008) due to the enhanced stability of RNA (Davis, 1995; Karikó et al., 2008). Moreover, Ψ diminishes PKR activation. Since PKR activation leads to the abrogation of protein synthesis, the presence of Ψ in mRNAs enhances the translation of transfected mRNA (Anderson et al., 2010). Furthermore, the presence of *N1*-Ψ ($m^1Ψ$) outperforms the Ψ and provides enhanced protein expression in translation activity (Andries et al., 2015). This is even more prominent if transfected mRNA contains $m^1Ψ$ in combination with m^5C (Andries et al., 2015).

Similarly, transfection of siRNAs activates the innate immune response via the TLR receptors activation. Interestingly, the substitution of uridines by U_m ablates the activation (Cekaite et al., 2007). Moreover, further studies showed that any N_m modification but C_m modification show the same effect (Eberle et al., 2008; Robbins et al., 2001).

In essence, understanding RNA modifications plays a pivotal role in the advancement of RNA-based therapies. Modified RNA molecules, such as those with m^5C, m^6A, m^5U, s^2U, or Ψ alterations, demonstrate decreased interferon expression and enhanced translational yields, offering a path to suppress the innate immune response. This knowledge not only augments our understanding of RNA behaviour but also catalyses the development of more potent and sophisticated RNA therapeutics and vaccines.

Mouse embryonic development

The intricate landscape of RNA modifications, as explored in the previous chapter, reveals their pivotal roles in gene expression regulation. Remarkably, perturbations in various RNA-modifying enzymes have been linked to a striking outcome: embryonic lethality. This chapter bridges the gap between RNA modifications and mouse embryonic development, delving into the profound impact of RNA-modifying enzymes on embryonic viability.

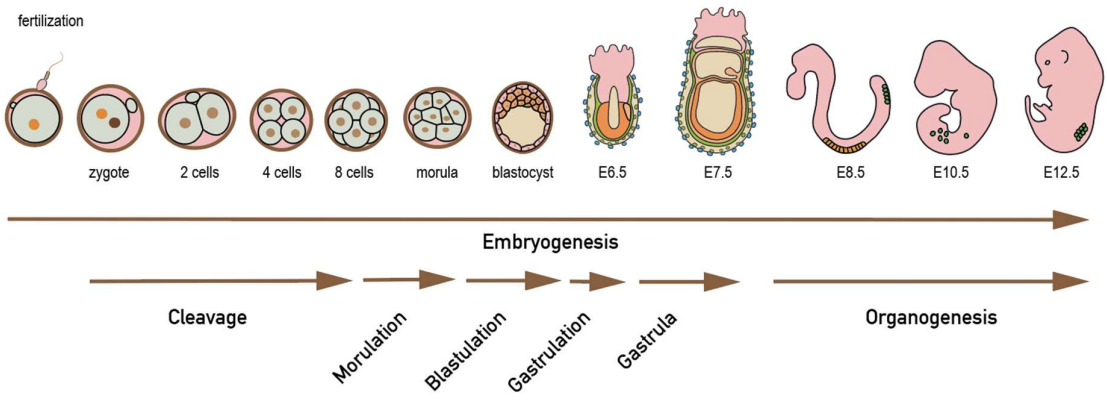


Figure 7. Embryogenesis. Embryogenesis in mice consists of 6 main stages, cleavage, morulation, blastulation, gastrulation, gastrula and organogenesis. E stands for Embryonic day when the stage is observed in a mouse. The green cells at organogenesis stages represent primordial germ cells.

Mouse embryonic development begins with the fusion of a sperm and an egg, resulting in a fertilised egg or zygote. The zygote then undergoes a series of cell divisions, forming a solid ball of cells called a morula (Kojima et al., 2014; Lei et al., 2020). The morula develops into a blastocyst, which consists of an outer layer of cells called the trophoblast and an inner cell mass. After the blastocyst stage, gastrulation is the next crucial step in mouse embryonic development. Gastrulation involves the rearrangement of cells within the embryo, forming three distinct germ layers: ectoderm, mesoderm, and endoderm. This process establishes the foundation for organogenesis, which follows gastrulation. Organogenesis is the period during which the germ layers differentiate and give rise to specific organs and tissues. Through a complex series of coordinated cellular interactions and morphogenetic events, organs such as the brain, heart, liver, and limbs begin to form, leading to the development of a fully functional mouse organism (Nakatsuji, 1992; Obata et al., 2000) (**Figure 7**).

Among the mRNA-modifying enzymes crucial for embryonic development are m^6A methyltransferases *Mettl16* and *Mettl3*. *Mettl16* mutants die during pre-implantation stages around E3.5. The underlying cause of lethality in these mutants is believed to be the absence of m^6A modification, specifically on the *Mat2a* transcript. *Mat2a* encodes an enzyme involved in generating the methyl group donor, S-adenosylmethionine (SAM), which is essential for various methylation reactions in the developing embryo. The loss of m^6A modification on the *Mat2a* transcript disrupts the production of SAM, leading to the abrogation of all methylations critical for embryonic development (Mendel et al., 2018). Interestingly, compared to *Mettl16*, which has been shown to methylate only hand full of transcripts, *Mettl3/14* complex methylates

most m^{6}A sites. Nevertheless, in comparison to *Mettl16* mutants, *Mettl3* mutants demonstrate a distinct pattern of embryonic lethality occurring at a later stage, specifically during implantation around E5.5 Studies have revealed that the presence of m^{6}A modification on pluripotency marker mRNAs leads to their decreased stability and reduced translation, thereby promoting the decay of these markers. In the case of Mettl3 mutants, the absence of proper m^{6}A modification results in the impaired regulation of pluripotency markers, preventing the embryo from progressing beyond E5.5 (Geula et al., 2015).

Another mRNA-modifying enzyme crucial for embryonic development is ADAR1. Adar1 mouse mutants die during organogenesis stages at E12.5 (Liddicoat et al., 2015a; Mannion et al., 2014). The embryonic lethality of Adar1 mutants is due to the activation of the innate immune system and Interferon expression, leading to erythropoiesis and liver failure (Liddicoat et al., 2015a; Mannion et al., 2014).

Other crucial mRNA-modifying enzymes are cap-specific modifiers, RNGTT, RNMT, CMTR1, CMTR2 and PCIF1. Notably, *Rnmt, Cmtr1, Cmtr2,* and *Pcif1* mouse mutants have been established, whereas no phenotype for *Rngtt* mutant mice was reported. *Rnmt* (Groza et al., 2023)*, Cmtr1* (Y. L. Lee et al., 2020) and *Cmtr2* (Groza et al., 2023) mutants are reported to cause embryonic lethality prior to organogenesis, whereas no embryonic phenotype was observed in *Pcif1* mutants (Pandey et al., 2020). It is important to note that the database (Groza et al., 2023) and the study (Y. L. Lee et al., 2020) have not explicitly determined the exact time point of embryonic lethality and the underlying causes.

To summarise, the interplay between RNA modifications and mouse embryonic development unveils the essential roles of mRNA-modifying enzymes in ensuring proper embryogenesis. The loss of specific modifications, such as m^{6}A and cap modifications, due to mutations in enzymes like METTL16 or METTL3 and some cap-specific modifiers leads to embryonic lethality at various stages. These findings underscore the intricate regulatory mechanisms orchestrated by RNA modifications during the crucial developmental stages of mouse embryos.

Mouse primordial germ cells

Similarly, as for embryonic development, RNA modifications were shown to be crucial for male and female fertility at different stages of the development of primordial germ cells.

In mice, primordial germ cells (PGCs) relocate from the base of the allantois to the genital ridges through a series of sequential movements. Initially forming a cluster of approximately

40 cells at around embryonic day 7.25 (E7.25), the PGCs migrate to the hindgut endoderm at E7.75, then progress to the mesentery at E9.5 and ultimately colonise the genital ridges by E10.5 (**Figure 7**). During the proliferative phase of PGCs in both males and females, an important event known as epigenetic reprogramming occurs. This process involves genome-wide DNA demethylation, which includes the erasure of genomic imprinting (Saitou et al., 2002; Saitou & Yamaji, 2012). The further development of male and female germ cells differs.

In male embryos, XY PGCs enter mitotic arrest upon entry to the genital ridges and stay in the G0/G1 phase of the cell cycle until birth to then resume proliferation around day P2-P5 (Saitou & Yamaji, 2012). The first wave of spermatogenesis is synchronised during testis maturation, with cells entering meiosis around P10. At prophase I of meiosis, spermatocytes are first detected in the testis at about day 10, Zygotene at P12 and Pachytene at P14. The spermiogenesis happens after P20 when round spermatids differentiate from elongated ones. Around 6-8 weeks, males reach sexual maturity (**Figure 8**) (Fayomi & Orwig, 2018).

In female XX embryos, primordial germ cells initiate the first phase of meiosis, prophase I, around the E13.5. By the E15.5, the majority of germ cells in the developing ovaries have progressed through the initial stages of meiotic prophase I, which include the leptotene, zygotene, and pachytene stages. By the E17.5, these cells pause their development at the diplotene stage of prophase I, becoming diplotene oocytes. Diplotene oocytes are located in primordial follicles. As the follicle matures, it becomes a primary follicle (P5), with the oocyte growing and the surrounding cells changing shape. The primary follicle then develops into a

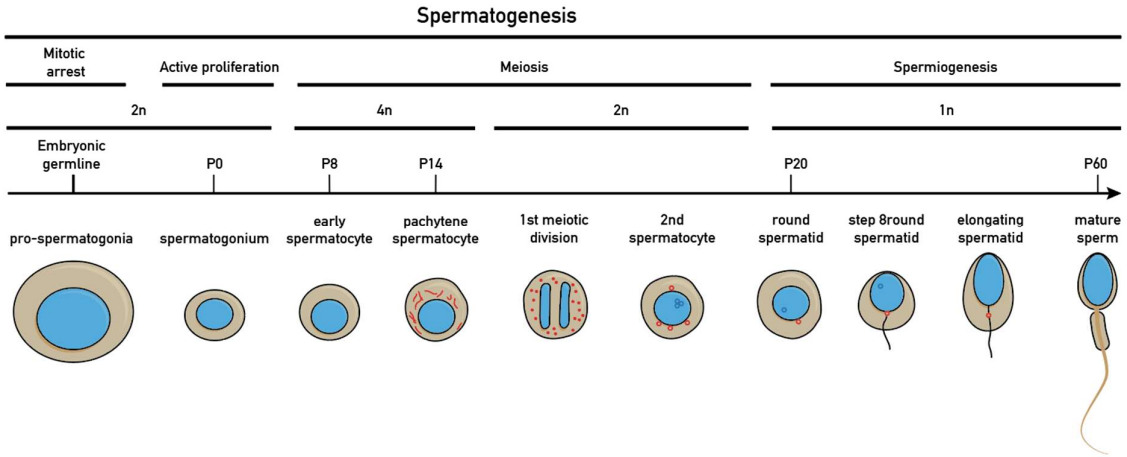


Figure 8. Mouse spermatogenesis. Embryonic germ cells are mitotically arrested till birth. Later the germ cells go through active proliferation, meiosis, and spermiogenesis, where the spermatids go through the last developmental changes to become fully developed sperms. 4n, 2n and 1n stand for DNA content. P0, P8, P14, P20, P60 stand for the animal's age abbreviated from postnatal day.

secondary follicle, characterised by multiple layers of surrounding cells and a small fluid-filled space. This space expands as the follicle grows, forming an intermediate called a pre-antral follicle and, finally, a fully formed antral follicle. Under the influence of hormones, some antral follicles typically outgrow the others to become preovulatory or Graafian follicles. The dominant follicle then releases its oocyte during ovulation. The released oocyte is called the secondary oocyte, which is arrested in the metaphase stage of meiosis II until fertilisation occurs (**Figure 9**) (Elvin, 1998; Saitou & Yamaji, 2012).

Among the mRNA-modifying enzymes crucial for fertility belongs METTL3 and METTL16. As both of the enzymes are embryonically lethal, the creation of conditional knockout (cKO) mice was needed. The fertility of METTL16 was studied in the presence of germline-specific expression of Mvh-Cre, which starts to be expressed around 14.5. cKO males are infertile as the cKO testes were atrophied, and the arrested germ cell development was observed. The effect of the lack of METTL16 on female fertility was not studied (Mendel et al., 2021).

In female mice, the lack of METTL3 during early oocyte development *(Mettl3f/fVasaCre+)* led to abnormal ovary morphology and sterility. In contrast, knockout at a later stage (*Mettl3f/f Zp3Cre+)* resulted in normal ovary morphology but still caused sterility. In male mice, early-stage knockout *(Mettl3f/fStra8Cre+)* significantly reduced testis volume, while knockout during the spermatid stage *(Mettl3f/fPrm1Cre+)* had no impact on fertility or seminiferous tubules morphology (Lasman et al., 2020).

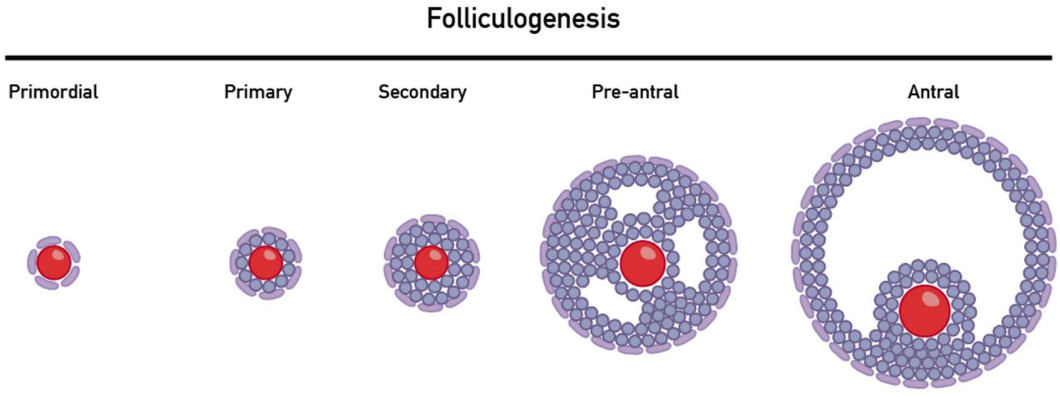


Figure 9. Mouse Folliculogenesis. Folliculogenesis begins with the development of the primordial follicle, which progresses to the antral follicle stage through primary, secondary, and pre-antral follicle stages.

Interestingly, compared to *Mettl3* KO mice that are lethal, Zebrafish *zMettl3* mutants are viable. Mutant fish of both sexes show defects in fertility. Males show defects in sperm maturation, and sperm motility is significantly reduced, whereas in females, oocyte maturation is affected. Nevertheless, defects in oocyte maturation in mutants can be rescued by sex hormone (Xia et al., 2018). Interestingly, many m^6A reader proteins showed fertility phenotype, such as RRC2A (Tan et al., 2023), YTHDF2 (Lasman et al., 2020) or m^6A eraser ALKBH5 (Zheng et al., 2013)

The disruption of RNA-modifying enzymes, such as *Mettl3* and *Mettl16*, significantly impacts the development of primordial germ cells and fertility in mice, underscoring the essential role of RNA modifications in these processes.

Aims of the study

The study's primary aim was to investigate the physiological role of mammalian RNA cap-proximal ribose methylations, specifically cap1 and cap2, in mice models and human cell culture and their significance in mouse embryonic development. The study focused on understanding the importance of cap-specific ribose methylations by examining the consequences of losing the responsible enzymes CMTR1 and CMTR2.

Additionally, I intended to determine the relationship between the loss of cap1 and the activation of the interferon pathway by studying the response of the interferon pathway in the absence of cap1 methylation in the liver conditional knockout model.

Finally, my aim was to explore the impact of cap1 loss on germ cell differentiation and gain insights into the involvement of cap1 methylation in the differentiation process of germ cells.

In summary, the study aimed to investigate the physiological functions of CMTR1 and CMTR2 and give insights into their roles in the innate immune system and beyond.

Results

Chapter 1 - Essential roles of RNA cap-proximal ribose methylation in mammalian embryonic development and fertility

This chapter contains a peer reviewed article published in Cell Reports with a name "Essential roles of RNA cap-proximal ribose methylation in mammalian embryonic development and fertility" published in July 2023. In this study we demonstrated the critical roles of RNA methyltransferases CMTR1 and CMTR2 in mouse embryonic development. Our findings indicated that both enzymes are crucial, for the switch from gastrulation to organogenesis stages. Interestingly, this developmental arrest in both *Cmtr1* and *Cmtr2* mutants occurred without triggering the innate immune response, suggesting that the role of these enzymes extends beyond merely distinguishing cellular RNAs as self. Furthermore, our study showed that conditional removal of CMTR1 in the mouse germline results in infertility. In addition, we found that a conditional deletion of CMTR1 leads to chronic activation of the interferon pathway in the liver, demonstrating the multifaceted roles these enzymes play in mammalian biology.

I maintained and genotyped all 4 mice colonies, organised all mouse experiments dissected organs as wells as embryos, perofimed cell lines expesiments and extracted RNA with the help of Fabienne Fleury-Olela. All bioinformatics analysis was done by Kyrylo Krasnikov and some preliminary ones by David Homolka. Mateusz Mendel designed *Cmtr1* loxP animals, Lingyun Li purified DHX15 and CMTR1 and confirmed their interaction, Olesya Panasenko performed ribosomal and polysomal profiling, Cathrine Broberg Vagbø conducted RNA MS analysis, RNA-seq libraries were prepared in Genomics Core Facility at EMBL Heidelberg or iGE3 Genomics Platform. Histology slides were prepared at Histology Facility at UNIGE. Illustrations of embryos, and graphical abstract were created by Nicolas Roggli.

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.

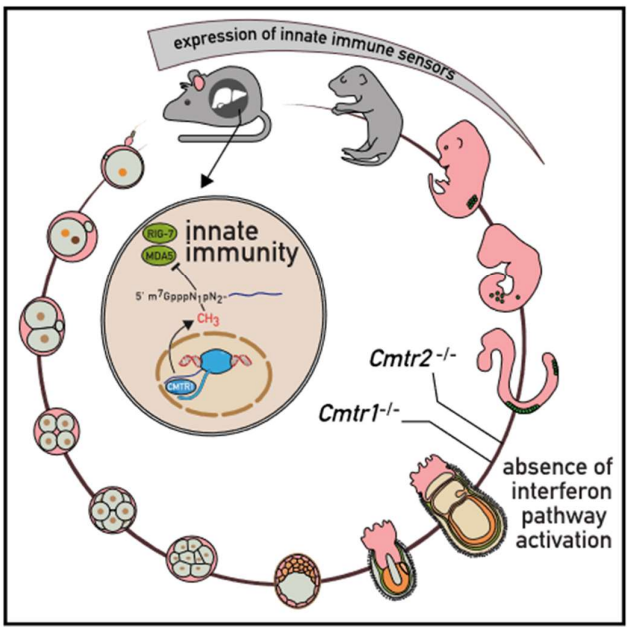
The manuscript was not modified for the purpose of this thesis, and thus, the figure numeration and bibliography are separate from the rest of the thesis.

Cell Reports

Report

Essential roles of RNA cap-proximal ribose methylation in mammalian embryonic development and fertility

Graphical abstract



Authors

Michaela Dohnalkova, Kyrylo Krasnykov, Mateusz Mendel, ..., Cathrine Broberg Vågbø, David Homolka, Ramesh S. Pillai

Correspondence

ramesh.pillai@unige.ch

In brief

Mammalian RNA cap-proximal ribose methylations are implicated in preventing the activation of the interferon pathway. Dohnalkova et al. reveal that loss of the mouse RNA methylases CMTR1 and CMTR2 causes embryonic developmental arrest without activation of the interferon pathway, pointing to gene regulatory roles.

Highlights

- CMTR1 and CMTR2 are essential for mouse embryonic development
- Absence of interferon pathway activation in arrested mutant mouse embryos
- Chronic activation of the interferon pathway in *Cmtr1* mutant livers
- Conditional loss of *Cmtr1* in the mouse germline leads to infertility

Dohnalkova et al., 2023, Cell Reports *42*, 112786
July 25, 2023 © 2023 The Author(s).
https://doi.org/10.1016/j.celrep.2023.112786

CellPress

0

Cell Reports

CellPress
OPEN ACCESS

Report

Essential roles of RNA cap-proximal ribose methylation in mammalian embryonic development and fertility

Michaela Dohnalkova,[1] Kyrylo Krasnykov,[1] Mateusz Mendel,[1] Lingyun Li,[1] Olesya Panasenko,[2] Fabienne Fleury-Olela,[1] Cathrine Broberg Vågbø,[3] David Homolka,[1] and Ramesh S. Pillai[1,4,*]

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

[2]Department of Microbiology and Molecular Medicine, Faculty of Medicine, University of Geneva, 1 Rue Michel Servet, 1211 Geneva 4, Switzerland

[3]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

[4]Lead contact

*Correspondence: ramesh.pillai@unige.ch

https://doi.org/10.1016/j.celrep.2023.112786

SUMMARY

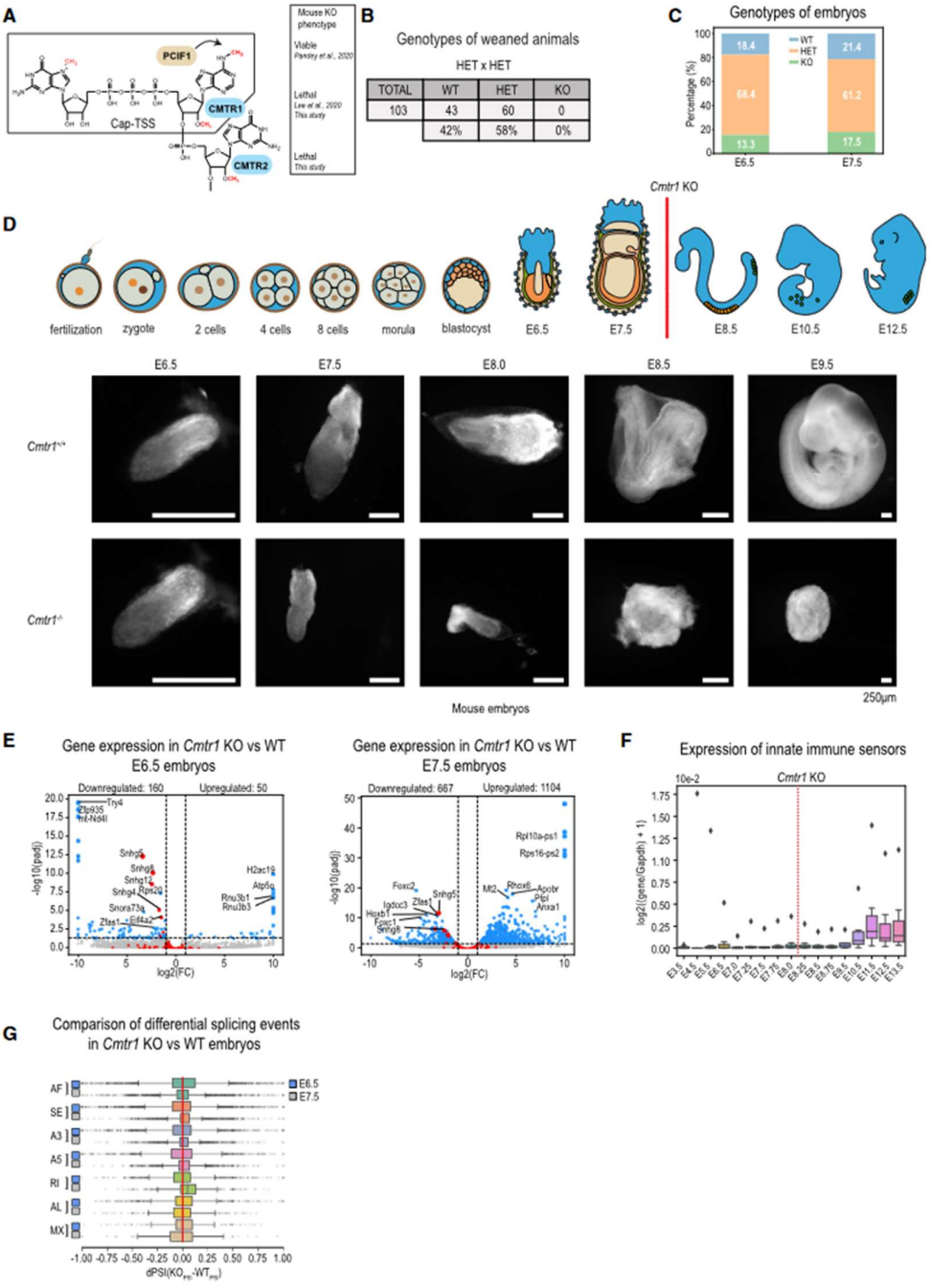
Eukaryotic RNA pol II transcripts are capped at the 5′ end by the methylated guanosine (m^7G) moiety. In higher eukaryotes, CMTR1 and CMTR2 catalyze cap-proximal ribose methylations on the first (cap1) and second (cap2) nucleotides, respectively. These modifications mark RNAs as "self," blocking the activation of the innate immune response pathway. Here, we show that loss of mouse *Cmtr1* or *Cmtr2* leads to embryonic lethality, with non-overlapping sets of transcripts being misregulated, but without activation of the interferon pathway. In contrast, *Cmtr1* mutant adult mouse livers exhibit chronic activation of the interferon pathway, with multiple interferon-stimulated genes being expressed. Conditional deletion of *Cmtr1* in the germline leads to infertility, while global translation is unaffected in the *Cmtr1* mutant mouse liver and human cells. Thus, mammalian cap1 and cap2 modifications have essential roles in gene regulation beyond their role in helping cellular transcripts to evade the innate immune system.

INTRODUCTION

RNA polymerase II transcripts receive a 5′ methylated guanosine (m^7G) cap that is attached via an inverted 5′–5′ triphosphate linkage to the transcription start site (TSS) nucleotide.[1–7] The m^7G cap is required for translation[8] and RNA stability.[9] This minimal cap structure termed cap0 (m^7GpppN, where N is the TSS nucleotide) is found in lower eukaryotes like yeast. In higher organisms, the TSS nucleotide is 2′-*O*-methylated on the ribose by CMTR1 to create the cap1 structure ($m^7GpppNm$),[10] with most of m^7G-capped RNAs having this methylation.[11] Another mammalian ribose methylase CMTR2, modifies the second cap-proximal nucleotide to create the cap2 structure ($m^7GpppNmpNm$),[12] which is found on approximately 50% of polyadenylated RNAs in human cell cultures.[1,4,13] When the TSS nucleotide is an adenosine (which is usually an Am due to cap1 methylation), mammalian PCIF1 catalyzes base methylation (N^6-methyladenosine, m^6A) to create the m^6Am mark[11] (Figure 1A).

Cap-proximal ribose methylations prevent cellular RNAs from activating the innate immune response pathway.[14] Vertebrate cells have a system of cytosolic sensor proteins[15,16] that recognize molecular features on bacterial and viral nucleic acids to trigger production of secreted cytokines like interferons.[17] The interferons in turn initiate the production of a large set of interferon-stimulated genes (ISGs) in the infected cell and those nearby to create an anti-pathogenic environment.[18] Cap0 is one such "non-self" molecular feature that is recognized by the host innate immune sensors like RIG-I[19] and MDA5.[20] The presence of cap-proximal ribose methylations on RNA substrates reduces binding and, as a consequence, interferon pathway activation, as the methyl group clashes with a conserved histidine (H830) in RIG-I.[21–23] Cap-proximal ribose methylations also protect cellular mRNAs from the negative effects of ISGs like IFIT1 (interferon-induced protein with tetratricopeptide repeats 1), which is a potent translation repressor that binds cap0 RNAs.[24–26] First- or second-position methylation alone can individually reduce binding of IFIT1 to the capped RNA, but cap2 (both first- and second-position methylations) has the strongest inhibitory effect.[24] Consistently, knockdown of host *CMTR1* in mammalian cells triggers expression of ISGs, presumably because of the unmethylated cellular cap0 RNAs being sensed as "non-self."[21,27] Similarly, transfected RNAs with a second-position ribose methylation are identified as "self."[28] In line with this, *CMTR2* KO human HEK293T cells show mild activation of the interferon pathway in an RIG-I-dependent manner.[13] It should be noted that the capacity to sense "non-self" RNAs is not a universal attribute of all cell types.[17]

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).



(legend on next page)

Cap-proximal methylations also have a role in cellular gene expression. CMTR1 is nuclear and uses its WW domain to associate with the C-terminal domain (CTD) of RNA pol II.[29] In fact, CMTR1 is found on the TSS of most RNA pol II genes in mouse embryonic stem cells (mESCs) and is required for the transcription of ribosomal protein and histone genes.[30] Knockout of mouse *Cmtr1* leads to embryonic lethality, while conditional deletion in mouse brain affects dendritic morphogenesis.[31] Interestingly, fly Cmtr1 and Cmtr2 act redundantly as the cap1 methylase but act on a distinct set of transcripts, with the double mutant flies showing reward-learning defects.[32] Here, we sought to investigate the tissue-specific roles of CMTR1 and the physiological relevance of CMTR2 using mouse mutants.

RESULTS

Loss of mouse *Cmtr1* leads to embryonic arrest without activating the interferon pathway

To investigate the physiological importance of RNA cap1 methylation (Figure 1A), we examined a *Cmtr1* knockout mouse mutant (STAR Methods and Figure S1A and S1B). Heterozygous *Cmtr1*$^{+/-}$ animals of both sexes are viable and fertile. Crosses between heterozygous individuals resulted in litters that only had wild-type and heterozygous animals at weaning age (Figure 1B), indicating pre-weaning lethality, as previously reported for a gene-trap allele of *Cmtr1*.[31] We isolated embryos at different post-implantation stages and identified them by genotyping (Figure S1C). Homozygous *Cmtr1*$^{-/-}$ embryos (hereafter referred to as the *Cmtr1* mutant) were present at Mendelian ratios at embryonic day 6.5 (E6.5) and E7.5 (Figure 1C), but they were mostly not detected beyond E8.5, with many turning up dead (Figure S1C). The *Cmtr1* mutant embryos are morphologically indistinguishable from the control wild type at E6.5 but are dramatically reduced in size at E7.5, indicating arrested development (Figure 1D).

We sequenced transcriptomes from wild-type and *Cmtr1* mutant embryos at E6.5 and E7.5. Lack of CMTR1 activity is expected to result in host RNA pol II transcripts being unmethylated on the ribose of the first transcribed nucleotide, and such "non-self" cap0 RNAs should normally trigger the interferon pathway.[19–21] Strikingly, there was a complete absence of activation of the interferon pathway genes (Figure 1E and Table S3). Explaining the absence of the interferon pathway activation, a survey of embryonic transcriptomes shows that the different innate immune sensor genes are not expressed in the E6.5 and E7.5 embryos (Figures 1F and S1D). The lethality of *Cmtr1* mouse embryos shows that cap1 RNA methylation has an essential role in early mouse embryonic development, and this is unrelated to its function in preventing activation of the innate immune pathway.

snoRNA host genes are downregulated in the *Cmtr1* mutant embryos

Mouse gastrulation initiates with the formation of the primitive streak at E6.5, through which epiblast cells ingress before being allocated as precursors of the two primary germ layers: the mesoderm and the definitive endoderm.[33] We used the bulk sequencing data to project information on possible cell composition in the embryos (STAR Methods). This reveals a downregulation of gene expression representative of the primitive streak and the mesoderm in the mutant embryos and upregulation of endoderm genes (Figure S1E). Of the over 200 genes altered in the E6.5 mutant embryos, several downregulated genes are those annotated as small nucleolar RNA (snoRNA) host genes (*Snhg12, Snhg8, Snhg5 and Snhg4*) (Figure 1E). Another downregulated gene is the long noncoding RNA *Zfas1*,[34] which is a regulator of epithelial-mesenchymal transition and a snoRNA host gene (for *snord12*). Other downregulated genes include ribosomal protein 20 (Rsp20) and the translation factor eIF4A2, both of which contain intron-encoded snoRNAs. Interestingly, multiple gene copies of the U3 snoRNA (*Rnu3b1*, *Rnu3b3*) show the opposite trend by being upregulated. Unlike the intron-resident snoRNAs, the U3 snoRNA is encoded from an snRNA-type poll II gene that expresses an independent longer precursor form of the RNA.[35]

snoRNAs guide modifications of ribosomal RNAs (rRNAs)[36,37] and spliceosomal small nuclear RNAs (snRNAs).[38] Mammalian snoRNAs are mostly hosted within introns of protein-coding or long noncoding RNAs. Proper splicing and liberation of the introns for further processing of the snoRNA is critical for their biogenesis.[39,40] RNA splicing is influenced by the presence of the m^{7}G cap structure and the proteins that bind it.[41–43] However, quantification of intronic reads from snoRNA host genes did not reveal any striking differences in the *Cmtr1* mutant transcriptomes (Figure S1F). Analysis of global splicing events also did not reveal any dramatic changes in the mutant embryos

Figure 1. Embryonic lethality in the *Cmtr1* mutant mice and downregulation of snoRNA host genes

(A) Chemical structure of the 5′ N^7-methylated guanosine (m^{7}G) cap and cap-proximal ribose methylations. TSS, transcription start site nucleotide. Phenotypes of the mouse knockouts of the enzymes involved are indicated.

(B) Genotypes of animals (at the weaning age) born from heterozygous *Cmtr1* mouse crosses. Numbers of animals (also given as a percentage) are shown. WT, wild type; HET, heterozygous; KO, homozygous knockout. See also Figure S1C.

(C) Genotypes of mouse embryos (shown as percentage) at indicated embryonic days obtained from heterozygous *Cmtr1* mouse crosses.

(D) Cartoon showing mouse embryonic development with the stage at which *Cmtr1* knockout embryos arrest indicated by a red line. Representative wild-type and *Cmtr1* KO embryos at different stages are shown. Notice the degeneration of KO embryos from E8.0 onward.

(E) Volcano plots of differential gene expression between *Cmtr1* KO and wild-type in E6.5 and E7.5 mouse embryos. snoRNA host genes are highlighted in red. Absolute log2 fold change (log2FC) cutoff = 1, adjusted p value (padj) cutoff = 0.05.

(F) Boxplot showing expression of selected innate immune sensors from publicly available transcriptome datasets from different mouse embryonic stages. See also Figure S1D.

(G) Comparison of several alternative splicing (AS) events between wild-type and *Cmtr1* KO E6.5 and E7.5 embryos. Delta percent spliced-in (dPSI) score was computed as a difference between *Cmtr1* KO and wild-type PSI scores per AS event. AF, alternative first exon; SE, skipped exon; A3, alternative 3′ splice site; A5, alternative 5′ splice site; RI, retained intron; AL, alternative last exon; MX, mutual exclusion. Further details are in the STAR Methods. See also Figure S1F.

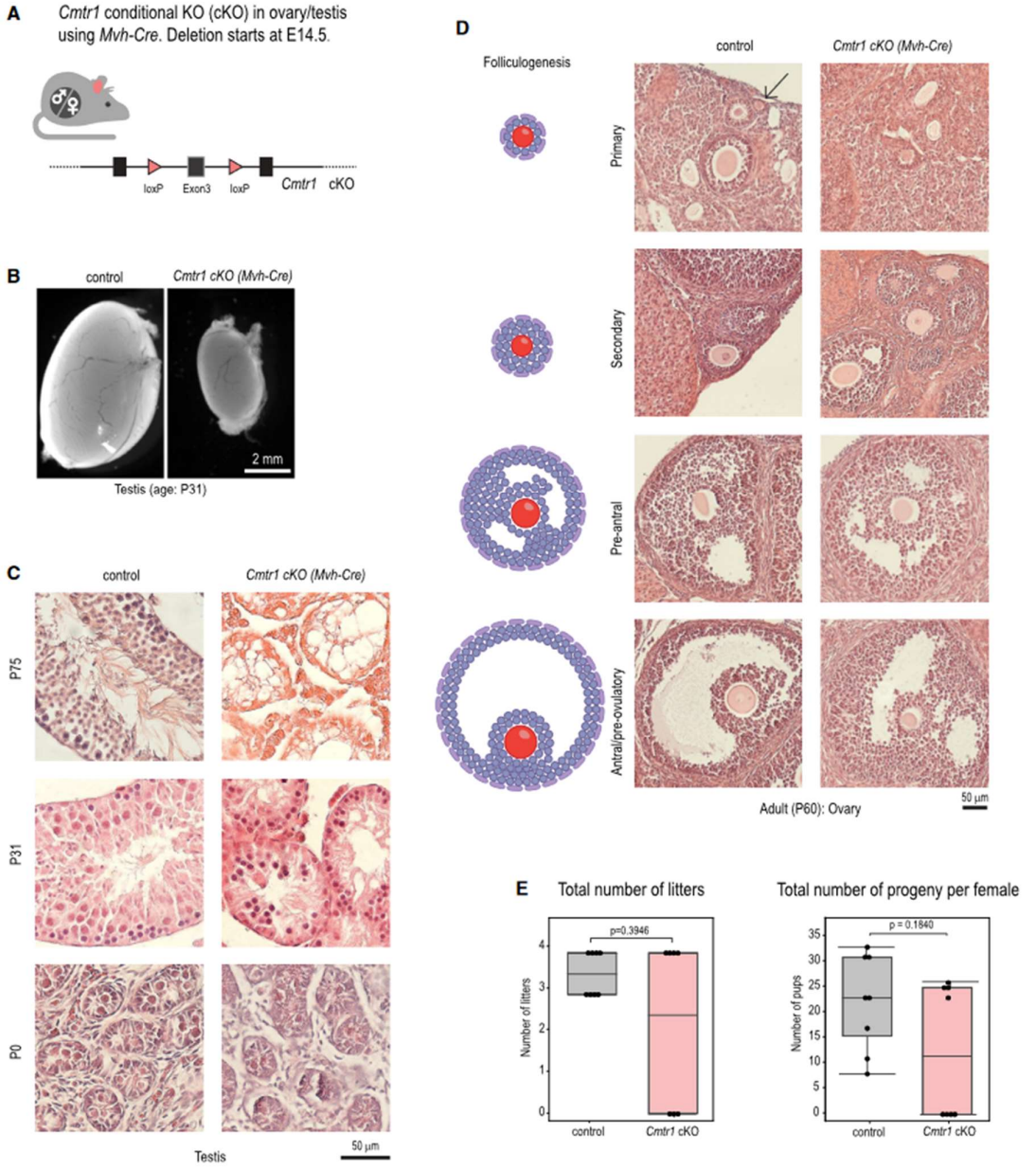


Figure 2. Mouse *Cmtr1* is essential for germline development and fertility

(A) Strategy for generating conditional knockout (cKO) deletion of *Cmtr1* in the mouse germline by deleting the coding exon3 using the *Mvh-Cre* line. See also Figure S2A.

(B) Atrophied mouse testes in the *Cmtr1* cKO ($Cmtr1^{-/loxP}$;Mvh-$Cre^{+/}$) vs. control ($Cmtr1^{+/loxP}$;Mvh-$Cre^{+/}$) mice. Scale bar for 2 mm is indicated.

(C) Histological analysis using hematoxylin and eosin staining of *Cmtr1* cKO and control mouse testes at different indicated postnatal stages. Scale bar for 50 μm is indicated.

(legend continued on next page)

(Figure 1G). In conclusion, the limited expression changes in E6.5 *Cmtr1* mutant embryos, including that of several snoRNA host genes, leads to large-scale changes in the E7.5 embryos causing the mid-gastrulation arrest.

Mouse CMTR1 is essential for fertility

To study the role of CMTR1 in the mouse germline, we obtained the conditional KO mice, where deletion in the germline is achieved by the germline-specific *Mvh-Cre* line (*Vasa-Cre*) (Figures 2A, S2A, and S2B and STAR Methods). Germline-specific expression of the *Mvh-Cre* starts from embryonic day E14.5, ultimately creating the *Cmtr1* conditional knockout mice ($Cmtr1^{loxP/-}$; *Mvh-Cre*, cKO). When *Cmtr1* cKO adult (>60 postnatal days, P60) males were crossed with wild-type females, no litters were obtained, indicating male infertility. Examination of testes from adult (P75) cKO males shows that they are highly atrophied when compared to those from control ($Cmtr1^{loxP/+}$; *Mvh-Cre*) littermates (Figure 2B). Histological examination shows that seminiferous tubules in the adult cKO testes are narrow and empty of all germ cells (Figure 2C). In contrast, tubules in the control testes are large and full of germ cells in all stages of development during spermatogenesis. Mitotic spermatogonia, meiotic spermatocytes, post-meiotic haploid round spermatids, elongate spermatids and sperm are all visible within the control seminiferous tubules (Figures 2C and S2C). To determine when the spermatogenic arrest manifests in the cKO males, we examined younger mice. Germ cells in the seminiferous tubules of the control and mutant neonates (P0) are comparable (Figure 2C). In P31 animals, the seminiferous tubules in the control testes are large and full of germ cells that have completed meiosis, while the tubules in the mutant testes are narrow and depleted of such germ cells (Figure 2C). We propose that germ cells in the *Cmtr1* cKO testes do not survive and are probably removed by apoptosis.

Mitotic oogonia in the embryonic female germline enter meiosis at E13.5, and the oocytes are in pachytene stage of prophase I of meiosis at birth,[44] which is immediately followed by the diplotene stage when the oocytes start to assemble a multi-cell layered follicle around them (Figure 2D). Folliculogenesis proceeds through multiple stages with primary follicles having a single layer of granulosa cells around the oocyte. Histological examination of adult ovaries from both the control and the cKO females reveals oocytes in all the different stages, including the large mature follicles (Figures 2D and S2D). When cKO females were crossed with wild-type partners, some produced litters, while others never had any progeny (Figure 2E). The overall number of pups tended to be lower with the cKO *Cmtr1* females. We note that conditional deletion of *Cmtr1* due to *Mvh*-driven Cre expression takes place only from E14.5, once meiosis is already initiated in the female germline at E13.5, explaining the low penetrance of the female infertility phenotype. Taken together, we conclude that CMTR1 is required for sustaining normal germline development and fertility in mice.

Loss of cap1 RNA methylation in mouse liver triggers the innate immune pathway

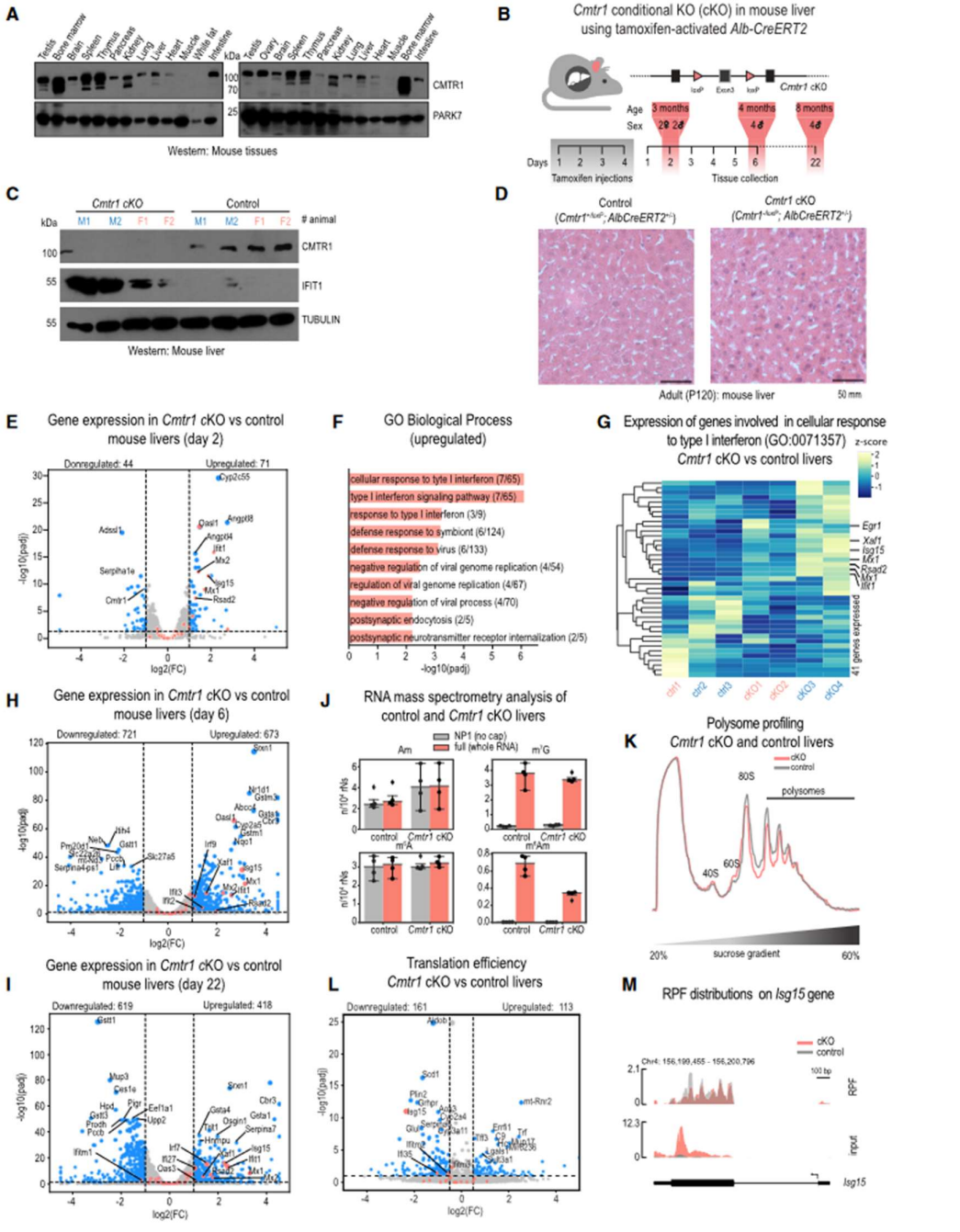
The above studies demonstrate that CMTR1 is important for developmental transitions, with its loss ultimately leading to death of the embryo or germ cells. CMTR1 is expressed in most mouse tissues (Figure 3A), and we decided to delete the *Cmtr1* in the adult mouse liver using tamoxifen-activated Cre-ERT2 that is expressed from the liver-specific albumin promoter (*Alb-CreERT2*) (Figures 3B, S3A, and S3B). Conditional knockout mice ($Cmtr1^{loxP/-}$;*Alb-CreERT2*, cKO) and control mice ($Cmtr1^{loxP/+}$;*Alb-CreERT2*) were both treated with intraperitoneal tamoxifen injections for a period of 4 days (days 1–4), and depletion of CMTR1 in the liver was monitored using western blot analysis 2 days later (Figure 3C). A complete absence of CMTR1 in the cKO mouse liver was observed (Figure 3C).

Histological analysis indicates an unchanged cellularity in the *Cmtr1* cKO liver (Figure 3D). Transcriptome analysis showed limited changes, with ~100 genes being altered in their expression (Figure 3E and Table S3). Strikingly, Gene Ontology analysis identifies an upregulation of transcripts involved in the anti-viral innate immune response pathway (Figure 3F). A number of ISGs are upregulated in the *Cmtr1* cKO liver (Figures 3E and 3G). Interferon mRNAs themselves are not detected, as they are usually expressed briefly before being turned off. We confirmed the induction of one such ISG by western blot analysis, which shows the specific expression of IFIT1 (interferon-induced protein with tetratricopeptide repeats 1) in all the *Cmtr1* cKO liver samples (Figure 3C). Analysis of the *Cmtr1* cKO mouse liver transcriptome at later time points of 6 days and 22 days post tamoxifen injection also revealed upregulation of ISG expression (Figures 3H and 3I). Thus, loss of *Cmtr1* in adult mouse liver can lead to chronic interferon pathway activation, which is accompanied by further alterations in metabolic pathway genes (Figure S3C). We note that the interferon-stimulated gene oligoadenylate synthetase like-1 (OASL1) that is induced in the mutant livers is reported as a negative regulator of interferon pathway, perhaps preventing tissue damage from chronic activation of the pathway.[45] RNA mass spectrometry analysis (STAR Methods) of polyA + RNAs from the *Cmtr1* cKO liver (sampled at day 2) confirms a sharp reduction in ribose methylation on the first nucleotide as represented by reduced levels of m^6Am (Figure 3J). m^6Am is a modification that depends on CMTR1-dependent TSS Am ribose methylation.[11,46] We propose that in the absence of CMTR1, the cellular RNA pol II ranscripts, which now have the cap0 structure, are sensed as "non-self," triggering interferon production[17] and expression of the ISGs.[47]

Loss of CMTR1 does not affect human cell viability and translation

To examine whether complete loss of CMTR1 is detrimental for cell survival, we examined a CRISPR-generated *CMTR1* knockout human HAP1 cell line (STAR Methods). HAP1 is a

(D) Histological analysis using hematoxylin and eosin staining of mouse ovaries from control and *Cmtr1* cKO is shown. The cartoon representation of the different stages of folliculogenesis is shown. Scale bar for 50 μm is indicated. Arrow points to a primary follicle.
(E) Fertility analysis of *Cmtr1* cKO females. Boxplots compare total number of litters (left) delivered and total number of progeny (right) per female. Mann-Whitney test was used to assess differences between wild-type and cKO animals. See also Figure S2D.



(legend on next page)

near-haploid cell line derived from the KBM-7 chronic myelogenous leukemia cell line. Complete loss of CMTR1 in these cells (Figure S4A) leads to complete loss of cap1 methylation (Figure S4B and STAR Methods). Although cell growth of the KO cell line was reduced (Figure S4C), this was not due to any impact on the cell cycle (Figure S4D). We note that reduced cell growth phenotype is observed in several unrelated mutant HAP1 cells that we have examined. We observed thousands of transcripts to be altered in the *CMTR1* KO cells but without activation of the interferon pathway (Figures S4E and S4H–S4I), probably because the innate immune sensors are not expressed in this cell line (Figure S4G).

We also examined translation status in the mutant cell line by ribosome profiling, but only a very few transcripts showed altered translation in the KO cell line (Figure S4F and Table S4). Similarly, sucrose-gradient centrifugation of *Cmtr1* cKO livers lysates did not reveal any changes in global translation as indicated by the largely identical patterns of the expected monosome and polysome peaks (Figures 3K and S3D). Ribosome profiling analysis showed that translation of only a few transcripts is altered (either increased or decreased) in the *Cmtr1* cKO liver (Figure 3L and Table S4). IFIT1 binds the cap0 structure to impair translation by competitively preventing recruitment of the translation initiation factor, cap-binding protein eIF4E.[24–26] However, we did not find any dramatic impact on translation *in vivo*. Consistent with a previous report,[48] we find that translation of *ISG15* is decreased in the absence of CMTR1 in mouse livers (Figures 3L–3M). Taken together, we show that loss of cap1 RNA methylation is not detrimental for translation of most mRNAs in the mouse liver and human HAP1 cells.

The cap2 methylase CMTR2 is essential for mouse embryonic development

Next, we obtained a mutant mouse where the entire CMTR2 coding sequence was deleted (Figures S5A and S5B and STAR Methods). Heterozygous $Cmtr2^{+/-}$ mice are viable and fertile. Intercrosses between heterozygous partners did not produce any homozygous $Cmtr2^{-/-}$ knockout (KO) mice (hereafter referred to as *Cmtr2* mutant) in the born litter, indicating pre-weaning lethality (Figures 4A and 4B). We set up crosses between heterozygous partners and genotyped embryos at different post-implantation stages (Figure S5C). This indicated that *Cmtr2* mutant embryos were present at above Mendelian ratios at E6.5 and E7.5 (Figure 4C). Mutant *Cmtr2* embryos recovered beyond E6.5 were infrequent, reduced in size, and found arrested at a preceding developmental stage (Figure S5C).

To examine the impact on gene expression in the *Cmtr2* KO, we chose E6.5 embryos that appeared morphologically similar to the control wild type. We also sequenced embryos isolated 1 day later at E7.5, where the *Cmtr2* mutant embryos are smaller than the control (Figure 4D). Sequence analysis indicates that CMTR2 regulates levels of hundreds of transcripts, most of which are upregulated in the mutant (Figure 4E and Table S3). Gene Ontology analysis revealed that genes active in many different pathways are impacted (Figures S5D–S5G). Similarly, hundreds of genes are altered in the *Cmtr2* mutant E7.5 embryos (Figure 4E). Estimation of cell composition changes from the bulk sequencing data shows that critical cell types required for normal embryonic development are affected in the *Cmtr2* mutant (Figure S5H and STAR Methods). Finally, comparison of the changes in the *Cmtr1* and *Cmtr2* mutant embryos shows that the two proteins regulate non-overlapping sets of genes (Figure 4F). In conclusion, we find that loss of the cap2 RNA methylase *Cmtr2* results in post-implantation lethality in mice, with most embryos arrested during mid-gastrulation at E7.5.

DISCUSSION

Here we showed that both *Cmtr1* and *Cmtr2* are essential for mouse embryonic development. While a role for cap1[21–23] and cap2[13,28] modifications in blocking activation of the innate immune response pathway is established, transcriptome analysis of the mutant embryos did not reveal signatures of interferon signaling. We rationalize that this is due to absence of expression of innate immune sensors at the early developmental time points (Figure 1F). Indeed, the mid-gastrulation arrest (at E7.5) of *Cmtr1*

Figure 3. Loss of *Cmtr1* in mouse liver activates the innate immune response pathway

(A) Western analysis of CMTR1 in adult (P60) mouse tissues.

(B) Strategy for generating conditional knockout (cKO) deletion of *Cmtr1* in the mouse liver using the tamoxifen activable CreERT2. Mice are given daily injections for 4 days and analyzed 2, 6, and 22 days later.

(C) Western analysis with four biological replicates of liver lysates each from control ($Cmtr1^{+/loxP}$; AlbCreERT2$^{+/-}$) and *Cmtr1* cKO ($Cmtr1^{-/loxP}$; AlbCreERT2$^{+/-}$) adult mice. TUBULIN is used as loading control. Sex of donor animal is indicated: M, male (blue); F, female (red).

(D) Histological analysis using hematoxylin and eosin staining of mouse liver from control and *Cmtr1* cKO.

(E) Volcano plot of differential gene expression in liver using transcriptome sequencing of three to four biological replicates of *Cmtr1* cKO and control adult mice. Type I IFN pathway genes are highlighted in red. Absolute log2 fold change (log2FC) cutoff = 1; adjusted p value (padj) cutoff = 0.05.

(F) Gene ontology analysis of the genes upregulated in the *Cmtr1* cKO adult mouse liver compared to the control.

(G) Heatmap showing expression changes of the 41 genes involved in cellular response to the type I interferon (GO:0071357) in the control and *Cmtr1* cKO mouse liver. Data for the three to four biological replicates are shown. Males (blue) and females (red) are shown in different colors.

(H and I) Volcano plot of differential gene expression between *Cmtr1* cKO and control adult mice from day 6 or day 22 post injection. Type I IFN pathway genes are highlighted in red. Absolute log2 fold change (log2FC) cutoff = 1; adjusted p value (padj) cutoff = 0.05. See also Figure S3C.

(J) PolyA+ RNA from *Cmtr1* control and cKO livers was subjected to mass spectrometry. Bar plots combined with boxplots show abundance (number of modified nucleotides/10^4 nucleotides) of some of the modifications (Am, m^7G, m^6A, m^6Am). Quadruplicate biological replicates were tested.

(K) Sucrose density gradient (linear 20%–60%) analysis of liver lysates from control and the *Cmtr1* cKO adult mouse. Positions of 40S and 60S subunits, 80S monosomes, and the polysome peaks are indicated. See also Figure S3D for second replicate.

(L) The volcano plot of differential translation efficiency of liver transcripts between *Cmtr1* cKO and control mice. *Isg15* is highlighted in red. Absolute log2 fold change (log2FC) cutoff = 0.5, and adjusted p value (padj) cutoff = 0.05.

(M) Normalized coverage tracks of *Isg15* counts (rpm) from the merged ribosome-protected fragment (RPF) and input samples. See also Figure S3E.

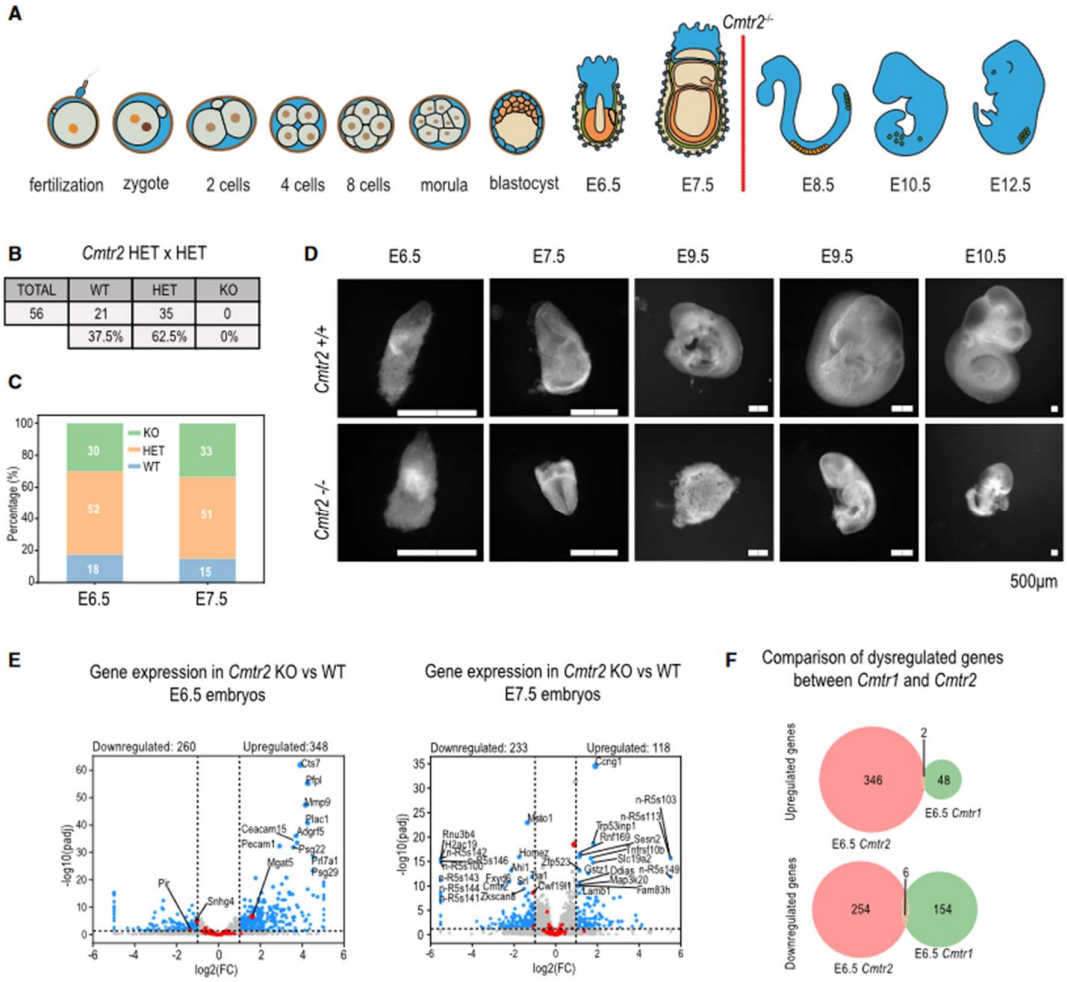


Figure 4. *Cmtr2* is essential for mouse embryonic development

(A) Scheme of embryonic development with lethality in *Cmtr2* mutants indicated by a red line.

(B) Genotypes of animals (at weaning age) born out of heterozygous *Cmtr2* mouse crosses. WT, wild type; HET, heterozygous; KO, homozygous knockout. See also Figure S5C.

(C) Genotypes of mouse embryos (shown as percentage) at indicated embryonic days.

(D) Pictures of representative wild-type and *Cmtr2* KO embryos at different stages are shown. The mutant embryos start to disintegrate already from E7.5 onward.

(E) Volcano plots showing differentially expressed genes between *Cmtr2* KO and wild-type E6.5 and E7.5 mouse embryos. snoRNA host genes are highlighted in red. Absolute log2 fold change (log2FC) cutoff = 1; adjusted p value (padj) cutoff = 0.05.

(F) Venn diagrams showing lack of overlap in gene expression changes in the *Cmtr1* and *Cmtr2* KO mouse embryos compared to their wild-type controls.

and *Cmtr2* mutant mouse embryos contrasts with the late arrest seen in the mutant embryos for *Adar1* (adenosine deaminase acting on RNA).[49,50] *Adar1* mutant embryos die due to activation of the interferon pathway in an MDA5-dependent manner only at E12.5 to E13.5, a time window when the innate immune receptors start to be expressed (Figures 1F and S1D). This suggests that cap1 and cap2 modifications have an essential role in shaping gene expression early during mammalian embryonic development that is unrelated to their role in marking host RNAs as "self."

Transcriptome analysis shows that the two proteins regulate non-overlapping sets of genes (Figure 4F). This points to distinct RNA targets, which is consistent with the distinct sub-cellular compartments where the two enzymes act.

CMTR1 can co-transcriptionally engage the nascent RNA via interaction of its WW domain with the CTD of RNA pol II.[29,30] Methylation by CMTR2 on the other hand is reported to be cytosolic.[51] How ribose methylations affect target gene expression is not clear. Perhaps proteins that bind the cap structure and sense the proximal nucleotides are influenced by the methylations. However, binding affinity of the recombinant human nuclear Cap-binding complex (CBC) for the m^7G cap analog (m^7GpppN vs. $m^7GpppNm$) is not influenced by the cap1 modification.[52] Likewise, affinity of the recombinant mouse translation initiation factor eIF4E for the cap structure is not affected by presence of the cap1 or cap2.[28] Nevertheless, analysis using fly nuclear extracts shows an increased association between the fly CBC component CBC80 and cap2 RNA, when compared to the cap0 RNA control.[32] It is also known that cap-proximal ribose methylations can influence cap-binding of factors involved in the innate immune response pathway.[21,24–26] Furthermore, protein factors that interact with the cap methylases may modulate their activity or targeting. CMTR1 is shown to interact with the RNA helicase DHX15 (Figure S4J), with the interaction reducing RNA methylation activity of CMTR1[53] or promoting its activity on RNAs with secondary structures.[54]

In the case of *Cmtr1* mutant embryos, we reported downregulation of several snoRNA host genes. We speculate that presence of cap1 modification might facilitate coordinated processing of the intron-resident snoRNAs and splicing of the intron. It also remains an exciting possibility that the unavailability of specific snoRNAs might underlie embryonic arrest phenotype in the *Cmtr1* mutant. However, this impact on snoRNA host genes was not seen in the *Cmtr1* mutant liver (Figure S3G) and human *CMTR1* KO HAP1 cells (Figure S4F). While we refer to genes being regulated as those that are increased or decreased in levels in the mutant environment, we are unable to precisely determine the molecular reason for these changes. There are different possibilities, as CMTR1 is proposed to promote transcription,[30] RNA splicing,[55] stability,[9] or association with cap-binding factors to promote RNA localization.[32] Our examination of translation in the *Cmtr1* mutant liver (Figures 3K and S3F) and human knockout cell line (Figure S4F) shows that global translation is unaffected.

Our *Cmtr2* mutant embryo analysis shows that the misregulated genes are distinct from those affected in the *Cmtr1* mutant (Figure 4F). Precise mapping of cap2 on transcripts expressed in human HEK293T cells shows that transcripts tend to accumulate the modification without any particular sequence.[13] The level of cap2 is also variable in different tissue and cell types. The same study concluded that translation and RNA stability were not influenced by the presence of the cap2 modification,[13] while another study showed that presence of cap2 reduces translation of reporter mRNAs in some cell types.[28] Cap2-modified RNA was shown to resist recognition by the innate immune sensor RIG-I, thereby preventing activation of the interferon signaling pathway[13] or by reducing binding of the interferon-stimulated gene IFIT1, which acts as a translation repressor.[24] In conclusion, our study sheds light on a role for cap1 and cap2 methylation in gene regulation, beyond their role in marking cellular RNAs as "self."

Limitations of the study

While we described the transcriptome changes in the mutant embryos, we are unable to say if these are direct effects due to loss of the respective proteins and due to loss of position-specific ribose methylations on these RNAs. While we documented reduced transcript levels of some of the snoRNA host genes in the *Cmtr1* mutant embryos, we could not evaluate if this affected the levels of the intron-encoded snoRNAs. We showed that CMTR1 is essential for germline development as mutant males are completely infertile, while females display infertility at low penetrance. However, we do not know if this is due to a direct role in germline gene expression or due to activation of the interferon pathway, as observed in the conditional *Cmtr1* mutant livers (Figure 2E). Finally, the molecular role of cap2 methylation in gene regulation and how CMTR2 selects its target RNAs remain open questions.

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 STUDY PARTICIPANT DETAILS
  - Animal work
  - Mouse mutants
  - Human HAP1 *CMTR1* KO cells
- METHOD DETAILS
  - *Cmtr1* knockout mouse
  - *Cmtr2* knockout mouse
  - Creation of *Cmtr1 loxP* mice
  - Tamoxifen-inducible conditional *Cmtr1* deletion in mouse liver
  - Conditional *Cmtr1* germline knockout mice
  - Mouse genotyping
  - Collection of mouse embryos
  - Human HAP1 *CMTR1* KO cells
  - Clones and constructs
  - Recombinant protein production
  - Purification of mouse CMTR2 protein
  - Purification of human CMTR1-DHX15 complex
  - Antibodies
  - Collection of RNA from HAP1 cells
  - RNA extraction from cell lines
  - Growth curve
  - Analysis of cell cycle by FACS
  - Collecting mouse tissues for western blot
  - Western Blot
  - Histological analysis of mouse tissues
  - Total RNA purification from mouse liver samples
  - PolyA+ RNA purification
  - Quantification of RNA modifications using LC-MS/MS
  - RNA library preparation and sequencing

- Polysome and ribosome profiling
- Polysome gradient centrifugation
- Ribosome profiling
- Ribosome protected fragment (RPF) library preparation
- Isolation of total RNA, library preparation and sequencing

- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Analysis of RNA-seq data
  - Analysis of ribosome profiling data
  - Analysis of alternative splicing events
  - Bulk-to-single-cell deconvolution (RNA-seq debulking)

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2023.112786.

ACKNOWLEDGMENTS

We thank Nicolas Roggli for scientific illustration and Pascal Gos for help with mouse work. We thank the Mouse Knockout Program and the Mutant Mouse Resource and Research Center (MMRRC), University of California at Davis for mouse mutants. We thank the following University of Geneva core facilities: iGE3 Genomics Platform, Transgenic Mouse Facility, Histology Facility. We also thank the EMBL Genomics core facility for deep sequencing. We thank fellowships to the students: M.D. (Swiss Government Excellence PhD Scholarship), K.K. (NCCR RNA&Disease network), M.M. (Boehringer Ingelheim Fonds PhD Fellowship), and L.L. (iGE3 PhD Fellowship). This work was supported by grants to R.S.P. from the Swiss National Science Foundation: Project Grant (#310030_207468), Sinergia Grant (#CRSII5_186266) and funding from the NCCR RNA & Disease (#51NF40-205601). Work in the Pillai lab is supported by the Republic and Canton of Geneva.

AUTHOR CONTRIBUTIONS

M.D. conducted a majority of the experiments including all the mouse experiments and cell culture work; K.K. did all computational analyses and prepared all the figures with M.D.; M.M. prepared the conditional *Cmtr1* KO mouse; L.L. produced recombinant proteins; O.P. did ribosome and polysome profiling; F.F.-O. prepared RNAs from mouse tissues; C.B.V. did RNA mass spectrometry analyses; D.H. helped with computational analyses at an early stage of this study; manuscript preparation was by R.S.P. with input from everyone.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 1, 2023
Revised: May 11, 2023
Accepted: June 25, 2023

REFERENCES

1. Wei, C.M., Gershowitz, A., and Moss, B. (1975). Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. Cell *4*, 379–386. https://doi.org/10.1016/0092-8674(75)90158-0.
2. Furuichi, Y., Morgan, M., Muthukrishnan, S., and Shatkin, A.J. (1975). Reovirus messenger RNA contains a methylated, blocked 5'-terminal structure: m-7G(5')ppp(5')G-MpCp. Proc. Natl. Acad. Sci. USA *72*, 362–366. https://doi.org/10.1073/pnas.72.1.362.
3. Furuichi, Y., Muthukrishnan, S., and Shatkin, A.J. (1975). 5'-Terminal m-7G(5')ppp(5')G-m-p in vivo: identification in reovirus genome RNA. Proc. Natl. Acad. Sci. USA *72*, 742–745. https://doi.org/10.1073/pnas.72.2.742.
4. Furuichi, Y., Morgan, M., Shatkin, A.J., Jelinek, W., Salditt-Georgieff, M., and Darnell, J.E. (1975). Methylated, blocked 5 termini in HeLa cell mRNA. Proc. Natl. Acad. Sci. USA *72*, 1904–1908.
5. Wei, C.M., and Moss, B. (1975). Methylated nucleotides block 5'-terminus of vaccinia virus messenger RNA. Proc. Natl. Acad. Sci. USA *72*, 318–322.
6. Perry, R.P., and Kelley, D.E. (1975). Methylated constituents of heterogeneous nuclear RNA: presence in blocked 5' terminal structures. Cell *6*, 13–19. https://doi.org/10.1016/0092-8674(75)90068-9.
7. Furuichi, Y., and Shatkin, A.J. (2000). Viral and cellular mRNA capping: past and prospects. Adv. Virus Res. *55*, 135–184.
8. Muthukrishnan, S., Filipowicz, W., Sierra, J.M., Both, G.W., Shatkin, A.J., and Ochoa, S. (1975). mRNA methylation and protein synthesis in extracts from embryos of brine shrimp, Artemia salina. J. Biol. Chem. *250*, 9336–9341.
9. Picard-Jean, F., Brand, C., Tremblay-Létourneau, M., Allaire, A., Beaudoin, M.C., Boudreault, S., Duval, C., Rainville-Sirois, J., Robert, F., Pelletier, J., et al. (2018). 2'-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. PLoS One *13*, e0193804. https://doi.org/10.1371/journal.pone.0193804.
10. Bélanger, F., Stepinski, J., Darzynkiewicz, E., and Pelletier, J. (2010). Characterization of hMTr1, a human Cap1 2'-O-ribose methyltransferase. J. Biol. Chem. *285*, 33037–33044. https://doi.org/10.1074/jbc.M110.155283.
11. 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 *363*, eaav0080. https://doi.org/10.1126/science.aav0080.
12. 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. https://doi.org/10.1093/nar/gkr038.
13. Despic, V., and Jaffrey, S.R. (2023). mRNA ageing shapes the Cap2 methylome in mammalian mRNA. Nature *614*, 358–366.
14. Habjan, M., and Pichlmair, A. (2015). Cytoplasmic sensing of viral nucleic acids. Curr. Opin. Virol. *11*, 31–37.
15. Kowalinski, E., Lunardi, T., McCarthy, A.A., Louber, J., Brunel, J., Grigorov, B., Gerlier, D., and Cusack, S. (2011). Structural basis for the activation of innate immune pattern-recognition receptor RIG-I by viral RNA. Cell *147*, 423–435.
16. Luo, D., Ding, S.C., Vela, A., Kohlway, A., Lindenbach, B.D., and Pyle, A.M. (2011). Structural insights into RNA recognition by RIG-I. Cell *147*, 409–422.
17. Ivashkiv, L.B., and Donlin, L.T. (2014). Regulation of type I interferon responses. Nat. Rev. Immunol. *14*, 36–49.
18. Schoggins, J.W., Wilson, S.J., Panis, M., Murphy, M.Y., Jones, C.T., Bieniasz, P., and Rice, C.M. (2011). A diverse range of gene products are effectors of the type I interferon antiviral response. Nature *472*, 481–485.
19. Daffis, S., Szretter, K.J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.Y., Schneller, S., Zust, R., Dong, H., et al. (2010). 2'-O methylation of the viral mRNA cap evades host restriction by IFIT family members. Nature *468*, 452–456. https://doi.org/10.1038/nature09489.
20. Züst, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B.W., Ziebuhr, J., Szretter, K.J., Baker, S.C., Barchet, W., Diamond, M.S., et al. (2011). Ribose 2'-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. Nat. Immunol. *12*, 137–143. https://doi.org/10.1038/ni.1979.
21. Schuberth-Wagner, C., Ludwig, J., Bruder, A.K., Herzner, A.M., Zillinger, T., Goldeck, M., Schmidt, T., Schmid-Burgk, J.L., Kerber, R., Wolter, S.,

et al. (2015). A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2'O-Methylated Self RNA. Immunity *43*, 41–51. https://doi.org/10.1016/j.immuni.2015.06.015.

22. Wang, Y., Ludwig, J., Schuberth, C., Goldeck, M., Schlee, M., Li, H., Juranek, S., Sheng, G., Micura, R., Tuschl, T., et al. (2010). Structural and functional insights into 5′-ppp RNA pattern recognition by the innate immune receptor RIG-I. Nat. Struct. Mol. Biol. *17*, 781–787.

23. 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. USA *113*, 596–601.

24. Abbas, Y.M., Laudenbach, B.T., Martínez-Montero, S., Cencic, R., Habjan, M., Pichlmair, A., Damha, M.J., Pelletier, J., and Nagar, B. (2017). Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and N2 ribose 2′-O methylations. Proc. Natl. Acad. Sci. USA *114*, E2106–E2115. https://doi.org/10.1073/pnas.1612444114.

25. Habjan, M., Hubel, P., Lacerda, L., Benda, C., Holze, C., Eberl, C.H., Mann, A., Kindler, E., Gil-Cruz, C., Ziebuhr, J., et al. (2013). Sequestration by IFIT1 impairs translation of 2′ O-unmethylated capped RNA. PLoS Pathog. *9*, e1003663.

26. Fleith, R.C., Mears, H.V., Leong, X.Y., Sanford, T.J., Emmott, E., Graham, S.C., Mansur, D.S., and Sweeney, T.R. (2018). IFIT3 and IFIT2/3 promote IFIT1-mediated translation inhibition by enhancing binding to non-self RNA. Nucleic Acids Res. *46*, 5269–5285.

27. Li, B., Clohisey, S.M., Chia, B.S., Wang, B., Cui, A., Eisenhaure, T., Schweitzer, L.D., Hoover, P., Parkinson, N.J., Nachshon, A., et al. (2020). Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. Nat. Commun. *11*, 164. https://doi.org/10.1038/s41467-019-13965-x.

28. Drazkowska, K., Tomecki, R., Warminski, M., Baran, N., Cysewski, D., Depaix, A., Kasprzyk, R., Kowalska, J., Jemielity, J., and Sikorski, P.J. (2022). 2′-O-Methylation of the second transcribed nucleotide within the mRNA 5′ cap impacts the protein production level in a cell-specific manner and contributes to RNA immune evasion. Nucleic Acids Res. *50*, 9051–9071.

29. Haline-Vaz, T., Silva, T.C.L., and Zanchin, N.I.T. (2008). The human interferon-regulated ISG95 protein interacts with RNA polymerase II and shows methyltransferase activity. Biochem. Biophys. Res. Commun. *372*, 719–724. https://doi.org/10.1016/j.bbrc.2008.05.137.

30. Liang, S., Silva, J.C., Suska, O., Lukoszek, R., Almohammed, R., and Cowling, V.H. (2022). CMTR1 is recruited to transcription start sites and promotes ribosomal protein and histone gene expression in embryonic stem cells. Nucleic Acids Res. *50*, 2905–2922. https://doi.org/10.1093/nar/gkac122.

31. Lee, Y.-L., Kung, F.-C., Lin, C.-H., and Huang, Y.-S. (2020). CMTR1-Catalyzed 2′-O-Ribose Methylation Controls Neuronal Development by Regulating Camk2α Expression Independent of RIG-I Signaling. Cell Rep. *33*, 108269.

32. Haussmann, I.U., Wu, Y., Nallasivan, M.P., Archer, N., Bodi, Z., Hebenstreit, D., Waddell, S., Fray, R., and Soller, M. (2022). CMTr cap-adjacent 2′-O-ribose mRNA methyltransferases are required for reward learning and mRNA localization to synapses. Nat. Commun. *13*, 1209–1213.

33. Tam, P.P.L., and Loebel, D.A.F. (2007). Gene function in mouse embryogenesis: get set for gastrulation. Nat. Rev. Genet. *8*, 368–381.

34. O'Brien, S.J., Fiechter, C., Burton, J., Hallion, J., Paas, M., Patel, A., Patel, A., Rochet, A., Scheurlen, K., Gardner, S., et al. (2021). Long noncoding RNA ZFAS1 is a major regulator of epithelial-mesenchymal transition through miR-200/ZEB1/E-cadherin, vimentin signaling in colon adenocarcinoma. Cell Death Discov. *7*, 61. https://doi.org/10.1038/s41420-021-00427-x.

35. Kiss, T. (2004). Biogenesis of small nuclear RNPs. J. Cell Sci. *117*, 5949–5951.

36. Decatur, W.A., and Fournier, M.J. (2002). rRNA modifications and ribosome function. Trends Biochem. Sci. *27*, 344–351. https://doi.org/10.1016/s0968-0004(02)02109-6.

37. Terns, M.P., and Terns, R.M. (2002). Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin. Gene Expr. *10*, 17–39.

38. Tycowski, K.T., Shu, M.D., Kukoyi, A., and Steitz, J.A. (2009). A conserved WD40 protein binds the Cajal body localization signal of scaRNP particles. Mol. Cell *34*, 47–57. https://doi.org/10.1016/j.molcel.2009.02.020.

39. Hirose, T., and Steitz, J.A. (2001). Position within the host intron is critical for efficient processing of box C/D snoRNAs in mammalian cells. Proc. Natl. Acad. Sci. USA *98*, 12914–12919.

40. Hirose, T., Shu, M.-D., and Steitz, J.A. (2003). Splicing-dependent and-independent modes of assembly for intron-encoded box C/D snoRNPs in mammalian cells. Mol. Cell *12*, 113–123.

41. Edery, I., and Sonenberg, N. (1985). Cap-dependent RNA splicing in a HeLa nuclear extract. Proc. Natl. Acad. Sci. USA *82*, 7590–7594.

42. Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I.W. (1994). A nuclear cap binding protein complex involved in pre-mRNA splicing. Cell *78*, 657–668.

43. 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.

44. Li, R., and Albertini, D.F. (2013). The road to maturation: somatic cell interaction and self-organization of the mammalian oocyte. Nat. Rev. Mol. Cell Biol. *14*, 141–152.

45. Lee, M.S., Kim, B., Oh, G.T., and Kim, Y.-J. (2013). OASL1 inhibits translation of the type I interferon–regulating transcription factor IRF7. Nat. Immunol. *14*, 346–355.

46. Boulias, K., Toczydłowska-Socha, D., Hawley, B.R., Liberman, N., Takashima, K., Zaccara, S., Guez, T., Vasseur, J.J., Debart, F., Aravind, L., et al. (2019). Identification of the m(6)Am Methyltransferase PCIF1 Reveals the Location and Functions of m(6)Am in the Transcriptome. Mol. Cell *75*, 631–643.e8. https://doi.org/10.1016/j.molcel.2019.06.006.

47. Schneider, W.M., Chevillotte, M.D., and Rice, C.M. (2014). Interferon-stimulated genes: a complex web of host defenses. Annu. Rev. Immunol. *32*, 513–545.

48. Williams, G.D., Gokhale, N.S., Snider, D.L., and Horner, S.M. (2020). The mRNA Cap 2′-O-Methyltransferase CMTR1 Regulates the Expression of Certain Interferon-Stimulated Genes. mSphere *5*, e00202-20–e00220.

49. Liddicoat, B.J., Piskol, R., Chalk, A.M., Ramaswami, G., Higuchi, M., Hartner, J.C., Li, J.B., Seeburg, P.H., and Walkley, C.R. (2015). RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. Science *349*, 1115–1120.

50. Mannion, N.M., Greenwood, S.M., Young, R., Cox, S., Brindle, J., Read, D., Nellåker, C., Vesely, C., Ponting, C.P., McLaughlin, P.J., et al. (2014). The RNA-editing enzyme ADAR1 controls innate immune responses to RNA. Cell Rep. *9*, 1482–1494.

51. Langberg, S.R., and Moss, B. (1981). Post-transcriptional modifications of mRNA. Purification and characterization of cap I and cap II RNA (nucleoside-2′-)-methyltransferases from HeLa cells. J. Biol. Chem. *256*, 10054–10060.

52. 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.

53. Inesta-Vaquera, F., Chaugule, V.K., Galloway, A., Chandler, L., Rojas-Fernandez, A., Weidlich, S., Peggie, M., and Cowling, V.H. (2018). DHX15 regulates CMTR1-dependent gene expression and cell proliferation. Life Sci. Alliance *1*, e201800092. https://doi.org/10.26508/lsa.201800092.

54. Toczydlowska-Socha, D., Zielinska, M.M., Kurkowska, M., Astha, Almeida, C.F., Almeida, C.F., Stefaniak, F., Purta, E., and Bujnicki, J.M. (2018). Human RNA cap1 methyltransferase CMTr1 cooperates with RNA helicase DHX15 to modify RNAs with highly structured 5′ termini. Philos. Trans. R. Soc. Lond. B Biol. Sci. *373*, 20180161.

55. Dönmez, G., Hartmuth, K., and Lührmann, R. (2004). Modified nucleotides at the 5' end of human U2 snRNA are required for spliceosomal E-complex formation. RNA *10*, 1925–1933. https://doi.org/10.1261/rna.7186504.

56. Schuler, M., Dierich, A., Chambon, P., and Metzger, D. (2004). Efficient temporally controlled targeted somatic mutagenesis in hepatocytes of the mouse. Genesis *39*, 167–172.

57. 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 Bioinf. *14*, 128. https://doi.org/10.1186/1471-2105-14-128.

58. 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*, W90–W97. https://doi.org/10.1093/nar/gkw377.

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

60. 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. https://doi.org/10.1186/s13059-014-0550-8.

61. 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. https://doi.org/10.1038/nmeth.3252.

62. 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. https://doi.org/10.1038/nmeth.4197.

63. Andrews, S. (2012). FastQC: A Quality Control Application for High Throughput Sequence Data. Babraham Institute Project page. http://www. bioinformatics. bbsrc. ac. uk/projects/fastqc.

64. Dodt, M., Roehr, J.T., Ahmed, R., and Dieterich, C. (2012). FLEXBAR—flexible barcode and adapter processing for next-generation sequencing platforms. Biology *1*, 895–905.

65. 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.

66. Waskom, M. (2021). Seaborn: statistical data visualization. J. Open Source Softw. *6*, 3021.

67. Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. Comput. Sci. Eng. *9*, 90–95.

68. Fang, Z., Liu, X., and Peltz, G. (2023). GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. Bioinformatics *39*, btac757.

69. 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. https://doi.org/10.1186/gb-2009-10-3-r25.

70. Toolkit, P. (2019). Broad institute, GitHub repository.

71. Xiao, Z., Huang, R., Xing, X., Chen, Y., Deng, H., and Yang, X. (2018). De novo annotation and characterization of the translatome with ribosome profiling data. Nucleic Acids Res. *46*, e61.

72. Kurtenbach, S., and Harbour, J.W. (2019). SparK: A Publication-Quality NGS Visualization Tool. bioRxiv845529, . Preprint at. https://doi.org/10.1101/845529.

73. Trincado, J.L., Entizne, J.C., Hysenaj, G., Singh, B., Skalic, M., Elliott, D.J., and Eyras, E. (2018). SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. Genome Biol. *19*. 40-11.

74. Dong, M., Thennavan, A., Urrutia, E., Li, Y., Perou, C.M., Zou, F., and Jiang, Y. (2021). SCDC: bulk gene expression deconvolution by multiple single-cell RNA sequencing references. Brief. Bioinform. *22*, 416–427.

75. Leek, J.T., Johnson, W.E., Parker, H.S., Fertig, E.J., Jaffe, A.E., Storey, J.D., Zhang, Y., and Torres, L.C. (2014). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics *28*, 882–883. https://doi.org/10.1093/bioinformatics/bts034.

76. Dunn, J.G., and Weissman, J.S. (2016). Plastid: nucleotide-resolution analysis of next-generation sequencing and genomics data. BMC Genom. *17*, 958–1012.

77. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., and Dubourg, V. (2011). Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. *12*, 2825–2830.

78. Gallardo, T., Shirley, L., John, G.B., and Castrillon, D.H. (2007). Generation of a germ cell-specific mouse transgenic Cre line. Genesis *45*, 413–417.

79. Ingolia, N.T., Ghaemmaghami, S., Newman, J.R.S., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. Science *324*, 218–223. https://doi.org/10.1126/science.1168978.

80. McGlincy, N.J., and Ingolia, N.T. (2017). Transcriptome-wide measurement of translation by ribosome profiling. Methods *126*, 112–129. https://doi.org/10.1016/j.ymeth.2017.05.028.

81. Ernst, C., Eling, N., Martinez-Jimenez, C.P., Marioni, J.C., and Odom, D.T. (2019). Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. Nat. Commun. *10*, 1251–1320.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
anti-CMTR1	Atlas antibodies	Cat. No. HPA029980
anti-IFIT1	Cell signaling	Cat. No. D2X9Z
anti-TUBULIN	Abcam	Cat. No. ab6046
anti-PARK7	Abcam	Cat. No. ab18257
Rabbit anti-Mouse IgG (H + L) Superclonal™ Secondary Antibody, HRP conjugate	Invitrogen	Cat. No. A27025
Amersham ECL Rabbit IgG, HRP-linked whole Ab (from donkey)	GE Healthcare	Cat. No. NA934-1ML
Bacterial and virus strains		
Top10	N/A	N/A
Biological samples		
pACEBac2SS-hDHX15	This study	6xHis-Strep-SUMO-TEV- tags on protein
pIDK vector-hCMTR1	This study	https://www.snapgene.com/plasmids/insect_cell_vectors/pIDK
pACEBac2SS-hDHX15-hCMTR1	This study	Tagged DHX15 and untagged CMTR1
pACEBac2SS-hCMTR2(15–759 aa)	This study	6xHis-Strep-SUMO-TEV- tags on protein
Chemicals, peptides, and recombinant proteins		
Sodium deoxycholate	Sigma	Cat. No. 30968
Complete EDTA-free protease inhibitor	Roche	Cat. No. 11 873 580 001
TRIzol™ Reagent	Invitrogen	Cat. No. 15596-026
Trypan blue	Sigma	Cat. No. 93595-50ML
Propidium Iodide	Sigma	Cat. No. P4170
Ponceau S	Sigma	P3504
IMDM Medium	Gibco	Cat. No. 12440046
Dulbecco's modified Eagle Medium	Invitrogen	Cat. No. 21969-035
fetal bovine serum	ThermoFisher	Cat. No. 10270106
Penicilline/Streptomycin	ThermoFisher	Cat. No. 15140122
Glutamine	ThermoFisher	Cat. No. 25030024
Trypsin-EDTA 0.05%	ThermoFisher	Cat. No. 25300-054
Propidium Iodide	Sigma	Cat. No. P4170
30% acrylamide (37.5:1)	National Diagnostic	Cat. No. EC-890
N,N,N′,N′-Tetramethylethylendiamin	Merck	Cat. No. 1107320100
Tween 20	SIGMA	Cat. No. P7949
Amersham Prime Western Blotting Detection Reagen	GE Healthcare	Cat. No. RPN2232
SuperSignal West Femto Maximum Sensitivity Substrate	ThermoFisher	Cat. No. 34095
Pierce ECL 2 Substrate	ThermoFisher	Cat. No. 1896433A
Phire Green Hot Start II PCR Master Mix	Thermo Scientific	Cat. No. F126L
RNA Ligase 2	NEB	Cat. No.M0242S
RNAlater reagent	Invitrogen	Cat. No. AM7021
SUPERaseIn RNase inhibitor	Ambion	Cat. No. AM2694
RNase A	Sigma	R6513
Fetal Bovine serum	Gibco	Cat. No. 10270

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RNAse I	RNAse I	Cat. No. AM2295
Washing buffer B	Thermo Scientific	11900D
Critical commercial assays		
MinElute Gel Extraction Kit	Qiagen	Cat. No. 28604
Dynabeads™ Oligo(dT)25	ThermoFisher	Cat. No. 61005
RiboCop rRNA Depletion Kit V2 H/M/R	Lexogen	Cat. No. 144
RNA Clean and Concentrator kit	Zymo Research	Cat. No. R1017
AllPrep DNA/RNA/Protein Mini Kit	Qiagen	ID: 80004
DC Protein Assay (Bio-Rad, Cat. No. 5000112)	DC Protein Assay (Bio-Rad, Cat. No. 5000112)	DC Protein Assay (Bio-Rad, Cat. No. 5000112)
Deposited data		
Deep sequencing datasets	This study	GEO: GSE235348
All raw gel data will be deposited at Mendeley Data.	This study	https://doi.org/10.17632/rv9kgtcjpv.1
Experimental models: Cell lines		
Hap1 cells	Horizon Discovery	Cat. No. C631
Hap1 *CMTR1* KO cells	Horizon Discovery	Cat. No. HZGHC004217c007
Experimental models: Organisms/strains		
Mouse: *Cmtr1* knockout	The Jackson Laboratory	Stock no. 46174-JAX
Mouse: *Cmtr2 knockout*	MMRRC, UC Davis	Stock no. 047142-UCD
Mouse: *Cmtr1* floxed mouse	This study	Available from lead contact
Mouse: *AlbCreERT2*	Gift from David Gatfield lab	Schuler et al.[56]
Mouse: *Vasa-Cre*	Jackson Laboratory	Stock no. 6954
Oligonucleotides		
DNA and RNA oligos		See Table S1
Software and algorithms		
ENRICHR	Chen et al.[57]; Kuleshov et al.[58]	http://amp.pharm.mssm.edu/Enrichr/
R	R Core Team[59]	https://www.r-project.org
DESeq2	Love et al.[60]	N/A
Bioconductor	Huber et al.[61]	https://www.bioconductor.org/
Salmon	Patro et al.[62]	N/A
FastQC	Andrews[63]	N/A
Flexbar	Dodt et al.[64]	N/A
STAR	Dobin et al.[65]	N/A
Python 3.10.6		N/A
seaborn (v0.12.1)	Waskom[66]	https://doi.org/10.21105/joss.03021
matplotlib (v3.6.2)	Hunter[67]	N/A
gseapy (v0.14.0)	Fang et al.[68]	https://doi.org/10.1093/bioinformatics/btac757
bowtie	Langmead et al.[69]	N/A
CollectRnaSeqMetrics	Toolkit[70]	N/A
RiboCode toolkit	Xiao et al.[71]	N/A
SparK (v2.6.2) python library	Kurtenbach and Harbor[72]	N/A
SUPPA2	Trincado et al.[73]	N/A
SCDC (v0.0.0.9000) R package	Dong et al.[74]	N/A
sva (v3.46.0) bioconductor package	Leek et al.[75]	N/A
plastid python package	Dunn and Weissman[76]	N/A
sklearn (v1.1.3)	Pedregosa et al.[77]	N/A
Kaluza software	Beckman	RRID:SCR_016182

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Other		
Amersham Hyperfilm ECL	GE Healthcare	Cat. No. 28906837
Amersham Protran 0.45 μm nitrocellulose membrane	GE Healthcare	Cat. No. 10600002

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 generated in this study are available from the lead contact without any restriction. The *Cmtr1* knockout mouse (MMRRC Stock No. 46174-JAX) was obtained from The Jackson Laboratory, while the *Cmtr2* knockout mouse (MMRRC Stock No. 047142-UCD) was from the MMRRC, UC Davis. The *Ddx4-Cre* (*Vasa-Cre*) transgenic line was obtained from the Jackson Laboratory (Jackson Laboratory, Stock no. 6954). And the Alb-CreERT2 mouse was a gift from David Gatfield, University of Lausanne, Switzerland. The HAP1 *CMTR1* knockout cell line (Horizon Discovery, Cat. No. HZGHC004217c007) and control wildtype cells (Horizon Discovery, Cat. No. C631) were purchased.

Data and code availability

- Deep sequencing data generated in this study are deposited with Gene Expression Omnibus (GEO) under accession no. GEO: GSE235348
- Code used in the current study is available from the lead contact upon reasonable request.
- Other raw data associated with this study are deposited with Mendeley Data (https://doi.org/10.17632/rv9kgtcjpv.1).

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Animal work

The *Cmtr1* floxed mouse was generated at the Transgenic Mouse Facility of University of Geneva, while the *Cmtr1* and *Cmtr2* knockout models were generated by the Knockout Mouse Project (KOMP). 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. Animals were maintained in ventilated cages with unrestricted supply of water and food. All adult experimental animals were sacrificed by intraperitoneal injection of 150 mg/kg pentobarbital followed by the cervical dislocation, while decapitation was used for P0 animals. This work was approved by the Canton of Geneva (GE/16/219C and GE297).

Mouse mutants

The *Cmtr1* (C57BL/6NJ-*Cmtr1*$^{tm1(IMPC)J}$/Mmjax) knockout mouse was purchased from The Jackson Laboratory (strain # 032957; MMRRC Stock no. 46174-JAX). The heterozygous mutant animals obtained from the Jackson Laboratory (via Charles River) were crossed with wildtype C57BL/6J (Janvier; stock no. SC-C57J-F; SC-C57J-M) partners to expand the colony.

The *Cmtr2* mutant mouse (C57BL/6N-*Cmtr2*$^{tm1.1(KOMP)Vlcg}$/JMmucd, MMRRC_047142-UCD, Stock no. 047142-UCD) was generated by the Knockout Mouse Phenotyping Program (KOMP) Repository, and obtained from the Mutant Mouse Resource and Research Center (MMRRC), University of California at Davis. We crossed the heterozygous mutant animals obtained from the UC Davis with wildtype C57BL/6J (Janvier; stock no. SC-C57J-F; SC-C57J-M) partners to expand the colony.

We created *Cmtr1* conditional knockout mice by knock-in of *loxP* sites in the same direction (recombination by the Cre recombinase should result in deletion of the intervening region) flanking the exon 3 of *Cmtr1* (Figure S2A). Founder mice were crossed with wildtype C57BL/6J (Janvier) partners to obtain germline transmission. Homozygous *Cmtr1*$^{loxP/loxP}$ animals are viable and fertile.

We prepared the conditional knockouts (cKO) (*Cmtr1*$^{loxP/-}$; *Alb-CreERT2*$^{KI/+}$ mice) and control (*Cmtr1*$^{loxP/+}$; *Alb-CreERT2*$^{KI/+}$) animals to delete the gene in mouse liver. Animals (n = 4) were intraperitoneally injected with Tamoxifen (75 μg/g of body weight) to

induce gene deletion in adult animals: 3 months-old (for day 2 experiment), 4 months-old (for day 6 experiment) and 8 months-old (for day 22 experiment) (Figure 3E, 3H, and 3I). We did these experiments in two batches. The first batch had only one time point (day2: 4 days of injection and analysis at day 2 post-tamoxifen injection) (Figures 3C–3G and 3J–3M) with control and conditional KO liver samples. To observe the chronic effects of loss of *Cmtr1*, we performed a second experiment with three time-points (day 2, 6 and 22) (GO term analysis in Figure S3C and volcano plots in Figures 3H and 3I) with control and conditional KO liver samples for each time point. At least three biological replicates were used for each time point (Table S2). Ribosome profiling (n = 4) was done with the first batch of liver samples from the day 2 time point. Sucrose-gradient analysis to obtain the polysome profiling data was conducted in duplicates only (Figures 3K and S3D).

The *Ddx4-Cre* (*Vasa-Cre*) transgenic line (Jackson Laboratory, Stock no. 6954) expresses the Cre recombinase from the *Ddx4* (*Vasa*) promoter.[78] We crossed $Cmtr1^{loxP/loxP}$ animals with $Cmtr1^{+/-}$;$Ddx4^{KI/+}$ mice to prepare the conditional knockouts (cKO) ($Cmtr1^{loxP/-}$; $Ddx4^{KI/+}$) and control animals ($Cmtr1^{loxP/+}$; $Ddx4^{KI/+}$). The testes (P0, P31 and P75) and ovaries (>P60) from cKO and control animals of the indicated ages were collected for histological and/or transcriptome analysis.

Mouse embryos (E6.5, E7.5 etc) were microdissected, imaged using stereomicroscope Discovery.V12 (Zeiss), and stored in RNA-later reagent (Invitrogen, Cat. No. AM7021) until RNA extraction. Samples were washed 3 times in ice-cold 1xPBS prior the RNA and DNA extraction by the AllPrep DNA/RNA/Protein Mini Kit (Qiagen, Cat. No./ID: 80004) and genotyped as described above.

Human HAP1 *CMTR1* KO cells

HAP1 is a near-haploid human cell line derived from the chronic myelogenous leukemia (CML) cell line KBM-7. The HAP1 *CMTR1* knockout cell line (Horizon Discovery, Cat. No. HZGHC004217c007) and control wildtype cells (Horizon Discovery, Cat. No. C631) were purchased. The *CMTR1* KO cell line has a 2 bp deletion in the target locus and was generated using the CRISPR-Cas9 technology. Western analysis confirms the complete lack of CMTR1 protein (Figure S4A) and RNA mass spectrometry reveals the complete absence of m^6Am (Figure S4B), a modification that depends on CMTR1-dependent cap1 methylation.[11]

METHOD DETAILS

Cmtr1 knockout mouse

The *Cmtr1* (C57BL/6NJ-$Cmtr1^{em1(IMPC)J}$/Mmjax) knockout mouse was purchased from The Jackson Laboratory (strain # 032957; MMRRC Stock no. 46174-JAX). It was generated by the Knockout Mouse Phenotyping Program (KOMP) at the Jackson Laboratory. The mouse *Cmtr1* gene locus on Chromosome 17 (NCBI: NM_028791) has 24 exons. The mutant was prepared by electroporation of two gRNAs (GCAGGACCCACACTAGACAT and GGTGGGGCACAAGTTAGCAC) targeting intronic regions flanking the exon 3 of the *Cmtr1* gene (Figure S1A). The guide RNAs and Cas9 endonuclease were introduced into single mouse embryos (C57BL/6NJ; The Jackson Laboratory Stock No. 5304). This makes a 344 bp deletion beginning at Chromosome 17 position 29,674,049 bp and ending after 29,674,392 bp (GRCm38/mm10). This deletes the entire exon 3 and 192 bp of flanking intronic sequence (Figure S1A). The heterozygous mutant animals obtained from the Jackson Laboratory (via Charles River) were crossed with wildtype C57BL/6J (Janvier; stock no. SC-C57J-F; SC-C57J-M) partners to expand the colony.

Cmtr2 knockout mouse

The *Cmtr2* mutant mouse (C57BL/6N-$Cmtr2^{tm1.1(KOMP)Vlcg}$/JMmucd, MMRRC_047142-UCD, Stock no. 047142-UCD) was generated by the Knockout Mouse Phenotyping Program (KOMP) Repository, and obtained from the Mutant Mouse Resource and Research Center (MMRRC), University of California at Davis. The mouse *Cmtr2* gene locus on Chromosome 8 has two exons, with the protein encoded by sequences in the exon 2. The mutant mice were created (Velocigene) by a targeted mutation of the locus via homologous recombination in mouse embryonic stem (ES) cells (VGB6 derived from C57BL/6NTac). This results in deletion of 2311 bp that includes the whole mouse CMTR2 coding sequence in exon 2 (mm10, chr8: 110,221,063–110,223,373 is deleted), and insertion of a cassette (*LacZ-loxP-Neo-loxP*) in its place (http://velocigene.com/komp/detail/15502) (Figure S5A). Such a *Cmtr2* knockout ES cell clone (15502A-C9) was injected into morulae or blastocysts. Resulting chimera founders were mated to C57BL/6N mice to obtain germline transmission. The obtained heterozygous animals were then bred to a ubiquitous Cre deleter mouse line for recombination of the *LoxP* sites to remove the *Neo* gene from the inserted cassette by Cre recombinase. The end result is that the *Cmtr2* knockout mouse lacks the CMTR2 coding sequence, leaving the *LacZ* coding sequence under control of the endogenous *Cmtr2* promoter. The Cre transgene was removed during the crosses. The MMRRC used C57BL/6N females for cryo-recovery. We crossed the heterozygous mutant animals obtained from the UC Davis with wildtype C57BL/6J (Janvier; stock no. SC-C57J-F; SC-C57J-M) partners to expand the colony.

Creation of *Cmtr1 loxP* mice

The *Cmtr1* genomic locus is located on mouse chromosome 17 and consist of 23 exons. We created *Cmtr1* conditional knockout mice by knock-in of *loxP* sites in the same direction (recombination by the Cre recombinase should result in deletion of the intervening region) flanking the exon 3 of *Cmtr1* (Figure S2A). The ssDNA had a central region with two *loxP* sites at positions 85 nt upstream and

55 nt downstream of exon 3, with 70nt homology arms at each end. We introduced mutations into the ssDNA to prevent repeat cleavage of the target sites: the upstream gRNA recognition site is disrupted by insertion of the *loxP* site, while we mutated the PAM sequence at the downstream site.

The tracrRNA and crRNAs (crRNA_Cmtr1_1 and crRNA_Cmtr1_2) (IDT) (Table S1) were annealed in thermocycler in two separate reactions: 2 μL tracrRNA (200 pmol; IDT; Cat. No. 1072533) and 2 μL crRNA (200 pmol; IDT) (Table S1) were mixed with 6 μL IDTE buffer (pH 7.5; IDT, Cat. No. 356429). Annealed gRNAs were stored at −70°C. The injection mix was prepared freshly on the day of mouse oocyte injections. The annealed gRNAs (to final concentration 0.6 pmol/μL) were mixed with the pre-diluted Cas9 3xNLS protein (to final concentration 30 ng/μL; IDT, Cat. No. 1081058) in a volume of 9 μL. The mix was incubated at room temperature for 10 min for complex formation and mixed with the ssDNA repair template (Genewiz; 10 ng/μL final concentration). The volume of this injection mix was adjusted with the IDTE buffer to a final volume of 100 μL. The injection mix was centrifuged at 13000 rpm for 5 min at 4°C, and 50 μL of supernatant was transferred to a new microcentrifuge tube and stored on ice. Mouse single-cell embryos of the B6D2F1/J hybrid line (also called B6D2; The Jackson Laboratory, stock no. 100006) were injected. The NMRI (Naval Medical Research Institute) mice, which have a white coat color, were used as foster mothers. Founder mice were identified by genotyping PCR and crossed with wildtype C57BL/6J (Janvier) partners to obtain germline transmission. Homozygous $Cmtr1^{loxP/loxP}$ animals are viable and fertile.

ssDNA repair template:

gRNAs are in italics, bold underlined are *loxP* sites insertions, in bold exon3, "A" in bold upstream of gRNA is a mutation in PAM sequence to avoid multiple cleavages.

ATACGTACGTATACAGCTGGCAAGAGTAGAGACGTCACTGTGACCTCCATTGAGTGCAGGACCC*ACACTA***ATAACTTCGTATAG**
CATACATTATACGAAGTTAT*GACATAGGTGGGAC*ATGTGGACTGTGGGTGCATGAGGCAGTCCTGTCATCCGGACCCACCTAACG
CTTCTCTTCTTCTTCCCCAG**CATCTGCTACAAGCCTCAGTGGATCTGACAGTGAGACCGAGGGGAAGCAGCCCTGCTCTGATGA**
TTTCAAAGATGCCTTCAAAGCAGATTCCCTTGTGGAGGGAACATCGTCCCGATATTCCATGTATAACAGTGTTTCCCAGAGGCT
TATGGTATGTCTTGGCTTAGAATGGACTTCTAAAGTTGCCCAAAAGAGGGAGAGGAAGA**ATAACTTCGTATAGCATACATTATACGA**
AGTTATGGGC**A***AGGGTGTACTGGTGTTGGGGA*GTGGGGTGGGGCACAAGTTAGCACAGGATATAGGTTCTGAGTAT.

Tamoxifen-inducible conditional *Cmtr1* deletion in mouse liver

The Alb-CreERT2 mouse[56] specifically expresses using the *Albumin* (Alb) promoter the tamoxifen-inducible CreERT2 in hepatocytes. As described by the authors,[56] a cassette consisting of an IRES with coding sequence for Cre-ERT2 was inserted (knock-in) downstream of the stop codon, in the 3′ UTR of the serum albumin (Alb) gene. The Alb-CreERT2 mouse was a gift from David Gatfield, University of Lausanne, Switzerland. The $Cmtr1^{loxP/loxP}$ mice were crossed with *Alb-CreERT2*$^{KI/KI}$ mice. In another cross, the $Cmtr1^{+/-}$ mice were crossed with the *Alb-CreERT2*$^{KI/KI}$ mice. Using the these lines we prepared the conditional knockouts (cKO) ($Cmtr1^{loxP/-}$; *Alb-CreERT2*$^{KI/+}$ mice) and control ($Cmtr1^{loxP/+}$; *Alb-CreERT2*$^{KI/+}$) animals. Animals (n = 4) were intraperitoneally injected with Tamoxifen (75 μg/g of body weight) to induce gene deletion in adult animals: 3 months-old (for day 2 experiment), 4 months-old (for day 6 experiment) and 8 months-old (for day 22 experiment) (Figures 3E, 3H, and 3I).

We did these experiments in two batches. The first batch had only one time point (day2: 4 days of injection and analysis at day 2 post-tamoxifen injection) (Figures 3C–3G and 3J–3M) with control and conditional KO liver samples. To observe the chronic effects of loss of *Cmtr1*, we performed a second experiment with three time-points (day 2, 6 and 22) (GO term analysis in Figure S3C and volcano plots in Figures 3H and 3I) with control and conditional KO liver samples for each time point. At least three biological replicates were used for each time point (Table S2). Ribosome profiling (n = 4) was done with the first batch of liver samples from the day 2 time point. Sucrose-gradient analysis to obtain the polysome profiling data was conducted in duplicates only (Figures 3K and S3D).

Tamoxifen (Sigma, Cat. No. T5648-1G) was diluted in corn oil (Sigma, Cat. No. C8267-500ML) and dissolved overnight at room temperature, protected from light, and stored at 4°C for up to 2 days. Daily injections of tamoxifen were given for four consecutive days after which the mice were sacrificed at different time points: 2 days later (6 days after start of the experiment); 6 days later (10 after start of the experiment) or 22 days later (26 after start of the experiment). After injections, the animals were monitored daily for change in body weight, signs of general discomfort and behavior changes. We observed total loss of the CMTR1 protein already at 2 days-post tamoxifen injections (Figure 3C), but the longer analysis time-points were used to detect long-term gene expression consequences of loss of the protein. By analysis of the transcriptome at the three time-points, chronic activation of the innate immune pathway was observed in the cKO *Cmtr1* liver tissue (Figures 3E, 3H, and 3I). Animals were euthanized by pentobarbital injection followed by cervical dislocation and livers were collected.

Conditional *Cmtr1* germline knockout mice

The *Ddx4-Cre* (*Vasa-Cre*) transgenic line (Jackson Laboratory, Stock no. 6954) expresses the Cre recombinase from the *Ddx4* (*Vasa*) promoter.[78] There are multiple copies of this transgene in this line. The obtained animals were first twice backcrossed with wildtype C57BL/6J prior to other crosses. We crossed $Cmtr1^{loxP/loxP}$ animals with $Cmtr1^{+/-}$;$Ddx4^{KI/+}$ mice to prepare the conditional knockouts (cKO) ($Cmtr1^{loxP/-}$; $Ddx4^{KI/+}$) and control animals ($Cmtr1^{loxP/+}$; $Ddx4^{KI/+}$). The testes (P0, P31 and P75) and ovaries (>P60) from cKO and control animals of the indicated ages were collected for histological and/or transcriptome analysis. Expression of the Cre recombinase starts at embryonic day 14.5 (E14.5) in the male and female germline. Meiosis is initiated at E13.5 in the female germline, and at P8 in the male germline. The cKO males were found to be infertile, while cKO females displayed low penetrant infertility

(Figures 2B–2E). We speculate that deletion of *Cmtr1* after initiation of meiosis in the female germline may be the reason for the not dramatically affect progression of oogenesis.

Mouse genotyping

Ear-punches of the weaned animals (21 days-old) were digested in 100 μL of Lysis buffer (10 mM NaOH, 0.1 mM EDTA) for 90 min at 95°C. After centrifugation at 3000 rcf for 10 min, 50 μL of the supernatant was transferred to a new tube containing 50 μL of TE buffer (20 mM Tris-HCl, pH 8.0 and 0.1 mM EDTA). An aliquot of 2 μL of the digestion mix was used for the genotyping PCR. Reaction mix for 20 μL PCR reactions: 10 μL of Phire Green Hot Start II PCR Master Mix (Thermo Scientific, Cat. No. F126L), 1.0 μL of each primer (10 nM), 2.0 μL DNA from ear punches (100–200 ng), and 5–6.0 μL water to make 20 μL final volume.

Primers to genotype knockout allele of *Cmtr1* were RR1185 and RR1186 (Table S1). The expected size of products was 916 bp (WT) and 661 bp (KO). Reactions were run using the following conditions: 98°C for 30 s, 35 cycles of [98°C for 5 s, 65°C for 5 s and 72°C for 12 s], 72°C for 1 min, and finally at 12°C to hold the reaction. Reactions were examined by 1.5% agarose gel electrophoresis (Figure S1B).

To identify knockout allele of *Cmtr2*, PCR reaction with three primers (MD280, MD282, MD283) were used (Table S1). Reactions were run using the following conditions: 98°C for 30 s, 35 cycles of [98°C for 5 s, 65°C for 5 s and 72°C for 12 s], 72°C for 1 min, and finally at 12°C to hold the reaction. PCR products of 244 bp (WT) and 339 bp (KO) were resolved by 2% agarose gel electrophoresis (Figure S5B).

To genotype *Cmtr1 LoxP* allele, the primers were MM458 and MM448 with the expected PCR product size 408 bp (WT) and 476 (KI) (Table S1). Reactions were run using the following conditions: 98°C for 30 s, 35 cycles of [98°C for 5 s, 65°C for 5 s and 72°C for 12 s], 72°C for 1 min, and finally at 12°C. Reactions were examined by 2% agarose gel electrophoresis (Figure S2B and S3A).

To genotype *DDX4* transgene, the primers were MM113 and MM114 with the expected PCR product size for the transgene fragment being 275 bp (Table S1). Reactions were run using the following conditions: 98°C for 30 s, 35 cycles of [98°C for 5 s, 65°C for 5 s and 72°C for 12 s], 72°C for 1 min, and finally at 12°C. Reactions were examined by 2% agarose gel electrophoresis (Figure S2B).

In case of the *Alb-CreERT2* animals, the WT allele was screened with ABT290 and ABV93 primers (Table S1), whereas the KI allele was detected with ABT290 and ABT294 in two different reactions with the following reaction conditions: 98°C for 30 s, 35 cycles of [98°C for 5 s, 55°C for 5 s and 72°C for 12 s], 72°C for 1 min, and finally at 12°C. PCR products were mixed together and examined by 2% agarose gel electrophoresis, where 444 bp (KI) and 229 bp (WT) bands were detected (Figure S3A).

Collection of mouse embryos

Adult (8 weeks or older) animals of the heterozygous genotypes for *Cmtr1* or *Cmtr2* knockout alleles were crossed together. Plugs were checked the morning after and considered as embryonic day 0.5 (E0.5). Plugged females were separated and sacrificed later at E6.5 to E10.5. Embryos were microdissected, imaged using stereomicroscope Discovery.V12 (Zeiss), and stored in RNAlater reagent (Invitrogen, Cat. No. AM7021) until RNA extraction. Samples were washed 3 times in ice-cold 1xPBS prior the RNA and DNA extraction by the AllPrep DNA/RNA/Protein Mini Kit (Qiagen, Cat. No./ID: 80004) and genotyped as described above.

Human HAP1 *CMTR1* KO cells

HAP1 is a near-haploid human cell line derived from the chronic myelogenous leukemia (CML) cell line KBM-7. The HAP1 *CMTR1* knockout cell line (Horizon Discovery, Cat. No. HZGHC004217c007) and control wildtype cells (Horizon Discovery, Cat. No. C631) were purchased. The *CMTR1* KO cell line has a 2 bp deletion in the target locus and was generated using the CRISPR-Cas9 technology. Western analysis confirms the complete lack of CMTR1 protein (Figure S4A) and RNA mass spectrometry reveals the complete absence of m^6Am (Figure S4B), a modification that depends on CMTR1-dependent cap1 methylation.[11]

Clones and constructs

The complementary DNA (cDNA) for human CMTR1 (NCBI: NP_055865.1), human DHX15 (NCBI: NP_001349.2) and mouse CMTR2 (NCBI: NP_666327.2) were obtained by reverse transcription-PCR (RT-PCR) amplification from human cell culture or mouse tissue RNA. To express the protein in insect cell expression system, full-length hCMTR1, hDHX15 or truncated mCMTR2 (15–759 aa) coding sequence was cloned into the modified pACEBac2SS vector for expression as an N-terminal 6xHis-Strep-SUMO-TEV fusion protein. For co-expression of hCMTR1 and hDHX15, hCMTR1 coding sequence was cloned into pIDK vector and then recombined with pACEBac2SS-hDHX15 via Cre-recombination. All constructs were verified by restricted digestion as well as by Sanger sequencing.

Recombinant protein production

Production of full-length hCMTR1, hDHX15 or truncated mCMTR2 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 *Sf*9 cells derived from the fall army worm *Spodoptera frugiperda*. Briefly, pACEBac2SS plasmids carrying target genes 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 h post-transfection. To expand the virus pool, 3.0 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 after infection. For expression of the CMTR1-DHX15 complex, recombined plasmid carrying both CMTR1 and DHX15 gene was transformed into DH10EMBacY competent cells, followed by bacmid extraction and baculovirus preparation as described above.

Purification of mouse CMTR2 protein

After protein expression, the cells were collected by centrifugation and lysed by sonication in buffer: 50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 5% Glycerol, 5 mM 2-mercaptoethanol, 40 mM Imidazole and protease inhibitor (Thermo Scientific, EDTA-free). Clear supernatant was collected by centrifugation at 18,000 rpm for 45 min at 4°C. The supernatant was incubated for 2 h with Ni^{2+} chelating Sepharose FF beads at 4°C, then the beads were washed by imidazole gradient washing buffer and finally bound protein was eluted with 250 mM imidazole in lysis buffer. Subsequently, the N-terminal tag was cleaved by the TEV protease overnight in the dialysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 5 mM 2-mercaptoethanol). The cleaved tag was removed by a second purification on Nickel beads. The protein was further purified by gel filtration chromatography (GE Healthcare, Superdex 200 increase 10/300) equilibrated with buffer (25 mM Tris-HCl pH 8.0, 150 mM KCl, 5% glycerol and 1 mM DTT). The pure fractions were verified by SDS-PAGE electrophoresis and flash-frozen in liquid nitrogen.

Purification of human CMTR1-DHX15 complex

Hi5 cells co-expressing untagged hCMTR1 and 6xHis-Strep-SUMO-TEV-hDXH15 were collected and resuspended in lysis buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 40 mM Imidazole, 5% glycerol and 5 mM 2-mercaptoethanol) supplemented with protease inhibitor (Thermo Scientific, EDTA-free). After sonication, the lysate was centrifuged at 18,000 rpm for 45 min at 4°C. The clarified supernatant was incubated at 4°C for 2h with the Ni^{2+} chelating Sepharose FF beads. The beads were washed with an imidazole gradient in the wash buffer (40 mM, 50 mM or 60 mM imidazole in lysis buffer) and bound protein complex was eluted by 250 mM Imidazole. The N-terminal His-Strep-SUMO tag was further removed by TEV protease overnight at 4°C in dialysis buffer (50 mM Tris-HCl pH 8.0, 200 mM NaCl, 5 mM 2-mercaptoethanol). After cleavage, second nickel column purification was performed and flow-through containing the cleaved protein (complex) was collected. The complex was further purified by gel filtration chromatography using Superdex 200 increase 10/300 (GE Healthcare) equilibrated with buffer (25 mM Tris-HCl pH 8.0, 150 mM KCl, 5% glycerol and 1 mM DTT). The pure fractions were verified by SDS-PAGE electrophoresis (Figure S4J) and flash-frozen in liquid nitrogen.

Antibodies

Commercial antibodies

Primary antibodies: rabbit or mouse anti-CMTR1 (Atlas antibodies, Cat. No. HPA029980), anti-IFIT1 (Cell signaling, Cat. No. D2X9Z), rabbit anti-TUBULIN (abcam, Cat. No. ab6046) and rabbit anti-PARK7 (abcam, Cat. No. ab18257).

Secondary antibodies: For Western blot analyses, the following secondary antibodies conjugated to HorseRadish Peroxidase were used: Amersham ECL Rabbit IgG, HRP-linked whole Ab (from donkey) (Cat. No. GE Healthcare, NA934-1ML) and Rabbit anti-Mouse IgG (H + L) Superclonal Secondary Antibody, HRP conjugate (Invitrogen, Cat. No. A27025).

Antibodies generated for this study

We generated rabbit polyclonal antibodies to mouse CMTR2 (mCMTR2). Two New Zealand White (NZW) rabbits were immunized with the soluble antigen (Biotem, France). The antigen used was the purified untagged mouse CMTR2 (15–759 aa) produced in insect cells. For each injection, 1 mg/mL protein was used. After six injections (at day 0, 14, 28, 56, 70 and 89) crude immune serum was collected (at day 96) and frozen. The anti-mCMTR2 crude sera detected the recombinant mouse protein by Western analysis, but failed to detect the protein in mouse tissue lysates. This could be due to low abundance of CMTR2 in mouse tissues or due to the low titer of the antibodies generated. Affinity purification of the antibodies with the antigen did not help to improve the situation.

Collection of RNA from HAP1 cells

Wildtype and *CMTR1* knockout human HAP1 cells were cultured in (High glucose) IMDM Medium (Gibco, Cat. No. 12440046) supplemented with 10% fetal bovine serum (FBS) and Penicillin-Streptomycin (10,000 U/mL) 1:100 (Gibco, Cat. No. 15140122) at 37°C in the presence of 5% CO2. At 70% confluency, the media was removed and TRIzol Reagent (Invitrogen, Cat. No. 15596-026) was added to monolayer of cells. Mixture was collected to a micro centrifuge tube, flash frozen and stored at −70°C until the RNA extraction.

RNA extraction from cell lines

After harvesting, pellets of human HAP1 cells were directly mixed with the TRIzol reagent (Invitrogen, Cat. No. 15596026), flash frozen and kept at −70°C. They were then processed according to the manufacture's protocol. To remove genomic DNA, approximately 10 μg of the extracted RNA was treated with TurboDNase (Invitrogen, Cat. No. AM2238) for 30 min and then cleaned by RNA Clean and Concentrator kit (Zymo Research, Cat. No. R1017).

Growth curve

We seeded 0.5 million wildtype or *CMTR1* knockout human HAP1 cells into wells of a 12-well plate and for the next 4 days cells were collected daily by trypsinization and counted in duplicate by Countess 3 Automated cell (Invitrogen) counter in trypan blue (Sigma, Cat. No. 93595-50ML) mixture 1:1. This experiment was repeated six times to obtain the data presented (Figure S4C).

Analysis of cell cycle by FACS

Wildtype and *CMTR1* knockout human HAP1 cells were grown to 60–70% confluency and then collected by trypsinization. Cell suspension was centrifuged at 500xg for 5 min, followed by removal of the trypsin/media. Cells were washed once in 1xPBS and centrifuged again. The PBS was removed and cells were completely resuspended in 100 μL of ice-cold 1xPBS. Afterward, cells were fixed by addition of ice-cold 100% Methanol, mixed and stored at −20°C until staining and further analysis. Later, fixed cells were centrifuged at 500xg at 4°C for 10 min. Supernatant was removed and cells were permeabilized with 150 μL PBS with 0.2% Triton X- and incubated for 30 min at room temperature. Cells were centrifuged, supernatant removed, and cells were resuspended in 500 μL of staining solution [15 μg/mL Propidium Iodide (Sigma, Cat. N P4170) and 6 μg/mL RNase A (Sigma Cat. no. R6513)]. Staining was performed for at least 1H in dark at 4°C. FACS analysis was performed in the staining solution using Gallios Flow Cytometer (Beckman) and analyzed by Kaluza software (Figure S4D).

Collecting mouse tissues for western blot

Multiple tissues were isolated from an adult (>P60) mouse. After flash-freezing in liquid nitrogen, a piece of different tissues were omogenized 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 14000xg for 30 min, and the supernatant collected. An aliquot was taken to measure the concentration by DC Protein Assay (Bio-Rad, Cat. No. 5000112). Lysate concentrations were normalized to 1 mg/μL. Protein extracts were stored at −70°C. The SDS loading buffer was added to the protein lysates and boiled at 95°C for 5 min, and 30 μg of protein per lane was loaded and resolved by SDS-PAGE (Figure 3A).

Western Blot

Whole cell lysates or tissues were separated via SDS-PAGE in order to detect proteins of interest in 10% or 12% polyacrylamide gels. Gel electrophoresis was performed at 120 V for 110 min. After separation, proteins were blotted on the Amersham Protran 0.45 μm 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 PBS and blocked for 1 h at room temperature with 5% dry milk in PBS with 0.05% Tween 20 (PBST) (Sigma, Cat. No. P7949). After this, membranes were incubated with primary antibody overnight at 4°C 1:10000 rabbit anti-TUBULIN (Abcam, Cat. No. ab6046), 1:100 anti-CMTR1 (Atlas antibodies, Cat. No. HPA029980), 1:200 anti-IFIT1 (Cell signaling, Cat. No. D2X9Z), 1:100 anti-PARK7 (abcam, Cat. No. ab18257). Then, membranes were washed 3 times for 10 min with PBST and incubated with HRP-conjugated secondary antibody at 1:10 000 dilution, either with anti-rabbit IgG HRP-linked (GE Healthcare, Cat. No. NA934) or anti-mouse IgG HRP-linked (Invitrogen, Cat. No. a27025) for 1 h at room temperature in 5% milk in PBST. After 1 h, membranes were washed 3 times for 10 min with PBST and incubated with one of detection reagents: Amersham Prime Western Blotting Detection Reagent (GE Healthcare, Cat. No. RPN2232), SuperSignal West Femto Maximum Sensitivity Substrate (ThermoFisher, Cat. No. 34095) or Pierce ECL 2 Substrate (ThermoFisher, Cat. No. 1896433A) for 5 min at room temperature. 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).

Histological analysis of mouse tissues

To prepare the paraffin sections, the mouse tissues were washed in 1xPBS, and fixed in 4% paraformaldehyde overnight at 4°C. After washing in 1xPBS, 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% ethanol (2 × 3h), 90% (1h), 95% (1h) and 100% ethanol (3 × 30 min) followed by incubation (3 × 30 min) in xylene. Xylene was removed and replaced with paraffin, and incubated at 56–58°C. Tissues were then transferred into plastic molds (Polysciences mold S-22; NC0397999) filled with paraffin, and paraffin was allowed to solidify at room temperature. The tissue sections (~5 μM thickness) were prepared using a microtome. The sections were allowed to stretch at 42°C and then stored at room temperature.

For histological analysis, the slides containing the paraffin sections were placed in a glass slide holder filled with xylene (3 × 5 min) to remove the paraffin. For rehydration, the slides were incubated in 100% ethanol, 96% ethanol, 70% ethanol, 50% ethanol (2 × 10 min for each step) and miliQ water (2 × 2 min for each step). Sections were stained with Hematoxylin solution (Merck) for 3 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. The sections were examined and pictures were taken using widefield (Zeiss Axio Imager Z1 or Axio M2) microscopy.

Total RNA purification from mouse liver samples

For total RNA extraction, we used multiple biological replicates (n = 4) of mouse liver from control (*Cmtr1*$^{loxP/+}$*; Alb-Cre-ERT2*) and *Cmtr1* cKO (*Cmtr1*$^{loxP/-}$*; Alb-Cre-ERT2*) animals after tamoxifen injections. Approximately 0.5 g tissue was taken and placed in a 50 mL conical tube (Sarstedt, Cat. No. 62.547.254) with 5 mL of extraction buffer. [Preparation of the extraction buffer: 250 g

guanidium thiocyanate (ITW Reagents, Cat. No. A1107), 17.6 mL sodium citrate, 0.75 M, pH 7.0 (Sigma-Aldrich, Cat. No. C8532) and 320 mL water were mixed at 60°C. Add 1/10 volume of sodium acetate, 2 M, pH 4.0 (Merck, Cat. No. 1.06268); 1/100 volume β-mercaptoethanol (Sigma, M3148), before use.]

Homogenize using a douncer (Kinematica AG, Cat. No. PT 2500E) in the 50 mL conical tube for 20 s until no fragments are left, and then add 5 mL phenol-H2O, mix well and stand on ice. Add 2 mL chloroform (VWR, Cat. No. 8.22265.2500): isoamyl alcohol (Merck, Cat. No. W205702) (49:1), mix well, stand on ice for 15 min, and transfer all the solution to a 15 mL TPP centrifuge tube (Thermo Fisher, Cat. No. 91016). Spin down for 20 min at 4000 rpm at 4°C, transfer the upper phase (approximately 5 mL) to a new 15 mL TPP tube. Then add 1 volume (~5mL) of phenol-chloroform: isoamyl alcohol (48:2) mix. Shake well to mix. Spin down for 15 min at 4000 rpm at 4°C, transfer upper phase (approximately 4 mL) to a new 15 mL TPP tube. Add 4 mL isopropanol to precipitate the nucleic acids, and leave it at −20°C for 25 min. After this, spin down at 5000rpm, 15 min at 4°C, remove all the solution, and completely dry the tube by keeping it up-side down on a tissue to remove all the liquid. Add 6 mL 4M lithium chloride (Merck, Cat. No. 1.05679) to resuspend the precipitate. The volume used depends on the size of the pellet. Shake until the pellet completely dissolves. Note that the pellet contains DNA along with the RNA. Leave on ice for 5 min, and spin down at 4500 rpm for 15 min at 4°C, remove all the solution and dry the tube completely. Add 7 mL 75% ethanol to resuspend the precipitate and leave it at room temperature for 10 min. Spin down at 4500 rpm for 15 min, remove all the solution and dry the tube completely with a tissue. A pellet will be visible if the amount of RNA is abundant. Add 75% ethanol and keep sample on ice or at −20°C for longer time. Repeat the 75% ethanol step once again. And dry the tubes at room temperature for 45min by keeping them open. Dissolve the RNA precipitation with ~700 µl DEPC treated water to a proper concentration. Dissolve RNA by gently pipetting. After a short spin, measure concentration and tRNA into a new 1.5mL tube. Store it under −80°C.

PolyA+ RNA purification

PolyA+ RNA was purified using magnetic Dynabeads Oligo(dT)25 (ThermoFisher, Cat. No. 61005). In brief, total RNA (75 µg) was adjusted to 100 µL with nuclease free water. The RNA was heated to 65°C for 2 min, and placed on ice, to disrupt secondary structures. 200 µL (1 mg) of Dynabeads were transferred to a microcentrifuge tube and washed twice with 100 µL Binding Buffer (1 M LiCl2, 2 mM EDTA, pH 8.0, 20 mM Tris-HCl, pH 7.5). Beads in 100 µL Binding Buffer were mixed with 100 µL of previously heated RNA. Beads were mixed thoroughly and rotated on a roller or mixer for 10 min at room temperature to allow the mRNAs to anneal to the oligo (dT)25 on the beads. Unbound fraction was discarded and beads were twice washed with 200 µL Washing buffer B. Washing buffer was removed and the elution was perform in the 20 µL of nuclease free water by heating at 80°C for 2 min, immediately placed on the magnet and the eluted PolyA+ RNA was transferred to a new RNase-free tube. For RNA mass spectrometry experiments, we repeated polyA + purification for a total of three times. The eluted RNA from the previous round was diluted to 100 µL in binding buffer for subsequent second or third round of purification. Used beads were washed twice in Wash buffer. After their resuspension in 100 µL of Binding Buffer, the purification was repeated as in the first round of purification.

Quantification of RNA modifications using LC-MS/MS

RNA was hydrolyzed to ribonucleosides by 20 U benzonase (Santa Cruz Biotech) and 0.2 U nuclease P1 (Sigma) 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, 0.05 U phosphodiesterase I and 0.1 U alkaline phosphatase (Sigma) were added, and incubated further at 37°C for 1 h. Digested samples were precipitated with 3 volumes of acetonitrile, centrifuged (16,000xg, 30 min, 4°C), and supernatants were lyophilized and dissolved in a solution of stable isotope labeled internal standards for LC-MS/MS analysis. 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 A: water and B: methanol (both added 0.1% formic acid) at 0.22 mL/min, for modifications starting with 5% B for 0.5 min followed by 2.5 min of 5–15% B, 3.5 min of 15–95% B, and 4 min re-equilibration with 5% B. Unmodified nucleosides were chromatographed with a 4 min gradient of 5–95% B and 4 min re-equilibration with 5% B. Mass spectrometric detection was performed using an Agilent 6495 Triple Quadrupole system operating in positive electrospray ionization mode, monitoring the mass transitions 269.1-150.1 (m^6Am), 282.1-150.1 (m^6A), 282.1-136.1 (Am), 268.1 (A), 284.1-152.1 (G), 244.1-112.1 (C), 245.1-113.1 (U), 296.1-164.1 ($m^6{}_2A$), and 298.1-166.1 (m^7G), 285.1-153.1 (d_3-m6A), 301.1-152.1 (d_3-Gm), 273.1-136.1 ($^{13}C_5$-A), and 246.1-114.1 (d_2-C). The m^6Am modification is catalyzed by PCIF1.[11] It carries out the N^6 methylation (m^6A) of the transcription start site adenosine that is already methylated on the ribose (Am; cap1) by CMTR1. Therefore, we used m^6Am levels as a readout for reduction in cap1 levels in the conditions lacking CMTR1 activity (Figures 3J and S4B).

RNA library preparation and sequencing

Library preparation for embryos form the heterozygous (HET) *Cmtr1*$^{+/-}$ crosses (n = 4) was performed at the *Genomics Core Facility* (GeneCore, EMBL Heidelberg) using 50 ng (for E6.5 embryos) and 200 ng (for E7.5 embryos) of total RNA with stranded rRNAminus RNAseq protocol and sequenced on NextSeq 500, HI, 75SE (85SE).

Library preparation and sequencing of a second batch of embryos from the heterozygous (HET) *Cmtr1*$^{+/-}$ crosses or *Cmtr2*$^{+/-}$ crosses was performed at iGE3 facility at the University of Geneva. We used 50 ng (for E6.5 embryos) and 200 ng (for E7.5 embryos)

of total RNA for the Smarter Ribodepletion kit. The prepared libraries were sequenced on the Illumina HiSeq 4000 sequencer (iGE3 Genomics Platform, University of Geneva).

Libraries were prepared with total RNA (500 ng) from human HAP1 WT and *CMTR1* KO (n = 4) cells at the *Genomics Core Facility* (GeneCore, EMBL Heidelberg) sequencing facility using the stranded rRNAminus RNAseq protocol and sequenced on the HiSeq2000.

All sequencing libraries prepared are listed in Table S2.

Polysome and ribosome profiling

Mouse liver was isolated from biological duplicates of control and *Cmtr1* cKO animals described previously. The tissue was rapidly cut up into small pieces and snap frozen in liquid nitrogen.

For one sample of human HAP1 cells, three 10 cm plates of ~70% confluency were used. Cells were treated with 100 μg/mL cycloheximide (CHX) at 37°C, media was removed and the plate was placed on ice. Cells were washed with ice-cold PBS containing 100 μg/mL CHX and the PBS was completely removed. Cells were collected by scraping in 1mL of ice-cold PBS supplemented with 100 μg/mL CHX. Cells were collected in Eppendorf tubes on ice and spun down for 5 min at 1000 rpm (200xg), 4°C. Supernatant was discarded and cell pellet was flash-frozen and stored at −80°C. Human HAP1 cell pellets were flash-frozen until further use. The ribosome profiling[79,80] and polysome fractionation was performed at the "BioCode: RNA to Proteins" Core Facility, Faculty of Medicine, UNIGE.

Polysome gradient centrifugation

Mouse liver tissues or HAP1 cell pellets were mechanically disrupted in liquid nitrogen and homogenized in a lysis buffer (50 mM Tris, pH 7.4, 100 mM KCl, 1.5 mM MgCl2, 1.0% Triton X-100, 0.5% Na-Deoxycholate, 25 U/mL Turbo DNase I, 1mM DTT, 100 μg/mL cycloheximide, and Protease inhibitors (Roche). 10 μL of SUPERaseIn RNase inhibitor (Ambion, #AM2694) was added to lysis buffer upon homogenization. Cell debris were pelleted (20000xg, 20 min, 4°C). Approximately 500 μL of tissue lysates containing 2 μg/mL of total RNA (1000 μg of total RNA) were loaded on the linear 20–60% sucrose gradients prepared on the gradient buffer (50 mM Tris, pH 7.4, 100 mM KCl, 1.5 mM MgCl2, 1mM DTT, 100 μg/mL cycloheximide). Ribosomes were fractionated at 247′600xg (38′000 rpm, rotor SW41 Ti (Beckman Coulter, #331362) for 3 h 30 min at 4°C. Fractionated ribosomes were monitored and collected using Density Gradient Fractionation System (ISCO) (Figures 3K and S3D).

Ribosome profiling

Mouse liver tissues or HAP1 cell pellets were mechanically disrupted in liquid nitrogen and homogenized in a lysis buffer (50 mM Tris, pH 7.4, 100 mM KCl, 1.5 mM MgCl2, 1.0% Triton X-100, 0.5% Na-Deoxycholate, 25 U/mL Turbo DNase I, 1mM DTT, 100 μg/mL cycloheximide, and Protease inhibitors (Roche, Cat. No. 04693132001). For determination of the optimal concentration of nuclease for RNA digestion, the extracts were treated with the different amounts of RNAse I (Ambion, #AM2295). This was followed by sucrose gradient centrifugation to determine shift of polysomes to monosomes, but without further degradation of monosomes.

To obtain ribosome footprints, 0.12 mL of total extracts containing 300 μg of total RNA were treated with RNAse I (Ambion, #AM2295) (250U/1 mg of total RNA), for 45 min at 20°C with slow agitation. 10 μL SUPERaseIn RNase inhibitor (Ambion, #AM2694) was added to stop nuclease digestion. Monosomes were isolated using MicroSpin S-400 HR spin columns (Amersham, #27514001). For isolation of ribosome protected mRNA fragments (RPF), 3× volumes of QIAzol (Qiagen, Cat. No. 79306) were added to the S-400 eluate, mixed thoroughly, and RNA extracted with Direct-Zol RNA Mini Prep Plus kit (Zymo Research, #R2070).

Ribosome protected fragment (RPF) library preparation

Libraries were prepared as described.[79,80] Briefly, ribosome protected fragments (RPFs) were size-selected (25–34 nt) by electrophoresis using a 15% TBE-Urea polyacrylamide gel electrophoresis (PAGE) and two RNA markers, 25-mer (OP-RNA25) and 34-mer (OP-RNA34) (Table S1). After dephosphorylation with T4 Polynucleotide Kinase (NEB, #M0201S) the adapter Linker-1 (OP-RNA45) was ligated to the 3′ end of the RPF using T4 RNA Ligase 2 (NEB, #M0242S). The ligated products were purified using 10% TBE-Urea PAGE. Ribosomal RNA was removed using RiboCop rRNA Depletion Kit V2 H/M/R (Lexogen #144) for mouse liver samples or using the Ribo-Zero Plus rRNA Depletion kit (Illumina, #20040892) for HAP1 cell samples. The adapter Linker-1 was used for priming reverse transcription (RT) with the RT primer Ni-Ni-9 (OP27) using ProtoScript II Reverse Transcriptase (NEB, #M0368S). RT products were purified using 10% TBE-Urea PAGE. The cDNA was circularized with CircLigase II ssDNA Ligase (Epicentre, #CL9021K). The final libraries were generated by PCR using forward index primer NI-N-2 (OP28) and reverse index primers. Amplified libraries were purified using 8% TBE-PAGE and analyzed with TapeStation. Libraries were sequenced on an Illumina HiSeq 4000, single-reads, 1 × 50 bp, 8 libraries in a pool.

Isolation of total RNA, library preparation and sequencing

Total RNA was isolated from the same nuclease-treated extracts, that were used to obtain RPFs, using Direct-Zol RNA Mini Prep Plus kit (Zymo Research, #R2070). RNA was sent to iGE3 Genomic Platform, University of Geneva for stranded mRNA library preparation. Libraries were sequenced on an Illumina HiSeq 4000, single-end reads, 1 × 50 bp, 12 libraries in pool.

All sequencing libraries prepared are listed in Table S2.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis of RNA-seq data

Quality control of the demultiplexed libraries was performed with FastQC.[63] Sequencing adapters were removed by Flexbar software.[64] All the sequenced mouse libraries were aligned to the mouse transcriptome (GRCm38 assembly) with STAR[65] (v2.7.10a) and reads were further quantified transcript-wise from STAR-generated BAM files by salmon quant (v1.8.0)[62] with the following options: –seqBias –gcBias –posBias –writeUnmappedNames. Transcript counts were collapsed into gene counts and DESeq function of the DESeq2 bioconductor package[60,61] was used to obtain log2 fold changes of gene expression between control and mutant samples and the adjusted p values. Hidden noise was inferred by the sva bioconductor package[75] and 2 surrogate variables (SV1 and SV2) were included into DESeq function model (design = ~ SV1 + SV2 + condition). Adjusted p value 0.05, as well as the absolute value of log2 fold change 1 were defined as thresholds of statistical significance. All the visualizations were done in Python 3.10.6. Volcano plots of differential gene expression was plotted using scatterplot function of the seaborn (v0.12.1) package[66] with additional matplotlib (v3.6.2) customizations[67] (Table S3). Boxplots of the expression of immune sensors (Figure 1F), comparison of the alternative splicing events (Figure 1G), z-scores of the log2-transformed counts of the snoRNA-containing introns (Figure S1F), log2 fold changes of selected gene classes expression and translational efficiency (Figure S1G, S3F–S3G, and S4F), as well as follicle proportions (Figure S2D) were plotted with the boxplot function of the seaborn package. Stacked bar charts of embryo genotypes (Figures 1C and 4C), and bulk-to-single-cell projections (Figure S1E and S5H) were created with the plot function (kind = 'bar', stacked = True) from the pandas (v1.5.1) package. Heat maps of the type I IFN pathway genes (Figure 3G) and selected innate immune sensors across early embryonic stages (Figure S1D) were plotted with clustermap function with z_score and weighted hierarchical clustering options of the seaborn package. The early embryonic transcriptome data used for analysis in Figure 1F was from published sources (GEO: GSE45719 and GSE119945; Arrayexpress: E-MTAB-6967). Venn diagrams comparing *Cmtr1* and *Cmtr2* dysregulated gene sets in E6.5 and E7.5 embryos (Figure 4F) were plotted with venn2 function of the matplotlib_venn (v0.11.7) package. Genes found to be significantly up- or down-regulated in the mutants were searched for enriched Gene Ontology terms in the Biological Process and Molecular Function ontologies using ENRICHR[58] and plotted with barplot function of the gseapy (v0.14.0) package.[68] Violin plot comparing expression levels of RIG-I between HAP1 cells and control *Cmtr1* mice (Figure S4G) was plotted by violinplot function of the seaborn (v0.12.1) package.

Analysis of ribosome profiling data

Quality control of the demultiplexed libraries was performed with FastQC.[63] Sequencing adapters were trimmed with Flexbar.[64] Ribosomal RNA filtering was done by aligning reads to the rRNA index with bowtie[69] with the following options: -S -v 1 -a –best –strata. After rRNA filtering step remaining pool of reads was aligned to a relevant genome: GRCm38 assembly for the mouse liver data and the GRCh38 assembly for human HAP1 cells, respectively. Genome alignment was performed by STAR[65] without any further read clipping by adding –alignEndsType EndToEnd option. Read coverage biases were checked from the BAM files by the CollectRnaSeqMetrics program from "Picard Toolkit".[70] Metagene profile analysis of the read phasing of the mouse liver data (Figure S3E) was performed by metagene program from the plastid python package.[76] RiboCode toolkit[71] was used to detect longest translated ORFs (-l yes). Count tables of the longest translated ORFs for further differential translation efficiency analysis were made by ORFcount program from the RiboCode toolkit. First and last 10 translated codons were masked from counting (-f 10 -L 10). Hidden noise was inferred by sva bioconductor package[75] and 2 surrogate variables (SV1 and SV2) were included into model. Differential translational efficiency analysis was done by DESeq function from DESeq2 bioconductor package with the following parameters specified: test = "LRT", reduced = ~SV1 + SV2 + assay + condition. Model design was as follows: ~SV1 + SV2 + assay + condition + assay:condition. Volcano plots showing differential translational efficiency (Table S4) between mouse *Cmtr1* cKO (*Cmtr1*$^{-/loxP}$; Alb-CreERT$^{+/-}$) and control (*Cmtr1*$^{+/loxP}$; AlbCreERT$^{+/-}$) (Figure 3L), as well as between human HAP1 *CMTR1* KO and WT cells (Figure S4F) were plotted using scatterplot function of the seaborn (v0.12.1) package with additional matplotlib (v3.6.2) customizations. Coverage tracks comparing ribosome protected fragments (RPFs) and input read distributions across gene model (Figure 3M) were plotted by SparK (v2.6.2) python library.[72]

Analysis of alternative splicing events

The annotation file of the alternative splicing (AS) events (SE – skipped exon, MX – mutually exclusive exons, A5/3 – alternative 5′/3′ site, RI – retained intron, AF – alternative first exon, AL – alternative last exon) was generated by generateEvents script from the SUPPA2 toolkit.[73] All the AS events were filtered based on the total gene expression counts (total TPM counts per gene per all conditions >10). The PSI (Percent Spliced-In) scores (Figure 1G) representing proportions of reads supporting the event over total amount of reads per gene were computed by psiPerEvent script from the SUPPA2 toolkit both per all the retained AS events and per transcripts (percentage of reads supporting a specific AS event over total number of events per gene, and percentage of reads supporting a specific transcript over total amount of reads per gene, respectively). Cryptic and/or rare AS events with PSI scores <0.05 or >0.95 as well as events with <25% of missing values per event were filtered out from further

analysis. The remaining missing values were imputed from the 2 neighbors by the KNN algorithm (KNNImputer function) from the sklearn library package (v1.1.3).[77]

Bulk-to-single-cell deconvolution (RNA-seq debulking)

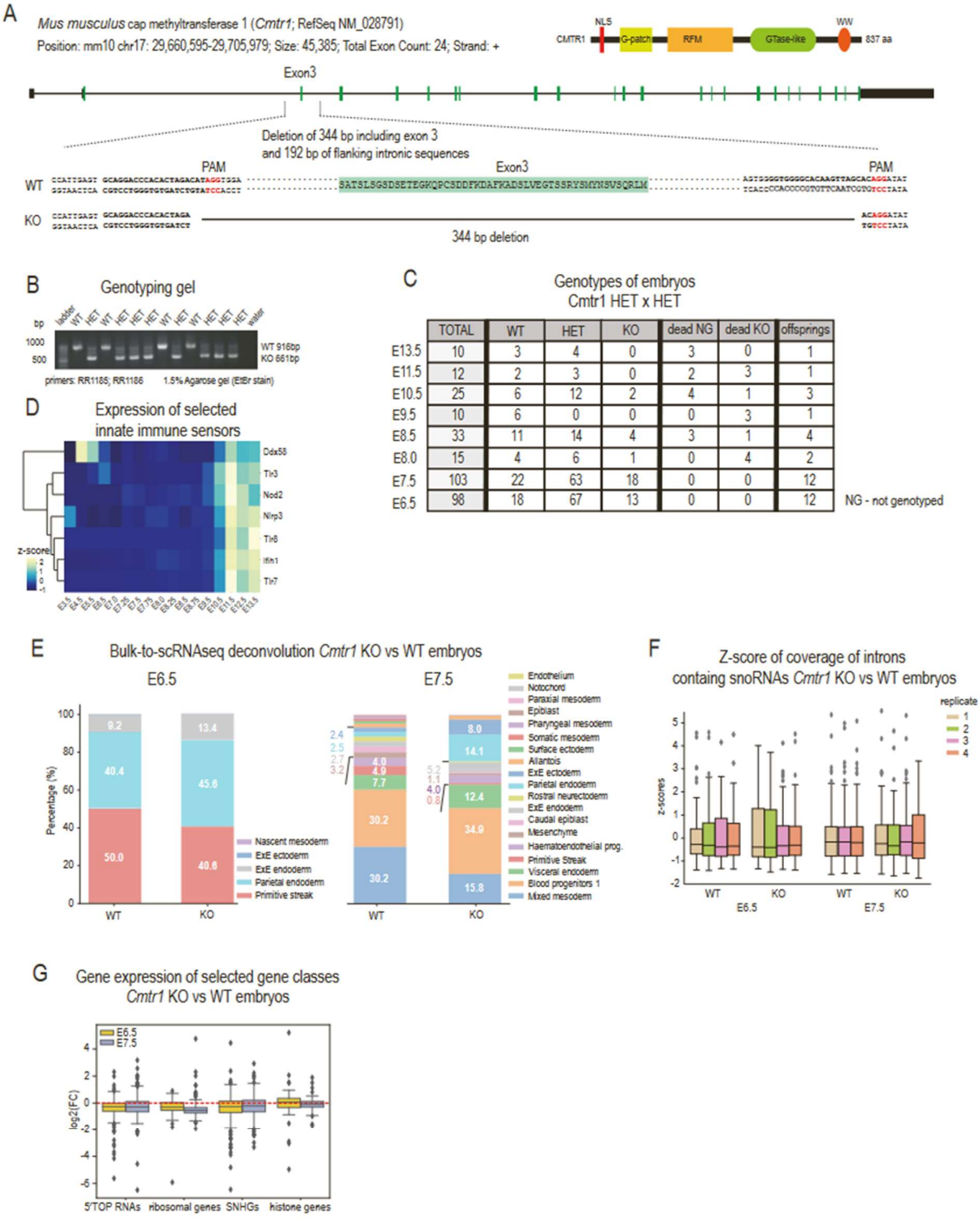
All the bulk-to-single-cell deconvolutions of the E6.5 and E7.5 *Cmtr1* and *Cmtr2* knock-out samples were performed with SCDC (v0.0.0.9000) R package[74] on the published single-cell embryonic atlases.[81]

Cell Reports, Volume *42*

Supplemental information

Essential roles of RNA cap-proximal ribose methylation in mammalian embryonic development and fertility

Michaela Dohnalkova, Kyrylo Krasnykov, Mateusz Mendel, Lingyun Li, Olesya Panasenko, Fabienne Fleury-Olela, Cathrine Broberg Vågbø, David Homolka, and Ramesh S. Pillai



	TOTAL	WT	HET	KO	dead NG	dead KO	offsprings
E13.5	10	3	4	0	3	0	1
E11.5	12	2	3	0	2	3	1
E10.5	25	6	12	2	4	1	3
E9.5	10	6	0	0	0	3	1
E8.5	33	11	14	4	3	1	4
E8.0	15	4	6	1	0	4	2
E7.5	103	22	63	18	0	0	12
E6.5	98	18	67	13	0	0	12

NG - not genotyped

Figure-S1

Figure S1. Characterization of the Cmtr1 knockout mouse

A) The domain architecture of the CMTR1 protein is shown to indicate the nuclear localization signal (NLS), the G-patch, the methyltransferase (RFM), the GTPase-like and the WW domains. Genome engineering of the mouse Cmtr1 locus with CRISPR using two guide RNAs. Exon3 is deleted, resulting in a disrupted translation frame. (B) Agarose gel showing typical genotyping PCR strategy using ear punches to identify wildtype (WT) and heterozygous (HET) animals at weaning age (P21, post-natal day 21). (C) Mouse embryos at indicated embryonic days were collected from heterozygous Cmtr1 females crossed with heterozygous males. The number of embryos of the different genotypes, including homozygous Cmtr1 knockout (KO) are shown. These are aggregated numbers from several offsprings (numbers shown in the last column). We identified several embryos that were clearly dead or were clearly 2 days delayed in the development that are degenerated (absorbed) and not genotyped (NG) or identified as KO. (D) Heatmap showing expression of selected innate immune sensors in mouse embryos at different developmental stages (using published sequencing data, see STAR Methods). (E) Deconvolution of cell type compositions from the bulk transcriptomics datasets of wildtype and Cmtr1 KO E6.5 and E7.5 embryos. The numbers on the charts correspond to a percentage of a specific cell group from total. (F) Z scores of the log2-transformed read counts for snoRNA-containing introns in different snoRNA host genes. Comparison is made for mouse *Cmtr1* KO vs WT embryos at E6.5 and E7.5 stages. (G) Log2 fold changes of gene expression levels for indicated gene classes in mouse Cmtr1 KO mouse embryos compared to wildtype controls at E6.5 and E7.5. Ribo genes, ribosomal protein genes; SNHG, snoRNA host genes; 5′TOP, 5′ terminal oligopyrimidine mRNAs. The numbers in brackets show how many genes from the list per each category were expressed in the mouse embryos.

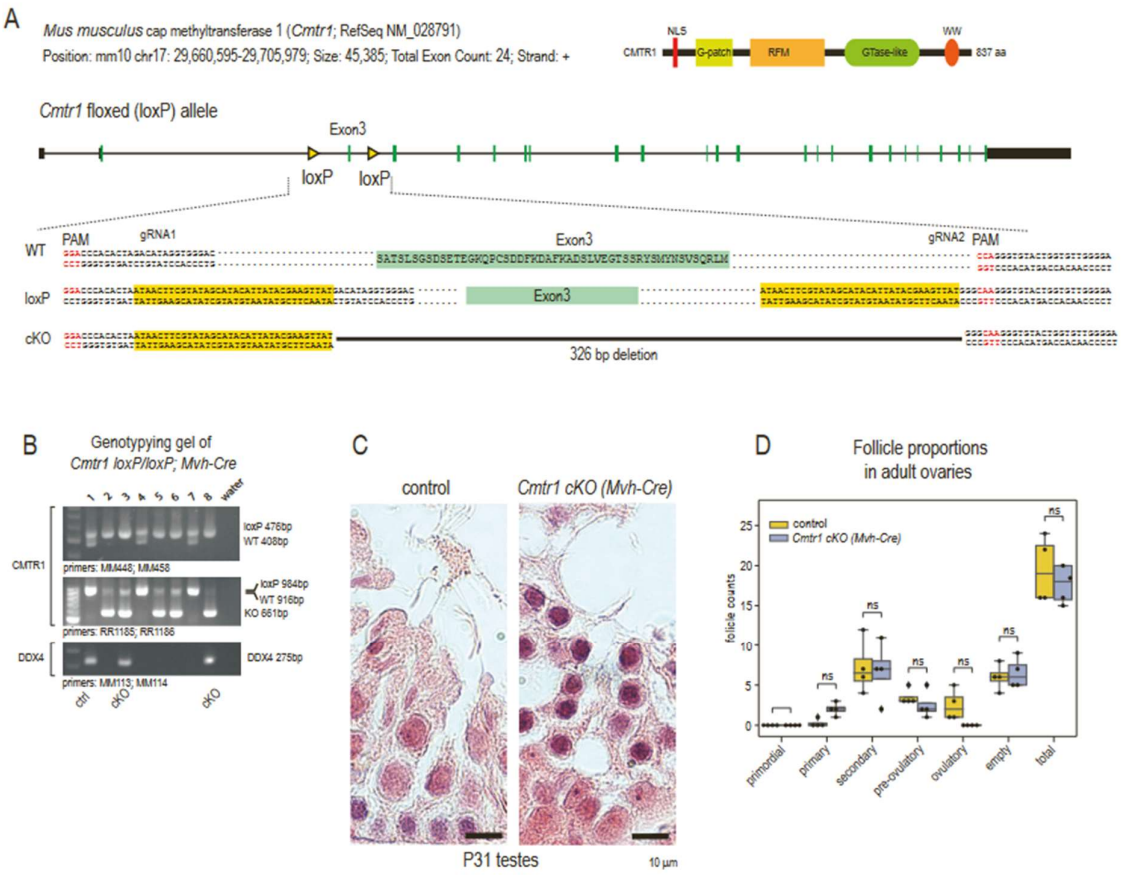

A
Mus musculus cap methyltransferase 1 (Cmtr1; RefSeq NM_028791)
Position: mm10 chr17: 29,660,595-29,705,979; Size: 45,385; Total Exon Count: 24; Strand: +
NLS
CMTR1
G-patch
RFM
GTase-like
WW
837 aa
Cmtr1 floxed (loxP) allele
Exon3
loxP
loxP
WT
PAM
gRNA1
Exon3
gRNA2
PAM
loxP
Exon3
cKO
326 bp deletion
B
Genotypying gel of
Cmtr1 loxP/loxP; Mvh-Cre
1 2 3 4 5 6 7 8 water
CMTR1
loxP 476bp
WT 408bp
primers: MM448; MM458
loxP 984bp
WT 916bp
KO 661bp
primers: RR1185; RR1186
DDX4
DDX4 275bp
primers: MM113; MM114
ctrl
cKO
cKO
C
control
Cmtr1 cKO (Mvh-Cre)
P31 testes
10 µm
D
Follicle proportions
in adult ovaries
control
Cmtr1 cKO (Mvh-Cre)
follicle counts
ns
primordial
primary
secondary
pre-ovulatory
ovulatory
empty
total

Figure-S2

Figure S2. Generation of the conditional *Cmtr1* knockout mouse

A) The domain architecture of the CMTR1 protein is shown to indicate the nuclear localization signal (NLS), the G-patch, the methyltransferase (RFM), the GTPase-like and the WW domains. Genome engineering of the mouse Cmtr1 locus with CRISPR using two guide RNAs to insert two loxP sites in the same direction, flanking the coding exon3. Deletion of the exon results in disrupted translation frame. (B) Agarose gel showing typical genotyping PCR strategy using ear punches to identify wildtype (WT), floxed (loxP) and deleted knockout (KO) alleles of *Cmtr1*, and the *Mvh- (Ddx4)-Cre* transgene at weaning age (P21, post-natal day 21). C) Histological analysis using haematoxylin and eosin staining of mouse testes (P31, postnatal day 31) from control and Cmtr1 cKO. (D) Box plot comparing the number of follicles in different developmental stages observed in histological sections of ovaries from adult (P60-90) control and *Cmtr1* cKO females. Empty: no nucleus in follicle identified, might indicate pre-ovulatory as well as ovulatory follicle. Mann-Whitney test was used to assess differences.

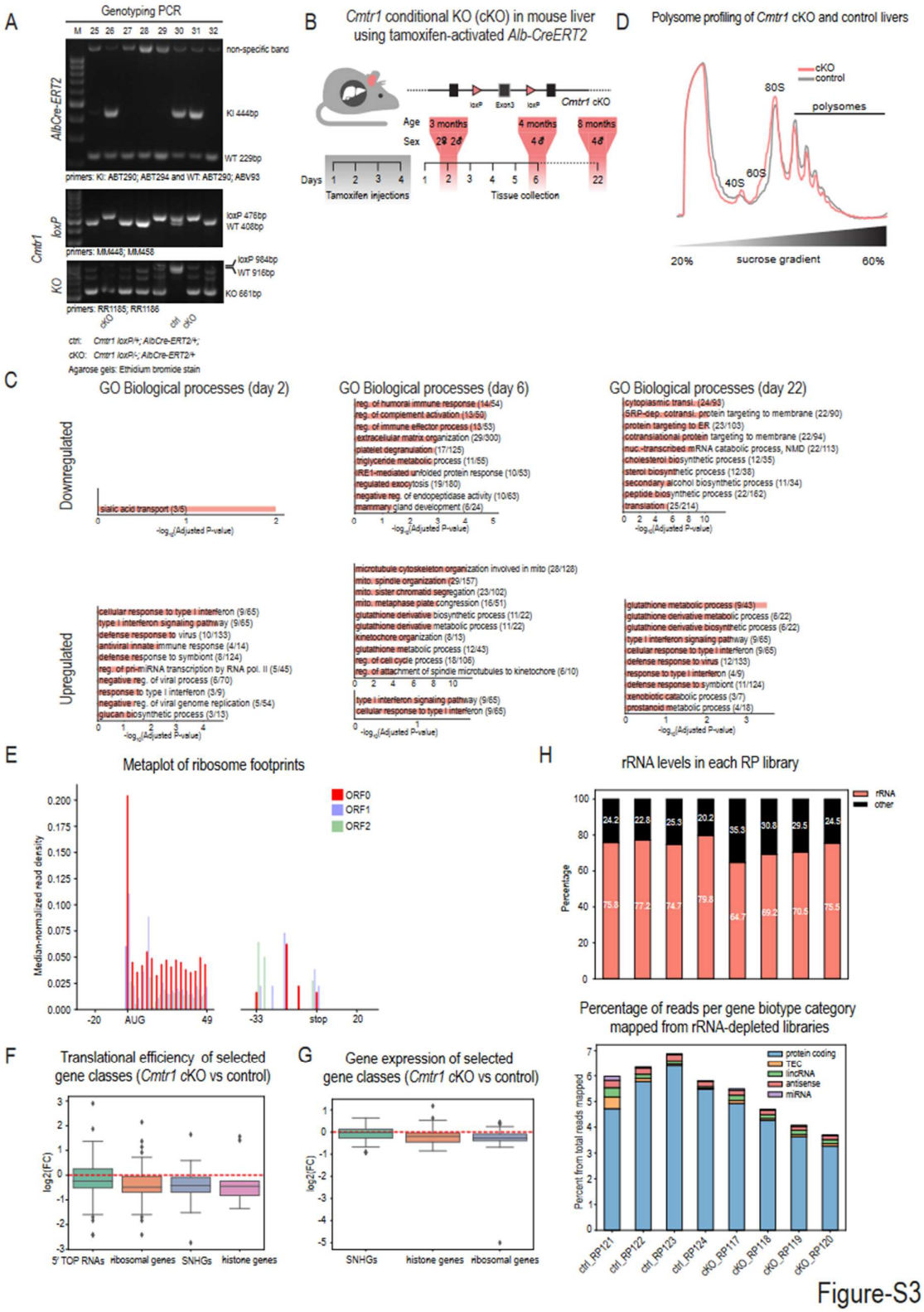

A
Genotyping PCR
M 25 26 27 28 29 30 31 32
non-specific band
Alb Cre-ERT2
KI 444bp
WT 229bp
primers: KI: ABT290; ABT294 and WT: ABT290; ABV93
Cmtr1 loxP
loxP 476bp
WT 408bp
primers: MM448; MM458
Cmtr1 KO
loxP 984bp
WT 916bp
KO 661bp
primers: RR1185; RR1186
ctrl cKO
ctrl: Cmtr1 loxP/+; Alb Cre-ERT2/+;
cKO: Cmtr1 loxP/-; Alb Cre-ERT2/+
Agarose gels: Ethidium bromide stain
B
Cmtr1 conditional KO (cKO) in mouse liver using tamoxifen-activated Alb-CreERT2
loxP Exon3 loxP
Cmtr1 cKO
Age 3 months 4 months 8 months
Sex 2♀ 2♂ 4♂ 4♂
Days 1 2 3 4 1 2 3 4 5 6 22
Tamoxifen injections
Tissue collection
D
Polysome profiling of Cmtr1 cKO and control livers
cKO
control
80S
polysomes
60S
40S
20%
sucrose gradient
60%
C
GO Biological processes (day 2)
Downregulated
sialic acid transport (3/5)
-log10(Adjusted P-value)
GO Biological processes (day 6)
reg. of humoral immune response (14/54)
reg. of complement activation (13/50)
reg. of immune effector process (13/53)
extracellular matrix organization (29/300)
platelet degranulation (17/125)
triglyceride metabolic process (11/55)
IRE1-mediated unfolded protein response (10/53)
regulated exocytosis (19/180)
negative reg. of endopeptidase activity (10/63)
mammary gland development (6/24)
-log10(Adjusted P-value)
GO Biological processes (day 22)
cytoplasmic transl. (24/93)
SRP-dep. cotransl. protein targeting to membrane (22/90)
protein targeting to ER (23/103)
cotranslational protein targeting to membrane (22/94)
nuc.-transcribed mRNA catabolic process, NMD (22/113)
cholesterol biosynthetic process (12/35)
sterol biosynthetic process (12/38)
secondary alcohol biosynthetic process (11/34)
peptide biosynthetic process (22/162)
translation (25/214)
-log10(Adjusted P-value)
Upregulated
cellular response to type I interferon (9/65)
type I interferon signaling pathway (9/65)
defense response to virus (10/133)
antiviral innate immune response (4/14)
defense response to symbiont (8/124)
reg. of pri-miRNA transcription by RNA pol. II (5/45)
negative reg. of viral process (6/70)
response to type I interferon (3/9)
negative reg. of viral genome replication (5/54)
glucan biosynthetic process (3/13)
-log10(Adjusted P-value)
microtubule cytoskeleton organization involved in mito (28/128)
mito. spindle organization (29/157)
mito. sister chromatid segregation (23/102)
mito. metaphase plate congression (18/51)
glutathione derivative biosynthetic process (11/22)
glutathione derivative metabolic process (11/22)
kinetochore organization (8/13)
glutathione metabolic process (12/43)
reg. of cell cycle process (18/106)
reg. of attachment of spindle microtubules to kinetochore (6/10)
type I interferon signaling pathway (9/65)
cellular response to type I interferon (9/65)
-log10(Adjusted P-value)
glutathione metabolic process (9/43)
glutathione derivative metabolic process (6/22)
glutathione derivative biosynthetic process (6/22)
type I interferon signaling pathway (9/65)
cellular response to type I interferon (9/65)
defense response to virus (12/133)
response to type I interferon (4/9)
defense response to symbiont (11/124)
xenobiotic catabolic process (3/7)
prostanoid metabolic process (4/18)
-log10(Adjusted P-value)
E
Metaplot of ribosome footprints
ORF0
ORF1
ORF2
Median-normalized read density
-20 AUG 49
-33 stop 20
H
rRNA levels in each RP library
rRNA
other
Percentage
24.2 22.8 25.3 20.2 35.3 30.8 29.5 24.5
75.8 77.2 74.7 79.8 64.7 69.2 70.5 75.5
Percentage of reads per gene biotype category mapped from rRNA-depleted libraries
protein coding
TEC
lincRNA
antisense
miRNA
Percent from total reads mapped
ctrl_RP121 ctrl_RP122 ctrl_RP123 ctrl_RP124 cKO_RP117 cKO_RP118 cKO_RP119 cKO_RP120
F
Translational efficiency of selected gene classes (Cmtr1 cKO vs control)
log2(FC)
5' TOP RNAs ribosomal genes SNHGs histone genes
G
Gene expression of selected gene classes (Cmtr1 cKO vs control)
log2(FC)
SNHGs histone genes ribosomal genes

Figure-S3

Figure S3. Analysis of translation in the *Cmtr1* conditional knockout mouse liver

(A) Agarose gel showing typical genotyping PCR strategy using ear punches to identify wildtype (WT), floxed (loxP) and conditional knockout (cKO) alleles of Cmtr1, and the Alb-CreERT2 transgene at weaning age (P21, post-natal day 21). (B) Conditional deletion of mouse Cmtr1 using a floxed allele with loxP sites flanking the coding exon3. (C) GO term analyses of the second batch of mouse control and Cmtr1 cKO liver RNA-seq. (D) Polysome analysis with sucrose density gradient (linear 20-60%) centrifugation of liver lysates from control and the Cmtr1 cKO adult mouse. Positions of 40S and 60S subunits, 80S monosomes and the polysome peaks are indicated. (E) Metaplot of the normalized (rpm) 5′ ends of ribosome footprints (RFP) within a correspondent ORF around start and stop codons with indicated up- and downstream distances. Notice the preference for one (red, ORF0) of the three reading frames. (F) Log2 fold changes of translation efficiency of genes belonging to the indicated classes in the Cmtr1 cKO mouse liver compared to the control. SNHGs, snoRNA host genes; 5′TOP RNAs, genes encoding 5′ terminal oligopyrimidine mRNAs. (G) Log2 fold changes of the expression changes between Cmtr1 cKO mouse liver and control in indicated gene classes. (H) Ribosome profiling quality control plots. Stacked bar charts showing percentage of rRNAs per library (upper) or percentage of reads mapped to mRNAs and several other gene classes (lower).

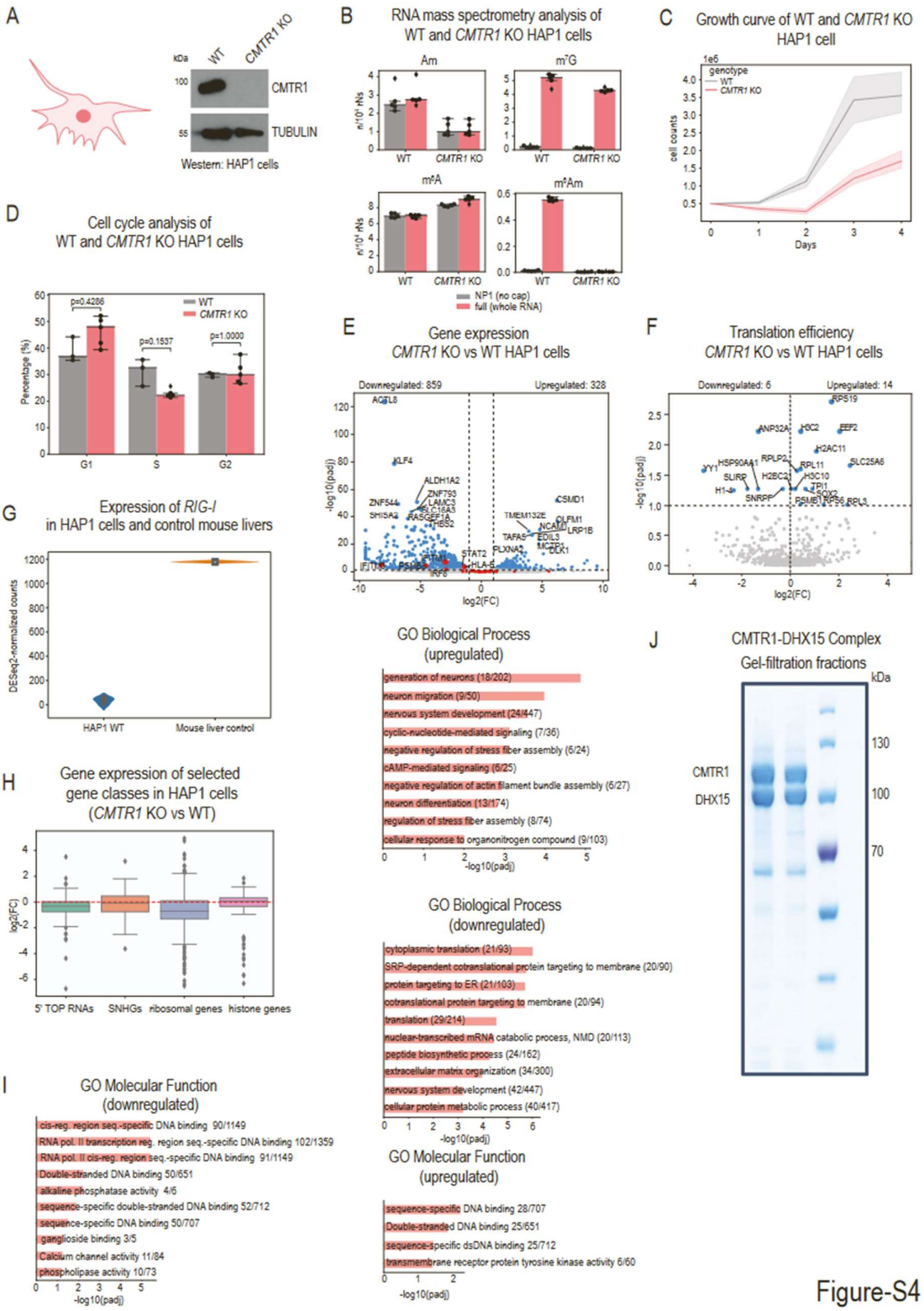

A
kDa
WT
CMTR1 KO
100
CMTR1
55
TUBULIN
Western: HAP1 cells
B
RNA mass spectrometry analysis of
WT and CMTR1 KO HAP1 cells
Am
m7G
m6A
m6Am
n/10^6 nNs
WT
CMTR1 KO
NP1 (no cap)
full (whole RNA)
C
Growth curve of WT and CMTR1 KO
HAP1 cell
1e6
genotype
WT
CMTR1 KO
cell counts
Days
D
Cell cycle analysis of
WT and CMTR1 KO HAP1 cells
p=0.4286
p=0.1537
p=1.0000
Percentage (%)
G1
S
G2
E
Gene expression
CMTR1 KO vs WT HAP1 cells
Downregulated: 859
Upregulated: 328
-log10(padj)
log2(FC)
F
Translation efficiency
CMTR1 KO vs WT HAP1 cells
Downregulated: 6
Upregulated: 14
G
Expression of RIG-I
in HAP1 cells and control mouse livers
DESeq2-normalized counts
HAP1 WT
Mouse liver control
H
Gene expression of selected
gene classes in HAP1 cells
(CMTR1 KO vs WT)
log2(FC)
5' TOP RNAs
SNHGs
ribosomal genes
histone genes
GO Biological Process
(upregulated)
generation of neurons (18/202)
neuron migration (9/50)
nervous system development (24/447)
cyclic-nucleotide-mediated signaling (7/36)
negative regulation of stress fiber assembly (6/24)
cAMP-mediated signaling (6/25)
negative regulation of actin filament bundle assembly (6/27)
neuron differentiation (13/174)
regulation of stress fiber assembly (8/74)
cellular response to organonitrogen compound (9/103)
-log10(padj)
GO Biological Process
(downregulated)
cytoplasmic translation (21/93)
SRP-dependent cotranslational protein targeting to membrane (20/90)
protein targeting to ER (21/103)
cotranslational protein targeting to membrane (20/94)
translation (29/214)
nuclear-transcribed mRNA catabolic process, NMD (20/113)
peptide biosynthetic process (24/162)
extracellular matrix organization (34/300)
nervous system development (42/447)
cellular protein metabolic process (40/417)
-log10(padj)
J
CMTR1-DHX15 Complex
Gel-filtration fractions
kDa
CMTR1
DHX15
130
100
70
I
GO Molecular Function
(downregulated)
cis-reg. region seq.-specific DNA binding 90/1149
RNA pol. II transcription reg. region seq.-specific DNA binding 102/1359
RNA pol. II cis-reg. region seq.-specific DNA binding 91/1149
Double-stranded DNA binding 50/651
alkaline phosphatase activity 4/6
sequence-specific double-stranded DNA binding 52/712
sequence-specific DNA binding 50/707
ganglioside binding 3/5
Calcium channel activity 11/84
phospholipase activity 10/73
-log10(padj)
GO Molecular Function
(upregulated)
sequence-specific DNA binding 28/707
Double-stranded DNA binding 25/651
sequence-specific dsDNA binding 25/712
transmembrane receptor protein tyrosine kinase activity 6/60
-log10(padj)

Figure-S4

Figure S4. CMTR1 is not essential in human HAP1 cells and does not regulate translation

A) Western analysis in wildtype (WT) and CMTR1 knockout (KO) human HAP1 cells. (B) PolyA+ RNA from CMTR1 WT and KO HAP1 cells was subjected to mass spectrometry. Bar plots combined with box plots show abundance (number of modified nucleotides/104 nucleotides) of some of the modifications (Am, m7G, m6A, m6Am). Quintuplicate biological replicates were tested. (C) Cell growth curve analysis for the WT and CMTR1 KO HAP1 cells. Curves of average cell counts of 5 WT and 5 CMTR1 KO replicates with 95% confidence interval are shown. (D) Cell cycle analysis of the CMTR1 WT and KO HAP1 cells. Whiskers show the minimum and maximum values. Mann-Whitney statistical test was used to assess differences. (E) Volcano plot of differentially expressed genes between KO and WT HAP1 cells. Type I IFN response genes are highlighted in red. Absolute log2 fold change (log2FC) cutoff = 1, adjusted p-value (padj) cutoff = 0.05. (F) Volcano plot of the differential translational efficiency between WT and KO HAP1 cells. Adjusted p-value (padj) cutoff = 0.1. (G) Violin plot of RIG-I expression levels from transcriptome sequencing data of wildtype (WT) HAP1 cells and control mouse livers. (H) Log2 fold expression changes for indicated gene classes in CMTR1 KO HAP1 cells compared to the control HAP1 cells. 5′TOP RNAs, 5′ terminal oligopyrimidine; SNHGs, snoRNA host genes. (I) Gene ontology analysis of genes dysregulated in the CMTR1 KO HAP1 cells compared to the control HAP1 cells. (J) Purification of a recombinant complex of human CMTR1 and DHX15. Two fractions obtained after size-exclusion chromatography were resolved by SDS-PAGE and stained with Coomassie blue stain. Protein markers in kilo Daltons (kDa) are indicated.

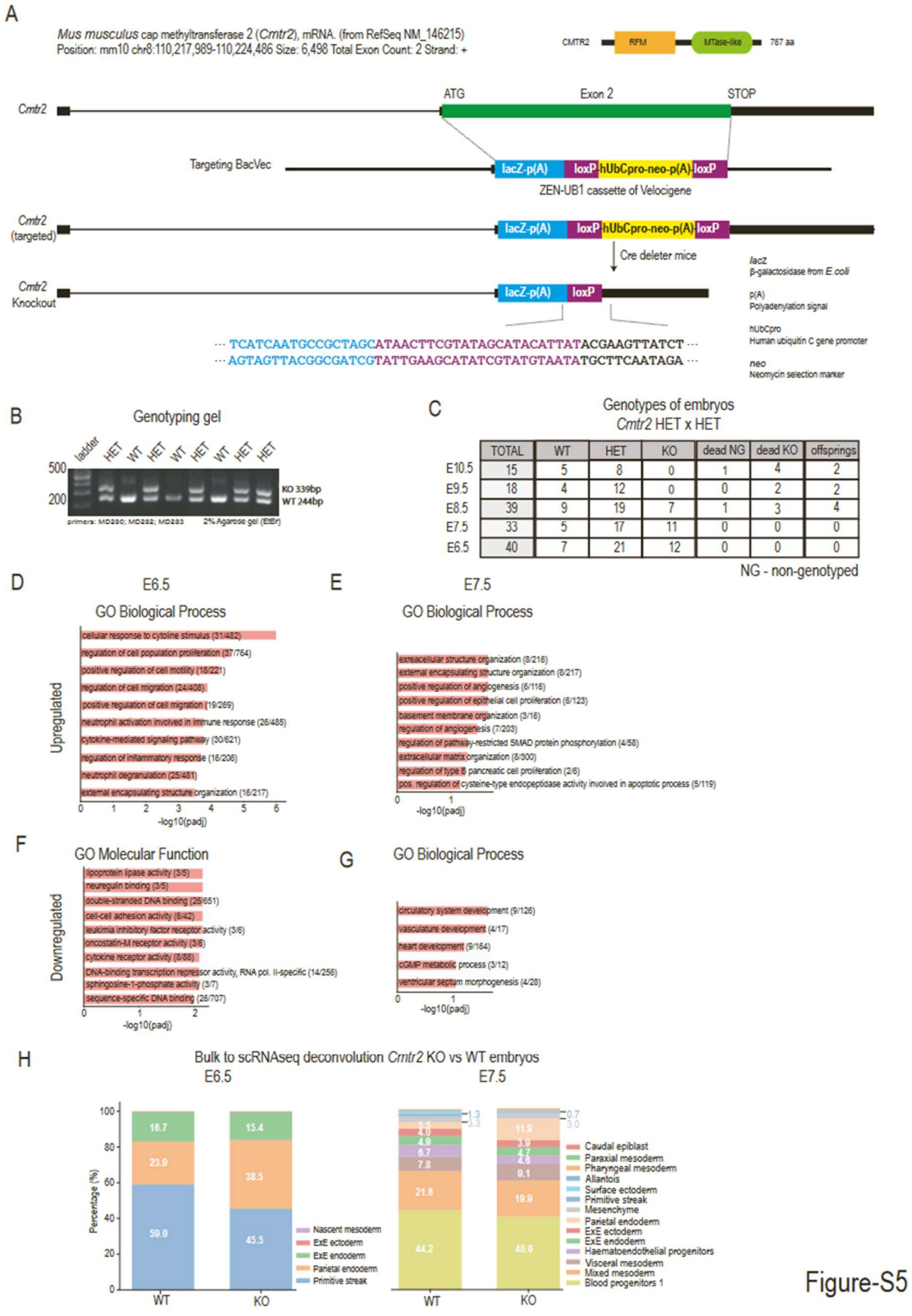

A
Mus musculus cap methyltransferase 2 (Cmtr2), mRNA. (from RefSeq NM_146215)
Position: mm10 chr8:110,217,989-110,224,486 Size: 6,498 Total Exon Count: 2 Strand: +
CMTR2
RFM
MTase-like
767 aa
Cmtr2
ATG
Exon 2
STOP
Targeting BacVec
lacZ-p(A)
loxP
hUbCpro-neo-p(A)
loxP
ZEN-UB1 cassette of Velocigene
Cmtr2 (targeted)
Cre deleter mice
Cmtr2 Knockout
lacZ
β-galactosidase from E.coli
p(A)
Polyadenylation signal
hUbCpro
Human ubiquitin C gene promoter
neo
Neomycin selection marker
··· TCATCAATGCCGCTAGCATAACTTCGTATAGCATACATTATACGAAGTTATCT ···
··· AGTAGTTACGGCGATCGTATTGAAGCATATCGTATGTAATATGCTTCAATAGA ···
B
Genotyping gel
ladder
HET
WT
HET
WT
HET
WT
HET
HET
500
200
KO 339bp
WT 244bp
primers: MD280; MD282; MD283
2% Agarose gel (EtBr)
C
Genotypes of embryos
Cmtr2 HET x HET
TOTAL
WT
HET
KO
dead NG
dead KO
offsprings
E10.5 15 5 8 0 1 4 2
E9.5 18 4 12 0 0 2 2
E8.5 39 9 19 7 1 3 4
E7.5 33 5 17 11 0 0 0
E6.5 40 7 21 12 0 0 0
NG - non-genotyped
D
E6.5
GO Biological Process
Upregulated
cellular response to cytoline stimulus (31/482)
regulation of cell population proliferation (37/764)
positive regulation of cell motility (18/221)
regulation of cell migration (24/408)
positive regulation of cell migration (19/269)
neutrophil activation involved in immune response (26/485)
cytokine-mediated signaling pathway (30/621)
regulation of inflammatory response (16/206)
neutrophil degranulation (25/481)
external encapsulating structure organization (16/217)
-log10(padj)
E
E7.5
GO Biological Process
exreacellular structure organization (8/216)
external encapsulating structure organization (8/217)
positive regulation of angiogenesis (6/116)
positive regulation of epithelial cell proliferation (6/123)
basement membrane organization (3/16)
regulation of angiogenesis (7/203)
regulation of pathway-restricted SMAD protein phosphorylation (4/58)
extracellular matrix organization (8/300)
regulation of type B pancreatic cell proliferation (2/6)
pos. regulation of cysteine-type endopeptidase activity involved in apoptotic process (5/119)
-log10(padj)
F
GO Molecular Function
Downregulated
lipoprotein lipase activity (3/5)
neuregulin binding (3/5)
double-stranded DNA binding (28/651)
cell-cell adhesion activity (6/42)
leukimia inhibitory factor receptor activity (3/6)
oncostatin-M receptor activity (3/6)
cytokine receptor activity (8/88)
DNA-binding transcription repressor activity, RNA pol. II-specific (14/256)
sphingosine-1-phosphate activity (3/7)
sequence-specific DNA binding (26/707)
-log10(padj)
G
GO Biological Process
circulatory system development (9/126)
vasculature development (4/17)
heart development (9/164)
cGMP metabolic process (3/12)
ventricular septum morphogenesis (4/28)
-log10(padj)
H
Bulk to scRNAseq deconvolution Cmtr2 KO vs WT embryos
E6.5
E7.5
Percentage (%)
16.7
23.9
59.0
15.4
38.5
45.5
WT
KO
Nascent mesoderm
ExE ectoderm
ExE endoderm
Parietal endoderm
Primitive streak
3.5
4.0
4.9
6.7
7.8
21.8
44.2
1.3
3.3
11.9
3.9
4.7
4.6
9.1
19.9
40.0
0.7
3.0
Caudal epiblast
Paraxial mesoderm
Pharyngeal mesoderm
Allantois
Surface ectoderm
Primitive streak
Mesenchyme
Parietal endoderm
ExE ectoderm
ExE endoderm
Haematoendothelial progenitors
Visceral mesoderm
Mixed mesoderm
Blood progenitors 1

Figure-S5

Figure S5. Characterization of the Cmtr2 knockout mouse

A) The domain architecture of the CMTR2 protein is shown to indicate the methyltransferase (RFM) and the methyltransferase-like (MTase-like) domains. Genome editing strategy removes the entire coding sequence of mouse Cmtr2 and replaces it with the lacZ reporter. (B) Agarose gel showing typical genotyping PCR strategy using ear punches to identify wildtype (WT), heterozygous (HET) and knockout (KO) alleles of Cmtr2 at weaning age (P21, post-natal day 21). (C) Genotypes of mouse embryos (shown as percentage) at indicated embryonic (E) days. The number of embryos of the different genotypes are shown. These are aggregated numbers from several offsprings (numbers shown in the last column). We identified several embryos that were clearly dead/degenerated and not genotyped (NG) or identified as KO. (D-G) Gene ontology analysis of gene expression changes in the Cmtr2 knockout mouse embryos compared to the wildtype controls. (H) Deconvolution of cell type compositions from the bulk transcriptomics datasets of wildtype and Cmtr2 KO E6.5 and E7.5 embryos. The numbers on the charts correspond to a percentage of a specific cell group from total.

Table S1: List of all oligonucleotides used in this study.		
Lab name	Sequence	Comments
Cmtr1 conditional KO mouse		
MM-RNA12	/AlTR1/rGrUrCrCrCrArCrCrUrArUrGrUrCrUrArGrUrGrU/AlTR2/	Alt-R® CRISPR-Cas9 crRNA ordered from IDT
MM-RNA13	/AlTR1/rUrCrCrCrCrArArCrArCrCrArGrUrArCrArCrCrC/AlTR2/	Alt-R® CRISPR-Cas9 crRNA ordered from IDT
MM446	ATACGTACGTATACAGCTGGCAAGAGTAGAGACGTCACTGTGACCTCCATTGAGTGCAG GACCCACACTAATAACTTCGTATAGCATACATTATACGAAGTTATGACATAGGTGGGAC ATGTGGACTGTGGGTGCATGAGGCAGTCCTGTCATCCGGACCCACCTAACGCTTCTCTT CTTCTTCCCCAGCATCTGCTACAAGCCTCAGTGGATCTGACAGTGAGACCGAGGGGAAG CAGCCCTGCTCTGATGATTTCAAAGATGCCTTCAAAGCAGATTCCCTTGTGGAGGGAAC ATCGTCCCGATATTCCATGTATAACAGTGTTTCCCAGAGGCTTATGGTATGTCTTGGCT TAGAATGGACTTCTAAAGTTGCCCAAAAGAGGGAGAGGAAGAATAACTTCGTATAGCAT ACATTATACGAAGTTATGGGCAAGGGTGTACTGGTGTTGGGGAGTGGGGTGGGGCACAA GTTAGCACAGGATATAGGTTCTGAGTAT	ssDNA repair template for introducing two loxP sites flanking exon 3 of Cmtr1
Genotyping		
RR1185	CGTTCCTGTTCTGTGGCTTG	Cmtr1 KO genotyping-forward. Gives different sizes with WT and KO alleles
RR1186	CATGTCTGAGGAGGCTGCT	Cmtr1 KO genotyping-reverse. Gives different sizes with WT and KO alleles
MM458	ATGTGTGCGTTCCTGTTCTG	Cmtr1 floxed allele genotyping-forward. Gives different sizes with WT and KI alleles
MM448	CCTATATCCTGTGCTAACTTGTGC	Cmtr1 floxed allele genotyping-reverse. Gives different sizes with WT and KI alleles
MD280	CCTCCCACACCTCCCCCTGAACCTGAAAC	Cmtr2 KO genotyping-KI forward for the LacZ insert in the KO locus
MD282	CTACAGAATGTTCATGACCTCTTGAGCAAGC	Cmtr2 KO genotyping-WT forward for the WT Cmtr2 sequence
MD283	GTCACAAAACGGCTGCACGTCATCAATGCCG	Cmtr2 KO genotyping-Common reverse primer downstream on Cmtr2 locus
ABT290	ATCATTTCTTTGTTTTCAGG	Alb-Cre-ERT2 genotyping-common forward gives different sizes with WT and KI alleles
ABT294	TTAAACAAGCAAAACCAAAT	Alb-Cre-ERT2 genotyping-KI allele reverse
ABV93	GGAACCCAAACTGATGACCA	Alb-Cre-ERT2 genotyping-WT allele reverse
MM113	CACGTGCAGCCGTTTAAGCCGCGT	Mvh-Cre genotyping-forward to detect transgene
MM114	TTCCCATTCTAAACAACACCCTGAA	Mvh-Cre genotyping-reverse to detect transgene
Ribosome profiling		
OP-RNA25	AUGUACACGGAGUCGAGCACCCGCA	25-nt RNA marker
OP-RNA34	AUGUACACGGAGUCGAGCACCCGCAACGCGAAUG	34-nt RNA marker
OP-RNA45	5' rAppCTGTAGGCACCATCAAT/3ddC/ 3'	Adapter Linker-1
OP27	AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGTGGTCGC5CACTCA5TT CAGACGTGTGCTCTTCCGATCTATTGATGGTGCCTACAG	RT primer Ni-Ni-9
OP28	AATGATACGGCGACCACCGAGATCTACAC	forward index primer NI-N-2

Table S2: All deep sequencing data generated for this study.				
experiment	sample	genotype	reads	Comments
RNA-seq of E6.5 embryos (PE)	MD_KO1	Cmtr1-/-	64165347	
RNA-seq of E6.5 embryos (PE)	MD_KO2	Cmtr1-/-	73633676	
RNA-seq of E6.5 embryos (PE)	MD_KO3	Cmtr1-/-	47339774	
RNA-seq of E6.5 embryos (PE)	MD_KO4	Cmtr1-/-	62758462	
RNA-seq of E6.5 embryos (PE)	MD_WT1	Cmtr1+/+	95967391	
RNA-seq of E6.5 embryos (PE)	MD_WT2	Cmtr1+/+	66340526	
RNA-seq of E6.5 embryos (PE)	MD_WT3	Cmtr1+/+	65704230	
RNA-seq of E6.5 embryos (PE)	MD_WT4	Cmtr1+/+	46737428	
RNA-seq of E7.5 embryos (PE)	MD_KO1	Cmtr1-/-	117885403	
RNA-seq of E7.5 embryos (PE)	MD_KO2	Cmtr1-/-	60211089	
RNA-seq of E7.5 embryos (PE)	MD_KO3	Cmtr1-/-	108585136	
RNA-seq of E7.5 embryos (PE)	MD_KO4	Cmtr1-/-	57474519	
RNA-seq of E7.5 embryos (PE)	MD_WT1	Cmtr1+/+	84283512	
RNA-seq of E7.5 embryos (PE)	MD_WT2	Cmtr1+/+	180370	
RNA-seq of E7.5 embryos (PE)	MD_WT3	Cmtr1+/+	71984647	
RNA-seq of E7.5 embryos (PE)	MD_WT4	Cmtr1+/+	83710007	
RNA-seq of mouse livers (PE)	MD_ctrl1_D2_b1	Cmtr1+/loxP; AlbCreERT2+/-	101020009	Day2 expt-1stbatch
RNA-seq of mouse livers (PE)	MD_ctrl2_D2_b1	Cmtr1+/loxP; AlbCreERT2+/-	72616359	Day2 expt-1stbatch
RNA-seq of mouse livers (PE)	MD_ctrl3_D2_b1	Cmtr1+/loxP; AlbCreERT2+/-	81799463	Day2 expt-1stbatch
RNA-seq of mouse livers (PE)	MD_cKO1_D2_b1	Cmtr1-/loxP; AlbCreERT2+/-	85041035	Day2 expt-1stbatch
RNA-seq of mouse livers (PE)	MD_cKO2_D2_b1	Cmtr1-/loxP; AlbCreERT2+/-	102687055	Day2 expt-1stbatch
RNA-seq of mouse livers (PE)	MD_cKO3_D2_b1	Cmtr1-/loxP; AlbCreERT2+/-	74711285	Day2 expt-1stbatch
RNA-seq of mouse livers (PE)	MD_cKO4_D2_b1	Cmtr1-/loxP; AlbCreERT2+/-	65837771	Day2 expt-1stbatch
RNA-seq of mouse livers (PE)	MiDo21_ctrl_D2_b2	Cmtr1+/loxP; AlbCreERT2+/-	79064254	Day2 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo22_ctrl_D2_b2	Cmtr1+/loxP; AlbCreERT2+/-	72578235	Day2 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo25_ctrl_D2_b2	Cmtr1+/loxP; AlbCreERT2+/-	101321669	Day2 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo26_ctrl_D2_b2	Cmtr1+/loxP; AlbCreERT2+/-	126006371	Day2 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo17_cKO_D2_b2	Cmtr1-/loxP; AlbCreERT2+/-	94450730	Day2 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo18_cKO_D2_b2	Cmtr1-/loxP; AlbCreERT2+/-	117914647	Day2 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo29_cKO_D2_b2	Cmtr1-/loxP; AlbCreERT2+/-	113886773	Day2 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo5_ctrl_D6_b2	Cmtr1+/loxP; AlbCreERT2+/-	97330461	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo6_ctrl_D6_b2	Cmtr1+/loxP; AlbCreERT2+/-	84329985	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo7_ctrl_D6_b2	Cmtr1+/loxP; AlbCreERT2+/-	63683694	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo8_ctrl_D6_b2	Cmtr1+/loxP; AlbCreERT2+/-	86025348	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo12_cKO_D6_b2	Cmtr1-/loxP; AlbCreERT2+/-	71044210	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo13_cKO_D6_b2	Cmtr1-/loxP; AlbCreERT2+/-	74273102	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo1_cKO_D6_b2	Cmtr1-/loxP; AlbCreERT2+/-	92403609	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo27_cKO_D6_b2	Cmtr1-/loxP; AlbCreERT2+/-	102243306	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo28_cKO_D6_b2	Cmtr1-/loxP; AlbCreERT2+/-	83858065	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo2_cKO_D6_b2	Cmtr1-/loxP; AlbCreERT2+/-	90041160	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo3_cKO_D6_b2	Cmtr1-/loxP; AlbCreERT2+/-	95788724	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo4_cKO_D6_b2	Cmtr1-/loxP; AlbCreERT2+/-	81888780	Day6 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo30_ctrl_D22_b2	Cmtr1+/loxP; AlbCreERT2+/-	86726077	Day22 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo14_ctrl_D22_b2	Cmtr1+/loxP; AlbCreERT2+/-	81706754	Day22 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo15_ctrl_D22_b2	Cmtr1+/loxP; AlbCreERT2+/-	94614595	Day22 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDi16_ctrl_D22_b2	Cmtr1+/loxP; AlbCreERT2+/-	69806471	Day22 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo10_cKO_D22_b2	Cmtr1-/loxP; AlbCreERT2+/-	82233224	Day22 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo11_cKO_D22_b2	Cmtr1-/loxP; AlbCreERT2+/-	84065377	Day22 expt-2ndbatch
RNA-seq of mouse livers (PE)	MiDo9_cKO_D22_b2	Cmtr1-/loxP; AlbCreERT2+/-	99328927	Day22 expt-2ndbatch
Ribosome profiling of mouse livers RPF (SE)	RP121	Cmtr1+/loxP; AlbCreERT2+/-	31183139	Footprints
Ribosome profiling of mouse livers RPF (SE)	RP122	Cmtr1+/loxP; AlbCreERT2+/-	36647766	Footprints
Ribosome profiling of mouse livers RPF (SE)	RP123	Cmtr1+/loxP; AlbCreERT2+/-	44268414	Footprints
Ribosome profiling of mouse livers RPF (SE)	RP124	Cmtr1+/loxP; AlbCreERT2+/-	78800229	Footprints
Ribosome profiling of mouse livers RPF (SE)	RP117	Cmtr1-/loxP; AlbCreERT2+/-	36283250	Footprints
Ribosome profiling of mouse livers RPF (SE)	RP118	Cmtr1-/loxP; AlbCreERT2+/-	45592718	Footprints
Ribosome profiling of mouse livers RPF (SE)	RP119	Cmtr1-/loxP; AlbCreERT2+/-	48374502	Footprints
Ribosome profiling of mouse livers RPF (SE)	RP120	Cmtr1-/loxP; AlbCreERT2+/-	24978277	Footprints
Ribosome profiling of mouse livers INP (SE)	RP321	Cmtr1+/loxP; AlbCreERT2+/-	33441118	Input
Ribosome profiling of mouse livers INP (SE)	RP322	Cmtr1+/loxP; AlbCreERT2+/-	30591131	Input
Ribosome profiling of mouse livers INP (SE)	RP323	Cmtr1+/loxP; AlbCreERT2+/-	26997713	Input
Ribosome profiling of mouse livers INP (SE)	RP324	Cmtr1+/loxP; AlbCreERT2+/-	32901242	Input
Ribosome profiling of mouse livers INP (SE)	RP317	Cmtr1-/loxP; AlbCreERT2+/-	40793141	Input
Ribosome profiling of mouse livers INP (SE)	RP318	Cmtr1-/loxP; AlbCreERT2+/-	34249233	Input
Ribosome profiling of mouse livers INP (SE)	RP319	Cmtr1-/loxP; AlbCreERT2+/-	24811221	Input
Ribosome profiling of mouse livers INP (SE)	RP320	Cmtr1-/loxP; AlbCreERT2+/-	30547251	Input
RNA-seq of HAP1 cells (PE)	MD_KO1	Cmtr1-/-	73056539	

RNA-seq of HAP1 cells (PE)	MD_KO2	Cmtr1-/-	83455006	
RNA-seq of HAP1 cells (PE)	MD_KO3	Cmtr1-/-	147985070	
RNA-seq of HAP1 cells (PE)	MD_KO4	Cmtr1-/-	82972876	
RNA-seq of HAP1 cells (PE)	MD_WT1	Cmtr1+/+	88417077	
RNA-seq of HAP1 cells (PE)	MD_WT2	Cmtr1+/+	101184569	
RNA-seq of HAP1 cells (PE)	MD_WT3	Cmtr1+/+	96551644	
RNA-seq of HAP1 cells (PE)	MD_WT4	Cmtr1+/+	77639067	
Ribosome profiling of HAP1 cells RPF (SE)	RPF229	Cmtr1-/-	17797304	Footprints
Ribosome profiling of HAP1 cells RPF (SE)	RPF230	Cmtr1-/-	27438285	Footprints
Ribosome profiling of HAP1 cells RPF (SE)	RPF231	Cmtr1-/-	26966840	Footprints
Ribosome profiling of HAP1 cells RPF (SE)	RPF232	Cmtr1-/-	46163108	Footprints
Ribosome profiling of HAP1 cells RPF (SE)	RPF225	Cmtr1+/+	1520195	Footprints
Ribosome profiling of HAP1 cells RPF (SE)	RPF226	Cmtr1+/+	19938583	Footprints
Ribosome profiling of HAP1 cells RPF (SE)	RPF227	Cmtr1+/+	82958026	Footprints
Ribosome profiling of HAP1 cells RPF (SE)	RPF228	Cmtr1+/+	19565670	Footprints
Ribosome profiling of HAP1 cells INP (SE)	RP529	Cmtr1-/-	28747106	Input
Ribosome profiling of HAP1 cells INP (SE)	RP530	Cmtr1-/-	40287010	Input
Ribosome profiling of HAP1 cells INP (SE)	RP531	Cmtr1-/-	32108785	Input
Ribosome profiling of HAP1 cells INP (SE)	RP532	Cmtr1-/-	36902310	Input
Ribosome profiling of HAP1 cells INP (SE)	RP525	Cmtr1+/+	32951952	Input
Ribosome profiling of HAP1 cells INP (SE)	RP526	Cmtr1+/+	35383220	Input
Ribosome profiling of HAP1 cells INP (SE)	RP527	Cmtr1+/+	43924304	Input
Ribosome profiling of HAP1 cells INP (SE)	RP528	Cmtr1+/+	36788561	Input
RNA-seq of E6.5 mouse embryos (SE)	MD_WT1	Cmtr2+/+	33137245	
RNA-seq of E6.5 mouse embryos (SE)	MD_WT2	Cmtr2+/+	21949267	
RNA-seq of E6.5 mouse embryos (SE)	MD_WT3	Cmtr2+/+	19711771	
RNA-seq of E6.5 mouse embryos (SE)	MD_WT4	Cmtr2+/+	19821136	
RNA-seq of E6.5 mouse embryos (SE)	MD_WT5	Cmtr2+/+	24606562	
RNA-seq of E6.5 mouse embryos (SE)	MD_KO1	Cmtr2-/-	16879565	
RNA-seq of E6.5 mouse embryos (SE)	MD_KO2	Cmtr2-/-	23156767	
RNA-seq of E6.5 mouse embryos (SE)	MD_KO3	Cmtr2-/-	24150328	
RNA-seq of E6.5 mouse embryos (SE)	MD_KO4	Cmtr2-/-	24656858	
RNA-seq of E6.5 mouse embryos (SE)	MD_KO5	Cmtr2-/-	22455342	
RNA-seq of E7.5 mouse embryos (SE)	MD_WT1	Cmtr2+/+	22071322	
RNA-seq of E7.5 mouse embryos (SE)	MD_WT2	Cmtr2+/+	22713400	
RNA-seq of E7.5 mouse embryos (SE)	MD_WT3	Cmtr2+/+	18232086	
RNA-seq of E7.5 mouse embryos (SE)	MD_WT4	Cmtr2+/+	26744940	
RNA-seq of E7.5 mouse embryos (SE)	MD_WT5	Cmtr2+/+	25379686	
RNA-seq of E7.5 mouse embryos (SE)	MD_KO1	Cmtr2-/-	26585383	
RNA-seq of E7.5 mouse embryos (SE)	MD_KO2	Cmtr2-/-	22964312	
RNA-seq of E7.5 mouse embryos (SE)	MD_KO3	Cmtr2-/-	31950020	
RNA-seq of E7.5 mouse embryos (SE)	MD_KO4	Cmtr2-/-	21348089	
RNA-seq of E7.5 mouse embryos (SE)	MD_KO5	Cmtr2-/-	21435906	

Discussion

My thesis aimed to understand the physiological function of extended mRNA cap structure by studying Cap-specific mRNA (nucleoside-*2'-O-*)-methyltransferases, *Cmtr1* and *Cmtr2*, in mammals.

Particularly, I investigated the role of the enzymes *Cmtr1* and *Cmtr2* in mouse embryonic development. Lack of *Cmtr1* (Chapter I, Figure 1) and *Cmtr2* (Chapter I, Figure 4) leads to developmental defects during gastrulation stages and embryonic lethality. Notably, the lack of cap1 and cap2 does not lead to any activation of the innate immune system in mutant embryos at the time of their embryonic arrest (Chapter I, Figure 1E, 4E), as could be expected due to their established roles in innate immunity. Interestingly, *Cmtr1* and *Cmtr2* seem to regulate different subsets of genes (Chapter I, Figure 4F). Lack of *Cmtr1* leads to the downregulation of snoRNA host genes (SNHG) (Chapter I, Figure 1E), whereas the most dysregulated class of genes in *Cmtr2* mutants seems to be 5s rRNA (Chapter I, Figure 4E).

To understand the role of cap-specific methyltransferases beyond the embryonic development, I first focussed on *Cmtr1,* generating conditional knockout (cKO) animals to analyse the role of *Cmtr1* in the germ line and liver (Chapter I, Figure 3). Investigation of *Cmtr1* liver cKO (*Cmtr1*$^{loxP/-}$*; AlbCreERT2*$^{ki/+}$) mutants uncover chronic interferon pathway activation in the livers of mice with *Cmtr1* mutations. *Cmtr1* germ cells cKO (*Cmtr1*$^{loxP/-}$*; Mvh-Cre*$^{ki/+}$) males are infertile, whereas females show low penetrance infertility (Chapter I, Figure 2).

Interestingly, despite the essential role of *Cmtr1* in embryonic development, the human HAP1 *CMTR1* KO cell line is viable, albeit showing reduced cell growth not caused by cell cycle defects (Chapter I, Figure S3C, D). The analysis of the influence of cap1 on translation shows that global translation remains unchanged in *Cmtr1* mutant livers (Chapter I, Figure 3K, L) and human cells (Chapter I, Figure S4F).

This investigation comprehensively examines CMTRs′ roles in development and selected organs and sheds new light on their complex functions and implications.

Embryonic lethality

As previously described, lack of any of the cap-specific mRNA (nucleoside-*2'-O-*)-methyltransferases leads to embryonic lethality in mice (Groza et al., 2023; Y. L. Lee et al.,

2020) (Chapter I, Figure 1B, 4B). However, the cause of embryonic arrest is not known. To understand the role of CMTR1 and CMTR2 on mouse embryonic development, we obtained *Cmtr1*$^{+/-}$ (Chapter I, Figure S1A) and *Cmtr2*$^{+/-}$ (Chapter I, Figure S5A) animals. Crosses between heterozygous animals for each of the two genes *Cmtr1* and *Cmtr2* confirmed that indeed both *Cmtr1* KO (*Cmtr1*$^{-/-}$) (Chapter I, Figure 1B) and *Cmtr2* KO (*Cmtr2*$^{-/-}$) (Chapter I, Figure 4B) are lethal. To understand the importance of both genes, we crossed *Cmtr1* and *Cmtr2* heterozygous animals and isolated embryos at post-implantation stages. Dissection of embryos showed that both *Cmtr1* (Chapter I, 1C, 1D and S1C) and *Cmtr2* (Chapter I, Figure 4C, 4D and S4C) mutant embryos show developmental defects during the gastrulation stages, although the phenotype slightly differs. *Cmtr1* mutant embryos start to exhibit size differences at E7.5 (Chapter I, Figure 1D) and do not reach E8.5, whereas *Cmtr2* mutants proceed gastrulation and some reach E8.5 with some escapers reaching organogenesis with size defects (Chapter I, Figure 4D). This indicates that the embryonic arrest of *Cmtr1* mutants occurs one day before *Cmtr2* mutants.

Embryonic lethality of both mutants in the mid-stage of their development points to the importance of the correct mRNA cap structure for proper entry into organogenesis but not for early embryonic developmentInterestingly, this is in contrast to the early studies of mRNA cap structure, which postulated an essential role of mRNA cap in the oocyte-to-the-embryo and maternal-to-zygotic transition. In moth, tobacco hornworm was shown that the mRNA cap structure differs between oocyte (non-methylated G cap) and developing embryo (m^7G cap; cap0), and the switch from cap0 to cap1 is crucial for the oocyte-to-the-embryo transition (Kastern et al., 1982; Kastern & Berry, 1976). In vertebrates, mos/MAPK pathways play a critical role in modulating oocyte meiotic cell-cycle progression. Full-grown oocytes have no Mos protein but contain translationally dormant c -*mos* mRNA, which is being translated soon after the oocytes are exposed to progesterone, the primary stimulus of maturation (Cooper, 1994). In *Xenopus*, cap1 methylation on c-mos mRNA was shown to be crucial for oocyte maturation *via* translational activation of c-mos mRNA (Kuge et al., 1998). On the other hand, dormant embryos and developed embryos of brine shrimp possess the mRNA cap1 structure, and the addition of methylation inhibitor, a SAM product, SAH, does not alter the translation (Muthukrishnan et al., 1975).

In contrast, mammalian cap1 and cap2 are not crucial for the oocyte-to-the-embryo or maternal-to-zygotic transition, as the *Cmtr1* and *Cmtr2* mutant embryos are viable till the

gastrulation stages. Both cap1 and cap2 rather play an essential role in gastrulation stages and ensuring entering the organogenesis.

One proposed argument why CMTR1 and CMTR2 mutants survive up to gastrulation stages could be the maternal contributions as oocyte contains proteins and mRNAs from the mother, which could contain CMTR1 and CMTR2 proteins and their mRNAs. Nevertheless, the maternal mRNAs start to be actively depleted around the 2-cell stage. Around the same time, zygotic gene expression occurs (Aoki, 2022). Moreover, the embryo has thousands of cells up to gastrulation, making maternal contribution improbable beyond day E3.5.

Further examination of the mutant embryos points towards some mechanisms that could contribute to embryonic lethality, which will be described in the following subsections.

Innate immunity during embryogenesis

As both cap1 and cap2 were earlier shown to be necessary for the distinguishing of self and non-self molecules (Despic & Jaffrey, 2023; Pichlmair et al., 2006; Schuberth-Wagner et al., 2015; Wang Yanli et al., 2010), one expected cause of the embryonic arrest and following lethality is the activation of the innate immune system, which could be detected in RNA sequencing datasets as elevated expression of the Interferon stimulated genes – ISGs (Schoggins & Rice, 2011). However, sequencing of both *Cmtr1* (Chapter I; Figure 1E) and *Cmtr2* (Chapter I; Figure 4E) mutant embryos at E6.5 and E7.5 did not show any ISGs being expressed.

Several studies showed that innate immunity, recognition of non-self RNA molecules, is inhibited in the early stages of embryonic development in human and mouse embryonic stem cells (ESC) (L. L. Chen et al., 2010; R. Wang et al., 2013, 2014). The expression of cytoplasmic sensors of viral-like mRNAs was shown to be inhibited by microRNAs (Witteveldt et al., 2019). Moreover, the lack of ADAR1, a protein critical for modyfing endogenous dsRNA molecules in order to avoid innate immune system activation, leads to embryonic lethality only around E11.5 due to the defects caused by the activation of the innate immune system. Knockout of downstream factors in the interferon signalling pathway like *Stat1*, *Ifnar1* (Mannion et al., 2014) or MDA5 results in prolonged survival of *Adar1* KO animals for two more days.

Taken together, studies on mESC show that innate immunity is initially inhibited during the embryonic development, but its overactivation leading to embryonic lethality in *Adar1* KO mice at E11.5, indicates that in mouse the innate immunity should be activated between E3.5

(mESC) and E11.5 (lethality of *ADAR1* KO mice). Since both *Cmtr1* and *Cmtr2* mutant embryos show lethality in this window, we analysed the expression of cytoplasmic innate immunity sensors and their downstream signalling pathway partners in publicly available RNAseq datasets and showed that the expression of those starts only during organogenesis stages around E10.5 (Chapter I; Figure 1F, S1D). This explains the lack of ISG expression in *Cmtr1* and *Cmtr2* mutants and points out to other regulatory roles of both mRNA cap structures.

CMTR1 and CMTR2 do not have a redundant role in mouse embryogenesis

Two recent studies from Mathias Soller′s group propose that in *Drosophila melanogaster* CMTr1 and CMTr2 could have overlapping roles, where both CMTr1 and CMTr2 could form cap1 structure. Moreover, only double mutants of *CMTr1* and *CMTr2* show reward learning defect phenotype (Dix et al., 2022; Haussmann et al., 2022). Finally, *in vitro,* methylation assay shows that hCMTR2 can methylate both the first and second transcribed nucleotide (Dix et al., 2022).

Compared to flies, the situation in mice differs. Both CMTR1 and CMTR2 are essential for embryonic development in mice. Analysis of gene expression changes between *Cmtr1* and *Cmtr2* mutant embryos at E6.5 showed that out of hundreds of dysregulated genes in mutant embryos, only 8 dysregulated genes are shared between those two (Chapter I; Figure 4F), indicating that both regulate different subsets of genes.

Gastrulation in mice is a crucial developmental stage marked by the emergence of the primitive streak around E6.5. Epiblast cells, destined to give rise to the embryo, migrate through this primitive streak and differentiate into precursors of the two primary germ layers: the mesoderm and the definitive endoderm (Tam & Loebel, 2007). Examining *Cmtr1* and *Cmtr2* mutant embryo's bulk RNA sequencing data deconvoluted to a single-cell level suggests distinct roles for both cap-specific methyltransferases.

At E6.5, both mutants already display a significant downregulation of primitive streak markers, demonstrating the immediate effects of these mutations. By E7.5, the developmental trajectory of these mutants begins to diverge further. *Cmtr1* mutants show a decreased presence of mixed mesoderm markers, retaining a high proportion of parietal endoderm and extraembryonic ectoderm cells (Chapter I; Figure S1E). In contrast, *Cmtr2* mutants exhibit an increased presence of pharyngeal mesoderm but fewer haematoendothelial progenitors and blood progenitors (Chapter I; Figure S4H). Interestingly, The International Mouse Phenotyping

Consortium (IMPC) reports that *Cmtr2* heterozygotic animals have increased heart weight supporting that CMTR2 is crucial for proper cardiovascular system development (Groza et al., 2023). These divergent phenotypes underscore the unique developmental consequences of each mutation.

Given the fact that hCMTR1 and hCMTR2 could both methylate the first transcribed nucleotide in *in vitro* methylation experiment (Dix et al., 2022; Haussmann et al., 2022), we examined polyA+ RNA from CMTR1 KO cells by mass spectrometry. Double-purified polyA+ RNA showed a decrease in N_m modification but not its complete lack (Chapter I; Figure S4B). The cause can be rRNA contamination, as the presence of $m^{6,6}A$ (rRNA-specific modification) in the polyA+ samples was detected (data not shown). Nevertheless, the complete lack of m^6A_m in *CMTR1* KO cells was observed. m^6A_m is cap specific modification that depends on both CMTR1 and PCIF1. The complete lack of m^6A_m indicates that CMTR2 could not methylate cap1. The discrepancy between *in vitro* and *in vivo* methylation results might be explained by cellular localisation of both proteins. CMTR1 localises to the nucleus, whereas CMTR2 predominately to the cytoplasm (Inesta-Vaquera et al., 2018; Liddicoat et al., 2015; Smietanski et al., 2014; Toczydlowska-Socha et al., 2018; Werner et al., 2011). Taken together, their overlapping functions in mice *in vivo* are unlike but cannot be excluded.

Embryonic lethality of *Cmtr1* mutants

Cmtr1 KO embryos cannot progress from the gastrulation stage to the organogenesis stage. The first noticeable differences between WT and *Cmtr1* KO embryos emerge around the mid-to-late gastrulation stage (E7.5), characterised by a reduced size. This highlights the crucial role of CMTR1 in gastrulation for the proper initiation of organogenesis.

Gastrulation initiates with the formation of the primitive streak on the surface of the epiblast, which consists of cells that will contribute to both the embryo and the placenta. During this process, cells within the primitive streak undergo invagination, moving inward to establish the three primary germ layers: ectoderm, mesoderm, and endoderm. This coordinated cell migration not only shapes the basic body plan but also establishes the embryo's anterior-posterior and dorsal-ventral axes. Another important event during gastrulation is the formation of the notochord. Additionally, the embryo undergoes a metabolic switch from anaerobic glycolysis to mitochondria-dependent aerobic oxidative phosphorylation during gastrulation and organogenesis. This metabolic shift is essential for supporting the rapid growth and development of both the embryo and the placenta.

To investigate the causes of embryonic arrest in *Cmtr1* KO embryos, we conducted RNA sequencing on embryos at gastrulation stages E6.5 and E7.5, aiming to capture early transcriptome changes. The RNAseq analysis of E6.5 *Cmtr1* mutants identified approximately 200 genes exhibiting altered expression, which increased to about 1600 genes one day later at E7.5. While the RNAseq analysis did not directly answer the question of what causes the embryonic arrest or the failure of proper cell differentiation, it provided some preliminary clues, such as dysregulations in few signalling pathways, snoRNA host genes or 5′TOP transcripts that will be discussed in the upcoming subchapters.

Dysregulation of signalling pathways

One of the critical factors that plays a role in the coordination of processes during gastrula development are signalling pathways such as signalling of bone morphogenetic proteins (BMPs), fibroblast growth factors (FGFs), Hedgehog, Nodal and Wnt (Heisenberg & Solnica-Krezel, 2008). It is not surprising that imbalances in signalling pathways, such as the distortion of NF-κB or WNT, can lead to embryonic arrest (Nguyen et al., 2018; Sidrat et al., 2021a, 2021b; Sokol, 2015; C. Xu et al., 2018). The WNT signalling pathway is crucial for establishing the anterior-posterior axes of embryos (Zou, 2006). Gradual inhibition of the WNT signalling in the posterior "tail" region is necessary for proper development (Nguyen et al., 2018; Sidrat, Rehman, Joo, Lee, Kong, et al., 2021). Disruptions in the WNT signalling pathway can lead to gastrulation defects due to its fundamental roles in cell fate determination and cell polarity (Sokol, 2015).

In our analysis of *Cmtr1* mutant E7.5 embryos, we observed the downregulation of genes associated with the WNT and hedgehog signalling pathways as observed by GOterms and KEGG pathways. Since the dysregulation of these pathways occurred at E7.5 but not at E6.5, it suggests that the dysregulation of signalling pathways might be rather a secondary effect than the primary cause of the lack of CMTR1.

snoRNA host genes (SNHGs)

The analysis of gene expression changes between WT and *Cmtr1* mutant embryos revealed dysregulation of snoRNA host genes (SNHGs) (Chapter I; Figure 1E, S1G). SNHGs are protein-coding or long non-coding genes that carry snoRNAs in their introns, and proper splicing is needed for snoRNA processing. We analysed the level of intron-containing snoRNAs and found no differences between the KO and WT embryos regarding snoRNA

levels (Chapter I; Figure S1F). It is important to note that the coverage of introns in RNA-seq data is generally low due to the degradation and technical limitations. Since neither the snoRNAs nor the introns containing snoRNAs appear to be dysregulated, it is more likely that if SNHGs contribute to any embryonic defects, it is due to the genes themselves rather than the snoRNAs they carry.

SNHGs with the most significant changes among the downregulated include *Snhg12, Snhg8, Snhg5, Snhg4* and *eIF4a2* (Chapter I; Figure 1E). *Snhg12* is a long non-coding RNA (lncRNA) that encodes four small nucleolar RNAs (SNORA66, SNORA61, SNORA16A, and SNORD99). Its downregulation leads to decreased cell proliferation. Additionally, *Snhg12* is a marker gene in many cancers.

Snhg8 is chromatin-localized lncRNA encoding for SNORA24. *Snhg8* plays a role in the differentiation of epithelial cells. Its downregulation contributes to the differentiation of epithelial cells through binding to DNA at histone H1 sites. H1s-mediated chromatin compaction and transcriptional repression have been proposed to play a critical role in human stem cell differentiation (P. He et al., 2022). Moreover, its knockdown inhibits cell proliferation and colony formation while promoting cell apoptosis in diffuse large B-cell lymphoma tissue (D.-H. Yang et al., 2021), and it is upregulated in various tumour types, increasing the proliferation, migration, invasion, and metastasis of cancer cells (Yuan et al., 2021).

Snhg5 carries SNORD50, and its knockdown decreased breast cancer cell proliferation (Chi et al., 2019; J. Li & Sun, 2018). Snhg4 is known to play an oncogenic role in tumours. The overexpression of Snhg4 positively impacted cell growth in MHCC-97H. Moreover, high expression of Snhg4 in acute myeloid cancer patients correlated with shorter survival (J. Li & Sun, 2018; Qiu et al., 2023).

eIF4a2 is eukaryotic initiation factor 4A belonging to the extensive DEAD-box RNA helicase family (W. T. Lu et al., 2014). Interestingly, EIF4A2 interacts with Ccr4-Not to promote miRNA-mediated translational repression of 5′ purine-rich mRNAs (Wilczynska et al., 2019). Similarly to other SNHGs, the knockdown of *EIF4A2* decreases growth and metastasis in colorectal cancer (Z.-H. Chen et al., 2018).

Overall, the expression of many SNHGs needs to be finely tuned, as their overexpression is observed in various cancers, many being marker genes, and their knockdown leads to decreased cell growth. Similarly, snoRNAs were shown to be regulated during differentiation (McCann et al., 2020). However, whether the observed phenotypes are caused by the genes themselves or the snoRNAs they carry remains to be elucidated.

5′TOP transcripts

Numerous SNHGs belong to the ribosomal gene family, many of which feature a 5′ terminal oligopyrimidine tract (5′TOP). 5′TOP is distinguished by a unique feature known as the 5′ terminal oligopyrimidine tract (5′ TOP). This characteristic element consists of a cytidine residue at the transcription start site nucleotide, followed by a continuous sequence of up to 13 pyrimidines. This feature plays a crucial role in coordinated translation control and it has been observed that the translation of mRNAs of ribosomal proteins is inhibited when cell growth is arrested. This phenomenon occurs in all cell lines studied, both in laboratory conditions and in living organisms (Avni et al., 1997).

Upon investigating these ribosomal genes and 5′TOP transcripts in Cmtr1 KO embryos, we observed a slight decrease in the overall representation of both categories (Chapter I; Figure S1). Interestingly, lack of CMTR1 has already been linked to downregulated gene expression of ribosomal protein and histone genes in mouse embryonic stem cells (mESC) (Galloway et al., 2021). 5′TOP mRNAs are targets of cap0 binding protein LARP1 (Galloway et al., 2021; Philippe et al., 2020). Interestingly, LARP1 anchors 5′ TOP transcripts within SGs and PBs upon cellular stress. Granule localisation in stress does not affect translation or decay during recovery (Wilbertz et al., 2019).Notably, the dysregulation of balance between LARP1 and 5′ TOP transcripts leads to p53 stabilisation and cell cycle arrest.

It is not clear what causes the downregulation of 5′ TOP transcripts and whether the slight downregulation of 5′ TOP transcripts observed in mutant embryos is sufficient to trigger embryonic arrest.

Cap1 regulation molecular functions.

Cap0 affects all aspects of mRNA life, from its stability, splicing, and mRNA export to translation, with most functions being affected *via* the cap-binding complex (CBC). CBC has a higher affinity to bind capped RNAs with purines at TSS than pyrimidines. Nevertheless, any effect of CBC binding to cap1 vs cap0 was not observed. Interestingly, the only increase in CBC binding was to the m^6A starting cap analogue (Worch et al., 2005); nevertheless, the involvement of cap1 in molecular processes, such as stability, splicing, and mRNA export to translation, needs to be better studied.

Since we observed no defects in early *Cmtr*1 mutant embryos (up to E6.5) as well as Hap1 KO line and liver conditional mutant cells are viable, cap1 seems to be not essential for

global transcription, mRNA splicing, mRNA export, translation or mRNA stability. However, failure of embryos to reach gastrulation suggest that it is still possible that cap1 regulates a subset of transcripts or affects all transcripts at low levels.

Regulation of translation

Cap1′s role in the regulation of translation has been demonstrated *in vivo* within *Xenopus* oocytes (Kuge et al., 1998) and in some cell lines (Drazkowska et al., 2022). Furthermore, cap1′s involvement has been established in antiviral environments where the expression of interferon and subsequent interferon-stimulated genes (ISGs) lead to the inhibition of cap0 mRNAs′ translation (Abbas et al., 2017; Fleith et al., 2018; Habjan et al., 2013; Pichlmair et al., 2011).

Our study sought to investigate cap1′s role in protein synthesis, specifically focusing on *Cmtr1* conditional mutant (cKO) livers. Remarkably, these mutant livers showed an upregulation of ISGs (Chapter I; Figure 3E), including IFIT1 (Chapter I; Figure 3C), a cap-binding protein that selectively recognises cap0 but not cap1. This selectivity allows it to inhibit cap0 mRNAs′ translation by preventing the recruitment of cap-binding eukaryotic translation initiation factors (Abbas et al., 2017; Fleith et al., 2018; Habjan et al., 2013; Pichlmair et al., 2011).

To assess the overall translation state, we performed sucrose-gradient centrifugation on liver lysates, generating polysome profiles from both control and *Cmtr1* cKO livers (Chapter I; Figure 3K, S3D). Although both profiles showed expected monosome and polysome peaks, the mutant livers display slight downregulation across all peaks in both duplicates. Notably, this suggests that the overall translation in cKO livers is not significantly affected, only showing a slight negative impact. Importantly, we normalised the input material to the total protein amount rather than the same number of cells. The normalisation of the input to the number of cells rather than protein concentration might have yielded different results, particularly considering the observed downregulation of ribosomal protein genes in RNA expression. In summary, the expression of IFIT1 in cKO livers does not affect global translation.

We conducted ribosome profiling to examine the translation status of individual mRNAs in mouse livers, which determines ribosome occupancy on mRNAs through deep sequencing (Ingolia et al., 2009). We normalised the ribosome footprint reads mapped to the coding sequence to transcript expression levels derived from the input lysate. Evaluating translation efficiency, we observed low coverage of mRNAs due to inefficient rRNA depletion

. Notably, the low coverage of mRNA reads in the ribosome profiling library due to rRNA depletion from fragmented RNA (Chapter I; Figure S3H) is a common method limitation. Nevertheless, the analysis of coding sequences revealed significant alterations in the translation of several transcripts in the *Cmtr1* cKO liver (Chapter I; Figure 3L, M), as well as slight downregulation of the whole class of genes belonging to histone genes, snoRNA host genes, ribosomal genes and 5′ TOP (Chapter I; Figure S3F). As previously reported, a lack of *Cmtr1* in mESCs leads to replication stress which might explain the downregulation of histone genes (Liang et al., 2022).

Previous research on interferon-treated cell cultures has demonstrated that the translation of three ISGs, ISG15, MX1, and IFITM1, depends on CMTR1, as these genes require cap1 methylation to avoid IFIT1-mediated inhibition (Williams et al., 2020). Our dataset echoes these published findings, showing that while ISG15′s expression is upregulated at the RNA level, its translation is paradoxically downregulated in the absence of CMTR1 (Chapter I; Figure 3H, L, M). These findings give rise to the hypothesis that IFIT1-mediated translational inhibition may affect mRNAs that share specific primary or secondary structures, potentially leading to their comparative translational inhibition. Furthermore, it raises the question of whether these structures are shared with other cellular genes, particularly among the 161 translationally downregulated genes we identified.

Conversely, *CMTR1* KO HAP1 cells exhibited over a thousand genes dysregulated relative to WT on the transcriptome level (Chapter I; Figure S4E, F). However, ribosome profiling showed only 20 genes as significantly dysregulated. Analysis of gene classes revealed mild downregulation of ribosomal genes and 5′ TOP mRNAs. Comparison of translational efficiency between liver and Hap1 cells showed no overlapping genes.

Additionally, no significant differences were found in the translation of mRNAs based on their Transcription Start Sites (TSS). We generally see slightly fewer translated C starting transcripts, which agrees with previously published data (Tamarkin-Ben-Harush et al., 2017). However, it is worth noting that we used database annotations to classify the mRNAs' TSS, which might differ in livers or after the conditional deletion of *Cmtr1*.

In conclusion, our study does not show a strong effect of cap1 on global translation and points out only to downregulation of histone, snoRNA host genes, and ribosomal genes. It is important to note that histones and ribosomal genes are sensitive to cellular stress. Therefore, we cannot conclusively determine if their decreased translation is a

primary effect of cap1 absence or a secondary effect due to the activation of the innate immune system or another form of cellular stress.

Regulation of splicing

Cap0 cap structure is crucial for mRNA splicing. Nevertheless, the involvement of cap1 in splicing still needs to be determined. Cap1 was shown to be important for splicing in chimeric U2 snRNA (Dönmez et al., 2004). We performed an alternative splicing events analysis to assess the changes in splicing between Cmtr1 KO and WT embryos.

We identified 195 significantly dysregulated events in mutant E6.5 embryos, corresponding to ~1% of all events (we identified 20763 events in our dataset), with the most dysregulated events being intron retention. The number of intron retention events further increases at E7.5 (Chapter I, Figure 1G). A more detailed analysis of intron retention events revealed that at the E6.5 stage the second intron is the most retained one, while at E7.5, it is intron 2, 3, and 4 (data not shown).

Similarly, we performed the same analysis for *Cmtr1* cKO livers at all three time points (D2, D6, D22). Interestingly, our preliminary search at time point D6 showed the most changes of all three time points with splicing events of alternative first exons, alternative 5′ splice sites, alternative 3′ splice sites, alternative last exon, skipped exon, as well as retained introns being between 2-4%.

Although the observed differences at alternative splicing are worth further examination, this goes beyond the scope of this study. Nevertheless, there are a few possibilities that cause the differences in alternative splicing. The causes might be the primary effect, lack of cap1 methylation at mRNAs, snRNAs, or secondary cause such as interferon expression. A few studies have already pointed out that splicing is altered upon viral infection or interferon (Liao & Garcia-Blanco, 2021; E. K. Robinson et al., 2021; Sertznig et al., 2022), and thus we cannot point out whether the differences observed in cKO livers are due to the lack of cap1, as primary effect or due to the activation innate immune system. Additionally, it is worth noting that the strongest interacting partner of CMTR1 is DHX15 (Chapter I, Figure S4J) (Inesta-Vaquera et al., 2018; Toczydlowska-Socha et al., 2018), a splicing factor that was found to repress suboptimal introns with weak splice sites, multiple branch points, and cryptic introns (J. Zhang et al., 2022).

Nevertheless, further studies are needed to confirm or disprove any primary effect of cap1 on alternative splicing.

Recognition of non-self

According to initial studies, the presence of cap1 is of utmost importance for distinguishing self and non-self RNA molecules. Remarkably, my work shows that developing embryos that lack *Cmtr1* show embryonic lethality before initiating the innate immune system. Our research focus shifted towards conditional mutants to investigate the role of CMTR1 outside of developing systems and facilitate the study of innate immunity. Western blot analysis revealed the expression of CMTR1 protein in all tested tissues except muscle and white fat tissues, displaying one to three bands (Chapter, Figure 3A). The liver is one organ known for possible interferon expression activation (Liddicoat et al., 2015a; Mannion et al., 2014). Additionally, the liver offers the advantage of being a large organ, enabling the execution of numerous diverse analyses. These characteristics make the liver an ideal organ for investigating innate immunity, particularly in CMTR1 conditional depletion background. We specifically generated conditional *Cmtr1* mutants in the liver using the tamoxifen-activated *AlbCreERT2* recombinase. The *AlbCreERT2* recombinase is expressed exclusively in the liver under the control of the *Alb* promoter. CreERT2 recombinase localises to the cytoplasm, and only upon the tamoxifen injection it translocases to the nucleus.

To investigate the role of the lack of *Cmtr1,* we examined cKO model animals (*Cmtr1*$^{loxP/-}$; *AlbCreERT2*$^{+/-}$). The experimental setup involved tamoxifen injections for four consequent days and collecting liver tissues at three different time points: 2, 6, and 22 days later (Chapter, Figure 3B). As lack of CMTR1 is expected to lead to activation of the innate immune system (Williams et al., 2020), and previous studies on Adar1 cKO mutants (Alb-ADAR1 cKO) with the deletion from E10.5 resulted in phenotypes such as growth retardation, smaller livers, and high mortality at early ages (G. Wang et al., 2015), we were monitoring the experimental animals daily for loss of weight or any discomfort phenotypes such as pain. In contrast to *Adar1* cKO, our *Cmtr1* cKO animals did not exhibit any growth retardation, as the deletion occurred only when the animals were adults. Still, no weight loss or discomfort was observed in the animals, even 22 days after the last injection. Additionally, we examined the size of the liver and performed histological examinations, but no significant changes in liver cellularity were observed (Chapter I, Figure 3C).

Evaluation of conditional mutants of *Cmtr1* in livers showed that already two days after the last injection, CMTR1 was depleted from livers (Chapter I, Figure 3C). We observed the upregulation of interferon-stimulated genes at all three-time points (Chapter I, Figure 3E-I).

Since the lack of cap1 leads to continuing interferon expression even after 22 days, this suggests that the absence of cap1 leads to chronic activation of the innate immune response.

Interestingly, *Adar1* liver cKO with constitutively expressed Cre recombinase under Alb promoter show more severe issues which might be caused by the activation of innate immune system during development of liver tissue (Hartner et al., 2008). The question is what would lead to more severe phenotype, lack of cap1 or lack of inosine, using cre recombinase under same promotor and conditions.

In summary, lack of cap1 in adult mouse liver leads to the chronic stimulation of the innate immune system. This offers a possibility to study *Cmtr1* liver cKO animals as a model to study autoimmune hepatitis as a long-term chronic liver disease that causes inflammation and liver damage.

Mouse germ cells conditional mutants

The conditional deletion of *Cmtr1* in mouse germ cells leads to varying phenotypes between genders. We accomplished this conditional deletion in the mouse germline using the germline-specific *Mvh-Cre* line (also known as *Vasa-Cre*). This line expresses the Cre recombinase transgene, regulated by the mouse vasa homolog (*Mvh*) promoter. Expression of *Mvh-Cre* starts at embryonic day E14.5 creating deletion id germ cells around this day. No offspring were observed when conditional knockout (Cmtr1loxP/-; Mvh-Cre+/-) males were bred with wild-type females. A dissection of the testes at postnatal day 31 (P31) revealed significant atrophy compared to the control group ($Cmtr1^{loxP/+}$; $Mvh\text{-}Cre^{+/-}$) (Chapter I, Figure 2B).

A histological examination of the knockout testes shows a pattern of progressive degeneration with ageing. At birth, P0, the cKO testes appeared normal compared to the control group. However, by P31, when the first germ cells should have differentiated, notable defects in differentiation were observed. The cells remained in one to three layers but failed to differentiate properly. This effect was even more pronounced in adults (P75). The seminiferous tubules in adult knockout testes were narrow and devoid of germ cells (Figure 2C). In stark contrast, tubules in the control testes were full of germ cells at all stages of development during spermatogenesis, including mitotic spermatogonia, meiotic spermatocytes, post-meiotic haploid round spermatids, elongate spermatids, and sperm (Chapter I, Figures 2C and S2C).

On the other hand, *Cmtr1* cKO females, when crossed with WT partners, showed decreased fertility as some females showed infertility and in general cKO females had slightly

lower number of progeny compared to WT females (Chapter I, Figures 2E). Notably, during the experimental setup, if no pregnancy was observed during the first month, the female was paired with a different partner to avoid potential issues related to individual mate preferences or compatibility, ensuring that the observed results were due to the genetic modification and not a mismatch between partners. Nevertheless, the histological examination of ovaries did not show any significant differences between follicles between cKO and control females. (Chapter I, Figures 2E).

As males are completely infertile and females display low penetrance infertility, few possibilities could explain the differences between male and female phenotypes. The most straightforward is that CMTR1 might be crucial for spermatogenesis but not oogenesis. Another possibility lies in the timing of Mvh-Cre expression, which starts at E14.5. This is a significant point in female development as meiosis has already begun, whereas male germ cells are in mitotic arrest before meiosis. Consequently, a hypothesis that CMTR1 may be essential for entering meiosis cannot be confirmed with the current *Mvh-Cre* cKO model. To ascertain whether CMTR1 plays a role in oocytes entering meiosis, other Cre lines would need to be established and tested.

Furthermore, another potential explanation for male infertility could be the expression of interferons, as activation of the innate immune system has been associated with infertility. This aspect should be considered in future investigations of the underlying mechanisms contributing to male infertility in the absence of CMTR1.

Embryonic lethality of Cmtr2 mutants

This subchapter will discuss the possible causes of *Cmtr2* mouse embryonic lethality. As previously discussed, *Cmtr2* mutant mice fail to survive beyond the gastrulation stages, underscoring the critical role of this gene in early embryonic development. RNA sequencing of embryos at E6.5 and E7.5 gastrulation stages reveals dysregulation of hundreds of genes. Gene ontology analysis of these genes points to numerous diverse pathways, implying a significant impact on organogenesis.

Ribosome biogenesis in *Cmtr2* mutant embryos

Despite the ribodepletion of the RNAseq libraries of *Cmtr2* embryos (E6.5 and E7.5), we could still detect 5s rRNA reads in our sequencing data. Interestingly, the 5s rRNA gene is notably downregulated among the most dysregulated genes (Chapter I, Figure 4E). At the E6.5 stage,

24 out of 39 detected 5s rRNA genes are downregulated. One day later, 8 out of 21 are downregulated and 3 upregulated.

The 5s rRNA is the only rRNA not transcribed by RNA pol I but rather by RNA pol III from a single cluster located on chromosome 8 (Kampen et al., 2021; Zhou et al., 2015). Dysregulation of ribosomal components can lead to the activation of p53, causing cell death (Kampen et al., 2021). The complex of 5s rRNA, RPL5, and RPL11, known as free 5s rRNP, is a vital player in the regulation of p53 by ribosomal proteins (Donati et al., 2013; Sloan et al., 2013). KEGG pathway enrichment showed 7 different P53 pathway components to be upregulated, such as p21, DR5, Noxa, PUMA, Sestrins, Cyclin G, and Np73. Collectively, the dysregulations in 5s rRNA and subsequent p53-mediated apoptosis might explain embryonic lethality. However, considering that the libraries were ribodepleted, the dysregulation of 5s rRNA might simply be an artefact. Therefore, techniques like northern blot or quantitative real-time PCR should be utilised to confirm this. If the downregulation of 5s rRNA is indeed confirmed, the regulatory mechanism underlying it would need further examination.

Splicing in Cmtr2

As an earlier study showed, cap2 on U2 snRNA is crucial for efficient splicing (Dönmez et al., 2004). Our study did not focus on splicing of CMTR2 as to dissect the primary and secondary causes in developing/dyeing embryos seems to be tricky; nevertheless, the focus on CMTR2 and its involvement in the splicing might be interesting in future research as the RNA sequencing revealed two snRNAs being downregulated - RNu11 and RNu12. Both RNu11 and RNu12 are snRNAs of the minor spliceosome. Minor spliceosome, is a spliceosome consisting of U11, U12, U4atac U6atac, and U5 snRNAs being crutial for recognition of specific subset of introns being call U12 dependent (Juan-Mateu & Valcárcel, 2023; Tarn & Steitz, 1996; Verma et al., 2018). Constitutive deletion of RNu11 in mice leads to embryonic lethality, whereas conditional deletion in the neocortex leads to microcephaly display cell cycle defects and p53-mediated cell death (Baumgartner et al., 2018). Mutations in RNU12 lead to congenital cerebellar ataxia characterised by delayed motor milestones in development, mild learning difficulties and hypotonia in infancy (Elsaid et al., 2017). Taken together, both U11 and U12 snRNA have vital roles in development. Nevertheless, it is unknown whether cap2 could stabilise the U11 and U12 or if their downregulation is a secondary effect. A non-developing model would be advantageous to investigate splicing in the genetic background of *Cmtr2*.

Remaining questions in the field

While this study provided several answers, it also paved the way for even more questions to be answered in next research studies. With the questions concerning CMTR1 and CMTR2 as well as more in the field touching innate immunity.

Could m^6A at the first transcribed nucleotide (m^7Gpppm^6A cap) contribute to the recognition of self and non-self-RNA molecules?

The presence of m^6A at the first transcribed nucleotide (m^7Gpppm^6A cap) raises intriguing questions about its potential role in distinguishing self from non-self-RNA molecules. It seems that the majority of cellular mRNAs start with adenosine (Akichika et al., 2018; Galloway et al., 2020), which only in 8% represent Am, whereas 92% of them form m^6A_m (Akichika et al., 2018).

It has been observed that many viruses possess m^6A_m at their cap structures, such as VSV (Tartell et al., 2021), dengue virus (FURUICHI, 2015), coronavirus (L. Wang et al., 2023), and vaccinia virus (Boone & Moss, 1977). While no viral methyltransferases have been identified, it is possible that PCIF1, known for its nuclear localisation (Pandey et al., 2020), translocated during viral infection to the cytoplasm and methylates viral RNA. PCIF1 has been shown to methylate viral cytoplasmic RNA in the case of VSV (Tartell et al., 2021) and coronavirus (L. Wang et al., 2023). Notably, removal of PCIF1 has been found to increase interferon β production in coronavirus-infected cells (L. Wang et al., 2023). The fact that vital infection promotes translocalisation of PCIF1 from the nucleus to cytoplasm supports the hypothesis that m^6A_m cap might have a vital role in sensing self- and non-self RNA molecules.

In summary, the presence of m^6A_m at the TSS nucleotide in both cellular and viral transcripts highlights the complex interplay between RNA modifications and viral replication strategies. The involvement of PCIF1, its subcellular localisation, and its role in the cytoplasmic replication cycles of certain viruses provide avenues for further investigation to unravel the intricate mechanisms underlying these phenomena.

How mRNA cap structure looks like in muscles and white fat?

Our current understanding postulates that all mRNAs carry the cap1 structure since its absence triggers the activation of interferon expression. A noteworthy observation from the protein expression analysis (conducted via western blotting) reveals the absence of CMTR1 in muscle

and white fat tissues (Chapter I, Figure 3A). Given that CMTR1 is the only known nuclear *2′-O*-methyltransferase, this implies that cap1 may not be present in these specific tissues. As cap0 has been shown to activate the innate immune system (Goubau et al., 2014; Hornung et al., 2006; P. Kumar et al., 2013; Pichlmair et al., 2006; Schuberth-Wagner et al., 2015; Wang Yanli et al., 2010; Züst et al., 2011), this, in turn, triggers a new question: Could cytoplasmic viral RNA sensors be present in these tissues?

Noteworthy is the fact, that the protein levels of CMTR1 levels in tissues were tested only by the western blot analysis, to confirm their lack further experiments should conducted such as real-time PCR, protein MS or western blot with different protein concentrations to see if the levels of CMTR1 are lower or CMTR1 is completely missing. Nevertheless, if CMTR1 is indeed not present in muscles and white fat tissues, it raises few alternative possibilities. Our first observation indicates that the protein expression of PCIF1 is noticeable in both muscle and white fat tissues (Pandey et al., 2020). Even though, our current understand of PCIF1 indicates that it preferentially methylates A_m modified transcripts, it is possible that it modifies A transcripts as well. Importantly m6A at the TSS nucleotide was already identified (J. Wang et al., 2019). This suggests that tissues with lower CMTR1 protein levels might contain m m^7Gpppm6A-RNA, as there is an assumption that m^6A at the transcription start site (TSS) nucleotide could have a role in sensing.

A second proposition is that CMTR2 might compensate for CMTR1's absence, thus tagging cellular mRNAs as self. This idea aligns with Mathias Soller's lab proposal that CMTR2 could potentially methylate the first transcribed nucleotide (Dix et al., 2022; Haussmann et al., 2022). Noteworthy is the fact the mammalian CMTR2 predominantly localises to the cytoplasm (Despic & Jaffrey, 2023; Werner et al., 2011).

Thirdly, it is conceivable that CMTR2 might achieve this by methylating the second transcribed nucleotide. The existence of Nm modification solely on the second transcribed nucleotide could shield the mRNA from being detected, albeit less efficiently than cap1 (Schuberth-Wagner et al., 2015; Wang Yanli et al., 2010). Given that roughly 50% of transcripts are believed to contain cap2 (Despic & Jaffrey, 2023; Furuichi et al., 1975; Wei Cha-Mer and Gershowitz, 1975), it is possible that CMTR2 has higher occupancy or just methylates mRNAs more effectively in these tissues.

Could CMTR2 contribute to the elimination of viral infection?

Lack of cap2 leads to mild expression of ISGs in Hek cells (Despic & Jaffrey, 2023). Additionally, cells transfected with *in vitro* synthesised RNA having N_m modification only on the second transcribed nucleotide show lower interferon expression compared to nonmethylated RNAs transfection (Schuberth-Wagner et al., 2015; Wang Yanli et al., 2010). As interferon production has a negative effect on viral growth, it raises the question if CMTR2 could be a potential target for viral therapies. To answer this question, I generated mutant *Cmtr2* lung cKO ($Cmtr2^{loxP/-}$;$SftpcCre^{+/-}$) animals expressing human COVID-19 receptor under constitutive promoter K18 (K18-hACE2) in order to study the role of *Cmtr2* in coronavirus infection.

Can CMTR1 and CMTR2 act redundantly?

Even though this study shows that CMTR1 and CMTR2 have different phenotypes and no m^6A_m modification was observed in CMTR1 KO cells indicating that CMTR2 is cannot methylate the pre-mRNAs co-transcriptionally, we cannot exclude that both CMTR1 and CMTR2 could act redundantly in some specific cases.

For example, can CMTR2, upon the lack of CMTR1, translocate to the nucleus? Or on the other hand, could CMTR1 localise to cytoplasm? For ADAR1, it was shown that upon IFN expression, it expresses a s shorter variant that localises to the cytoplasm (Galipon et al., 2017; Pestal et al., 2015). Similarly, PCIF1 is known to localise to the nucleus. Nevertheless, it was shown that during viral infection PCIF1 can translocate to cytoplasm to methylate viral RNA. (Tartell et al., 2021). This opens a question of how many mRNA modifying enzymes could change their localisation upon viral infection.

When do RNAs get cap2 methylated?

Analysis of RNAs showed that cap2 is present in mRNA, and snRNAs (Despic & Jaffrey, 2023; Werner et al., 2011), with mRNAs being cap2-modified in 50% of cases, whereas snRNAs in nearly 100% in various cell types (Krogh et al., 2017). Unlike nuclear CMTR1, CMTR2 localises to the cytoplasm (Werner et al., 2011).As CMTR2 recognizes and binds to m^7G-capped RNAs (Smietanski et al., 2014), it opens the question of when CMTR2 could methylate the RNAs. Both mRNAs and snRNAs are exported to the cytoplasm by the CBC complex, which binds to the m^7G-cap structure (Carmody & Wente, 2009; Cheng et al., 2006; Elisa Izaurralde et al., 1995). CBC stays attached to the mRNAs till it is exchanged by eIF4e

during translation (Sato & Maquat, 2009). This proposes a question of when CMTR2 can methylate the mRNAs. One possibility is that CMTR2 outcompetes CBC binding to methylate mRNA. The second would be that CMTR2 methylates the mRNA when the CBC is exchanged to eIF4e during translation.

On the other hand, the snRNAs cap2 methylation by CMTR2 is probably simpler. In the cytoplasm, snRNAs are capped when their m7G cap is further methylated forming trimethyl cap. It is probable that the cap2 methylations happened during this process as well. Taken together, cytoplasmic capping remains an interesting question for further research.

When do cytoplasmic sensors sense cap0 mRNAs?

Cellular mRNAs carrying cap0 when reach cytoplasm are recognised as non-self by cytoplasmic RIG-I like sensors to trigger interferon expression mRNAs (Chapter I, Figure 3) (Despic & Jaffrey, 2023; Williams et al., 2020) and subsequent establishment of antiviral environment by ISGs expression the question is at which moment the RNAs possessing cap0 are recognised by such sensors. Given the fact that cap-binding proteins bind and protect mammalian cellular mRNAs throughout the whole mRNAs lifespan which raises a similar question to the previous one, when cytoplasmic sensors sense cellular RNAs and at which moment the mRNA cap is naked.

Why is the innate immune system inhibited in early embryos?

In general, embryos are unable to produce interferon as showed in mouse and human embryonic stem cells (ESC) (L. L. Chen et al., 2010; R. Wang et al., 2013, 2014). On the contrary, even though the ESC cells cannot produce Interferon, they are able to respond to interferon stimulation by expression of low levels of ISGs (R. Wang et al., 2014) indicating some possibility of maternal interferon expression in embryo protection. Nevertheless, mESC retained the ability to sense viral dsRNA via Protein kinase R (PKR) leading to translation inhibition and subsequent inhibition of cell proliferation (R. Wang et al., 2013). Interferon expression upon viral infection starts to be detected only from E7.5, but remains limited to trophoblast (Barlow et al., 1984) Nevertheless the response to pathogens in trophoblast is limited. It is able to respond to viral-like dsRNA, poly I:C. the poly I:C treatment induces the expression of interferon in trophoblast through the Toll-like 3 receptor (TLR3), endosomal sensor of dsRNAs, activation and not the cytoplasmic RNA sensors. On the other hand the trophoblast is unable to respond to gram-negative bacteria as lipopolysaccharides treated

trophoblast cells did not show any interferon expression (Abrahams et al., 2006). Interestingly, trophoblast is believed to protect the cells from viral infection as when human trophoblast cells are infected with a virus, they produce protective molecules. If the media from these treated trophoblast cells is then transferred to unrelated cells before they are infected, it results in a reduced rate of viral infection compared to cells in untreated media (Delorme-Axford et al., 2013). On the other hand the cytoplasmic sensors of RIG-I family seems to be inactive till E10.5 (Chapter I, Figure 1F, S1D) The mechanism of inhibition of cytoplasmic RIG-I like sensors in mESC was shown to be due to the expression of miRNA miR-673-5p (Witteveldt et al., 2019).

This raises a question whether miR-673-5p is expressed throughout the whole early- and mid-embryonic development or is there different mechanism? What would happen is the interferon expression pathway could be activated in ESC cells and early embryos? Would it help embryos to overcome some viral infections?

Among the viruses known to trigger severe abnormalities when infection happens during embryonic development is Zika virus (ZIKV). ZIKV infection contracted during pregnancy manifests many congenital anomalies and postnatal developmental complications. These encompass not only foetal loss and intrauterine growth restriction (IUGR) but extend to grievous conditions like microcephaly. Furthermore, ZIKV can induce a variety of motor and neurodevelopmental disorders, underscoring its substantial impact on the developmental health of affected neonates. Interestingly, when the infection in murine uterus happens during gastrulation stages, E6.5 - E8.5, it is fatal, whereas infection at later stages at E13.5 leads to viable offspring with some pups with mild encephalitis (Nakayama et al., 2021). Another study using different viral and mouse strains showed, that ZIKV infection at E.4.5 has fatal outcomes whereas E8.5 leads to viable litter. Nevertheless, when the same experiments were performed with IFN receptor 1 deficient mice, ZIKV infection was much more severe and led to the demise of the embryo (Yockey et al., 2016). Taken together, activation of interferon expression in developing embryo might have beneficial roles in overcoming viral infections.

Another pathogen transmitted disease that is especially dangerous during pregnancy is toxoplasmosis. Treatment of toxoplasmosis-infected mouse embryonic fibroblast (MEF) cells with IFN-β or IFN-γ, to induce innate immune reaction, inhibited T. gondii growth (Mahmoud et al., 2015). Moreover, study on two different mouse lines showed that expression IFN-γ has a protective function on maternal-fetal transmission (Shiono et al., 2007). These findings

confirm the importance of innate immune system in viral infection and opens a potential for innate immune stimulators during pregnancy with known infection to some diseases.

Is cap1 presence conserved as a mark of self-molecules?

As interferon expression is common only for vertebrates, it raises a question whether same time of molecules could be sensed in invertebrates but different innate immune mechanism. Interestingly, studies on fruit flies (*Drosophila*) suggest that enzymes responsible for cap methylation — CMTr1 and CMTr2 — might also be involved in the immune response. When both enzymes were mutated in fruit flies, we observed an upregulation in specific immune response genes (**Figure 10**) (Haussmann et al., 2022). This suggests that cap methylation could potentially be a conserved mechanism for differentiating 'self' and 'non-self' molecules in invertebrates. Interestingly, a similar pattern has been noted for another RNA modification enzyme, ADAR, which edits RNA by converting adenosine to inosine. Initially, fruit fly mutants for ADAR were not studied for innate immunity as the mutant display quite low innate immune responce, but later studies focusing on innate immunity confirmed that the presence of inosine is a conserved indicator of 'self' RNA molecules. This reinforces the potential role of RNA modifications in shaping the immune response (Deng et al., 2020; Palladino et al., 2000b; Robinson et al., 2015).

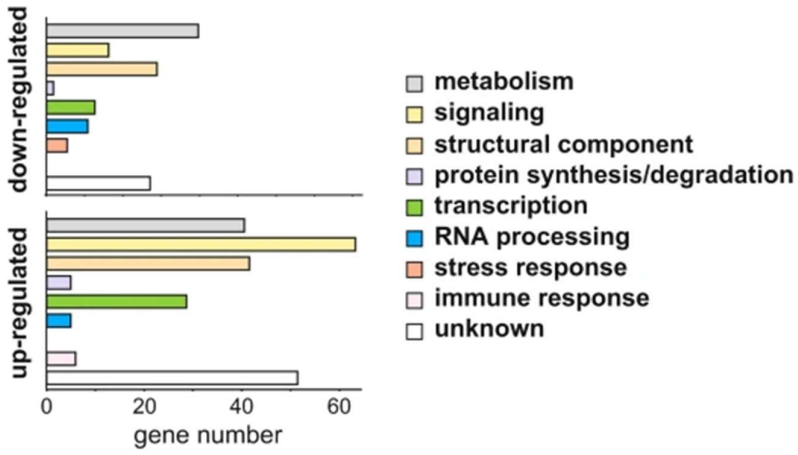


Figure 10. Functional classification of upregulated (bottom) and downregulated (top) genes in CMTr1^{13A}; CMTr2^{M32} double mutant flies compared to control flies reported by Haussmann et al., 2022.

Do SNHGs have any role in embryonic development?

This study showed significant dysregulation of numerous SNHGs in the *Cmtr1* KO embryo. Is their dysregulation in *Cmtr1* mutants contributing to the lethal phenotype? While many SNHGs function as prognostic markers in cancer, their physiological role in mice has yet to be explored. The studies on SNHGs are limited, focusing only on single cancer cell lines and lacking developmental models such as mice or screenings in mouse embryonic stem cells (mESCs) (Zimta et al., 2020) . A broader and more detailed investigation could provide important insights into the function of SNHGs in embryonic development. Moreover, an insight into the function of the SNHG versus its encoded snoRNA would also be insightful.

Conclusion

In conclusion, my PhD research advanced our understanding of CMTR1 and CMTR2 functions in mammals. Utilising mouse mutants and cell culture models, I demonstrated that both proteins are crucial for mouse embryonic development—a role that is distinct from their previously described function in the innate immune system. The absence of cap1 in livers confirmed its vital role in distinguishing self- and non-self- RNA molecules in differentiated cells, leading to chronic activation of the innate immune system. Furthermore, conditional mouse models indicated that CMTR1 is pivotal in post-embryonic development, as its absence in germ cells results in male infertility and decreased female fertility. Notably, the comparison of dysregulated genes between *Cmtr1* and *Cmtr2* mutants revealed that cap1 and cap2 structures each modulate a unique subset of genes. Among the dysregulated genes in *Cmtr1* mutants are snoRNA host genes, 5′TOP transcripts, and ribosomal protein genes. Overall, this research illuminated the roles of cap1 and cap2 methylation in gene regulation, extending beyond their role in designating cellular RNAs as 'self.' While this study provided several answers, it also paved the way for new questions in the field

Bibliography

Abbas, Y. M., Laudenbach, B. T., Mart\'{\i}nez-Montero, S., Cencic, R., Habjan, M., Pichlmair, A., Damha, M. J., Pelletier, J., & Nagar, B. (2017). Structure of human IFIT1 with capped RNA reveals adaptable mRNA binding and mechanisms for sensing N1 and N2 ribose 2^-O methylations. *Proceedings of the National Academy of Sciences*, *114*(11), E2106–E2115. https://doi.org/10.1073/pnas.1612444114

Abebe, J. S., Verstraten, R., & Depledge, D. P. (2022). Nanopore-Based Detection of Viral RNA Modifications. In *mBio* (Vol. 13, Issue 3). American Society for Microbiology. https://doi.org/10.1128/mbio.03702-21

Abrahams, V. M., Schaefer, T. M., Fahey, J. V., Visintin, I., Wright, J. A., Aldo, P. B., Romero, R., Wira, C. R., & Mor, G. (2006). Expression and secretion of antiviral factors by trophoblast cells following stimulation by the TLR-3 agonist, Poly(I : C). *Human Reproduction*, *21*(9), 2432–2439. https://doi.org/10.1093/humrep/del178

Acera Mateos, P., Sethi, A. J., Ravindran, A., Guarnacci, M., Srivastava, A., Xu, J., Woodward, K., Yuen, Z. W. S., Mahmud, S., Kanchi, M., Zhou, Y., Sneddon, A., Hamilton, W., Gao, J., Starrs, L. M., Hayashi, R., Wickramasinghe, V., Zarnack, K., Preiss, T., … Eyras, E. (2023). Simultaneous identification of m6A and m5C reveals coordinated RNA modification at single-molecule resolution. *BioRxiv*. https://doi.org/10.1101/2022.03.14.484124

Akichika, S., Hirano, S., Shichino, Y., Suzuki, T., Nishimasu, H., Ishitani, R., Sugita, A., Hirose, Y., Iwasaki, S., Nureki, O., & Suzuki, T. (2018). Cap-specific terminal N6-methylation of RNA by an RNA polymerase IIassociated methyltransferase. *Science*, *363*(6423), eaav0080. https://doi.org/10.1126/science.aav0080

Ambros, V., & Baltimore, D. (1978). Protein is linked to the 5’ end of poliovirus RNA by a phosphodiester linkage to tyrosine. *Journal of Biological Chemistry*, *253*(15), 5263–5266. https://doi.org/10.1016/s0021-9258(17)30361-7

Anderson, B. R., Muramatsu, H., Nallagatla, S. R., Bevilacqua, P. C., Sansing, L. H., Weissman, D., & Karikó, K. (2010). Incorporation of pseudouridine into mRNA enhances translation by diminishing PKR activation. *Nucleic Acids Research*, *38*(17), 5884–5892. https://doi.org/10.1093/nar/gkq347

Andries, O., Mc Cafferty, S., De Smedt, S. C., Weiss, R., Sanders, N. N., & Kitada, T. (2015). N1-methylpseudouridine-incorporated mRNA outperforms pseudouridine-incorporated mRNA by providing enhanced protein expression and reduced immunogenicity in mammalian cell lines and mice. *Journal of Controlled Release*, *217*, 337–344. https://doi.org/10.1016/j.jconrel.2015.08.051

Aoki, F. (2022). Zygotic gene activation in mice: profile and regulation. In *Journal of Reproduction and Development* (Vol. 68, Issue 2). https://doi.org/10.1262/jrd.2021-129

Apponi, L. H., Leung, S. W., Williams, K. R., Valentini, S. R., Corbett, A. H., & Pavlath, G. K. (2010). Loss of nuclear poly(A)-binding protein 1 causes defects in myogenesis and mRNA biogenesis. *Hum Mol Genet*. https://doi.org/10.1093/hmg/ddp569

Arnez, J. G., & Steitz, T. A. (1994). Crystal Structure of Unmodified tRNAGln Complexed with Glutaminyl-tRNA Synthetase and ATP Suggests a Possible Role for Pseudo-Uridines in

Stabilization of RNA Structure. *Biochemistry*, *33*(24), 7560–7567. https://doi.org/10.1021/bi00190a008

Ashour, J., Morrison, J., Laurent-Rolle, M., Belicha-Villanueva, A., Plumlee, C. R., Bernal-Rubio, D., Williams, K. L., Harris, E., Fernandez-Sesma, A., Schindler, C., & García-Sastre, A. (2010). Mouse STAT2 restricts early dengue virus replication. *Cell Host and Microbe*, *8*(5), 410–421. https://doi.org/10.1016/j.chom.2010.10.007

Asthana, S., Martin, H., Rupkey, J., Patel, S., Yoon, J., Keegan, A., & Mao, Y. (2022). The Physiological Roles of the Exon Junction Complex in Development and Diseases. In *Cells* (Vol. 11, Issue 7). MDPI. https://doi.org/10.3390/cells11071192

Avni, D., Biberman, Y., & Meyuhas, O. (1997). The 5′ terminal oligopyrimidine tract confers translational control on TOP mRNAs in a cell type-and sequence context-dependent manner. In *Nucleic Acids Research* (Vol. 25, Issue 5). Oxford University Press.

Baltimore, D. (1971). Expression of Animal Virus Genomes. In *BACTERIOLOGICAL REVIEWS* (Vol. 35, Issue 3). https://journals.asm.org/journal/br

Barlow, D. P., Randle, B. J., & Burke, D. C. (1984). Interferon synthesis in the early post-implantation mouse embryo. *Differentiation*, *27*(1–3), 229–235. https://doi.org/10.1111/j.1432-0436.1984.tb01433.x

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., & Chang, H. Y. (2014). M6A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell*, *15*(6), 707–719. https://doi.org/10.1016/j.stem.2014.09.019

Baumgartner, M., Olthof, A. M., Aquino, G. S., Hyatt, K. C., Lemoine, C., Drake, K., Sturrock, N., Nguyen, N., al Seesi, S., & Kanadia, R. N. (2018). Minor spliceosome inactivation causes microcephaly due to cell cycle defects and death of self-amplifying radial glial cells. *Development*. https://doi.org/10.1242/dev.166322

Bazak, L., Haviv, A., Barak, M., Jacob-Hirsch, J., Deng, P., Zhang, R., Isaacs, F. J., Rechavi, G., Li, J. B., Eisenberg, E., & Levanon, E. Y. (2014). A-to-I RNA editing occurs at over a hundred million genomic sites, located in a majority of human genes. *Genome Research*, *24*(3), 365–376. https://doi.org/10.1101/gr.164749.113

Beelman, C. A., & Parker, R. (1995). Degradation of mRNA in eukaryotes. *Cell*, *81*(2), 179–183. https://doi.org/10.1016/0092-8674(95)90326-7

Begik, O., Lucas, M. C., Pryszcz, L. P., Ramirez, J. M., Medina, R., Milenkovic, I., Cruciani, S., Liu, H., Graziele, H., Vieira, S., Sas-Chen, A., Mattick, J. S., Schwartz, S., & Novoa, E. M. (2021). Quantitative profiling of pseudouridylation dynamics in native RNAs with nanopore sequencing. *NAtuRE BIOtECHNOLOGy |*, *39*. https://doi.org/10.1038/s41587-021-00915-6

Bélanger, F., Stepinski, J., Darzynkiewicz, E., & Pelletier, J. (2010). Characterization of hMTr1, a Human Cap1 2-O-Ribose Methyltransferase. *Journal of Biological Chemistry*, *285*(43), 33037–33044. https://doi.org/10.1074/jbc.m110.155283

Berget, S. M., Moore, C., & Sharp, P. A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA* (adenovirus 2 mRNA processing/5' tails on mRNAs/electron microscopy of mRNA-DNA hybrids). In *Biochemistry* (Vol. 74, Issue 8). https://www.pnas.org

Birkedal, U., Christensen-Dalsgaard, M., Krogh, N., Sabarinathan, R., Gorodkin, J., & Nielsen, H. (2015). Profiling of Ribose Methylations in RNA by High-Throughput Sequencing. *Angewandte Chemie International Edition*, *54*(2), 451–455. https://doi.org/https://doi.org/10.1002/anie.201408362

Bobrovs, R., Kanepe, I., Narvaiss, N., Patetko, L., Kalnins, G., Sisovs, M., Bula, A. L., Grinberga, S., Boroduskis, M., Ramata-Stunda, A., Rostoks, N., Jirgensons, A., Tars, K., & Jaudzems, K. (2021). Discovery of SARS-CoV-2 Nsp14 and Nsp16 Methyltransferase Inhibitors by High-Throughput Virtual Screening. *Pharmaceuticals*, *14*(12), 1243. https://doi.org/10.3390/ph14121243

Boccaletto, P., MacHnicka, M. A., Purta, E., Pitkowski, P., Baginski, B., Wirecki, T. K., De Crécy-Lagard, V., Ross, R., Limbach, P. A., Kotter, A., Helm, M., & Bujnicki, J. M. (2018). MODOMICS: A database of RNA modification pathways. 2017 update. *Nucleic Acids Research*, *46*(D1), D303–D307. https://doi.org/10.1093/nar/gkx1030

Boone, R. F., & Moss, B. (1977). Methylated 5'-Terminal Sequences of Vaccinia Virus mRNA Species Made in Vivo at Early and Late Times after Infection'. In *VIROLOGY* (Vol. 79).

Boulias, K., Toczydlowska-Socha, D., Hawley, B. R., Liberman, N., Takashima, K., Zaccara, S., Guez, T., Vasseur, J.-J., Debart, F., Aravind, L., Jaffrey, S. R., & Greer, E. L. (2019). Identification of the m6Am Methyltransferase PCIF1 Reveals the Location and Functions of m6Am in the Transcriptome. *Molecular Cell*, *75*(3), 631-643.e8. https://doi.org/https://doi.org/10.1016/j.molcel.2019.06.006

Bowie, A. G., & Unterholzner, L. (2008). Viral evasion and subversion of pattern-recognition receptor signalling. *Nature Reviews Immunology*, *8*(12), 911–922. https://doi.org/10.1038/nri2436

Brusa, R., Zimmermann, F., Koh, D.-S., Feldmeyer, D., Gass, P., Seeburg, P. H., & Sprengel, R. (1995). Early-Onset Epilepsy and Postnatal Lethality Associated with an Editing-Deficient *GluR-B* Allele in Mice. *Science*, *270*(5242), 1677–1680. https://doi.org/10.1126/science.270.5242.1677

Burke, J. M., & Sullivan, C. S. (2017). DUSP11–An RNA phosphatase that regulates host and viral non-coding RNAs in mammalian cells. In *RNA Biology* (Vol. 14, Issue 11, pp. 1457–1465). Taylor and Francis Inc. https://doi.org/10.1080/15476286.2017.1306169

Carmody, S. R., & Wente, S. R. (2009). mRNA nuclear export at a glance. *Journal of Cell Science*, *122*, 1933–1937. https://doi.org/10.1242/jcs.041236

Casey, J. L., Koeller, D. M., Ramin, V. C., Klausner, R. D., & Harford, J. B. (1989). Iron regulation of transferrin receptor mRNA levels requires iron-responsive elements and a rapid turnover determinant in the 3' untranslated region of the mRNA. *EMBO Journal*, *8*(12), 3693–3699. https://doi.org/10.1002/j.1460-2075.1989.tb08544.x

Cekaite, L., Furset, G., Hovig, E., & Sioud, M. (2007). Gene Expression Analysis in Blood Cells in Response to Unmodified and 2'-Modified siRNAs Reveals TLR-dependent and Independent Effects. *Journal of Molecular Biology*, *365*(1), 90–108. https://doi.org/10.1016/j.jmb.2006.09.034

Chang, H., Lim, J., Ha, M., & Kim, V. N. (2014). TAIL-seq: Genome-wide determination of poly(A) tail length and 3' end modifications. *Molecular Cell*, *53*(6), 1044–1052. https://doi.org/10.1016/j.molcel.2014.02.007

Chen, C., Zhao, X., Kierzek, R., & Yu, Y.-T. (2010). A Flexible RNA Backbone within the Polypyrimidine Tract Is Required for U2AF 65 Binding and Pre-mRNA Splicing In Vivo. *MOLECULAR AND CELLULAR BIOLOGY*, *30*(17), 4108–4119. https://doi.org/10.1128/MCB.00531-10

Chen, L. L., Yang, L., & Carmichael, G. G. (2010). Molecular basis for an attenuated cytoplasmic dsRNA response in human embryonic stem cells. *Cell Cycle*, *9*(17), 3552–3564. https://doi.org/10.4161/cc.9.17.12792

Chen, Y. G., & Hur, S. (2022). Cellular origins of dsRNA, their recognition and consequences. In *Nature Reviews Molecular Cell Biology* (Vol. 23, Issue 4, pp. 286–301). Nature Research. https://doi.org/10.1038/s41580-021-00430-1

Chen, Z., Li, Y., & Krug, R. M. (1999). Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery. *EMBO Journal*, *18*(8), 2273–2283. https://doi.org/10.1093/emboj/18.8.2273

Chen, Z.-H., Wu, Q., Lu, J., QI, J., Liu, Z., Wang, Y., Wu, X., Chen, Y., Zeng, Z.-L., Ju, H., & Xu, R. (2018). Effect of knockdown of eukaryotic initiation factor 4A2 (eIF4A2) on growth, metastasis, and oxaliplatin sensitivity in colorectal cancer. *Journal of Clinical Oncology*, *36*(15_suppl), e15658–e15658. https://doi.org/10.1200/JCO.2018.36.15_suppl.e15658

Cheng, H., Dufu, K., Lee, C. S., Hsu, J. L., Dias, A., & Reed, R. (2006). Human mRNA Export Machinery Recruited to the 5' End of mRNA. *Cell*, *127*(7), 1389–1400. https://doi.org/10.1016/j.cell.2006.10.044

Chi, J. R., Yu, Z. H., Liu, B. W., Zhang, D., Ge, J., Yu, Y., & Cao, X. C. (2019). SNHG5 Promotes Breast Cancer Proliferation by Sponging the miR-154-5p/PCNA Axis. *Molecular Therapy - Nucleic Acids*, *17*, 138–149. https://doi.org/10.1016/j.omtn.2019.05.013

Choi, J., Indrisiunaite, G., Demirci, H., Ieong, K. W., Wang, J., Petrov, A., Prabhakar, A., Rechavi, G., Dominissini, D., He, C., Ehrenberg, M., & Puglisi, J. D. (2018). 2'-O-methylation in mRNA disrupts tRNA decoding during translation elongation. *Nature Structural and Molecular Biology*, *25*(3), 208–216. https://doi.org/10.1038/s41594-018-0030-z

Choi, K. H. (2012). *Viral Polymerases* (pp. 267–304). https://doi.org/10.1007/978-1-4614-0980-9_12

Choi, U. Y., Kang, J. S., Hwang, Y. S., & Kim, Y. J. (2015). Oligoadenylate synthase-like (OASL) proteins: dual functions and associations with diseases. In *Experimental and Molecular Medicine* (Vol. 47, Issue 3). Springer Nature. https://doi.org/10.1038/EMM.2014.110

Chow, L. T., & Broker, T. R. (1978). The spliced structures of adenovirus 2 fiber message and the other late mRNAs. *Cell*, *15*(2), 497–510. https://doi.org/10.1016/0092-8674(78)90019-3

Clancy, M. J., Shambaugh, M. E., Timpte, C. S., & Bokar, J. A. (2002). Induction of sporulation in Saccharomyces cerevisiae leads to the formation of N 6-methyladenosine in mRNA: a potential mechanism for the activity of the IME4 gene. *NAR*, 4509–4518. https://doi.org/10.1093/nar/gkf573.

Cobb, M. (2017). 60 years ago, Francis Crick changed the logic of biology. *PLOS Biology*, *15*(9), e2003243. https://doi.org/10.1371/journal.pbio.2003243

Cockman, E., Anderson, P., & Ivanov, P. (2020). TOP mRNPs: Molecular Mechanisms and Principles of Regulation. *Biomolecules*, *10*(7), 969. https://doi.org/10.3390/biom10070969

COHN, W. E., & VOLKIN, E. (1951). Nucleoside-5'-Phosphates from Ribonucleic Acid. *Nature*, *167*(4247), 483–484. https://doi.org/10.1038/167483a0

Cooper, G. M. (1994). *Expression and Function of c-mos in Mammalian Germ Cells* (Vol. 3, pp. 127–148). Academic Press. https://doi.org/10.1016/S1064-2722(08)60008-0

Coots, R. A., Liu, X. M., Mao, Y., Dong, L., Zhou, J., Wan, J., Zhang, X., & Qian, S. B. (2017). m6A Facilitates eIF4F-Independent mRNA Translation. *Molecular Cell*, *68*(3), 504-514.e7. https://doi.org/10.1016/J.MOLCEL.2017.10.002

Coppola, J. A., Field, A. S., & Luse, D. S. (1983). *Promoter-proximal pausing by RNA polymerase II in vitro: Transcripts shorter than 20 nucleotides are not capped (initiation of transcription)* (Vol. 80). https://www.pnas.org

Crawford, N. M., & Baltimore, D. (1983). *VPg-pUpU in poliovirus-infected cells (RNA replication/picornavirus)* (Vol. 80). https://www.pnas.org

Crick, F. (1970). Central Dogma of Molecular Biology. In *NATURE* (Vol. 227).

Daffis, S., Szretter, K. J., Schriewer, J., Li, J., Youn, S., Errett, J., Lin, T.-Y., Schneller, S., Zust, R., Dong, H., Thiel, V., Sen, G. C., Fensterl, V., Klimstra, W. B., Pierson, T. C., Buller, R. M., Jr, M. G., Shi, P.-Y., & Diamond, M. S. (2010). 2^-O methylation of the viral mRNA cap evades host restriction by IFIT family members. *Nature*, *468*(7322), 452–456. https://doi.org/10.1038/nature09489

Dai, Q., Moshitch-Moshkovitz, S., Han, D., Kol, N., Amariglio, N., Rechavi, G., Dominissini, D., & He, C. (2017). *nm-seq maps 2'-O-methylation sites in human mrna with base precision*. *14*, 695. https://doi.org/10.1038/Nmeth.4294

Davis, D. R. (1995). Stabilization of RNA stacking by pseudouridine. In *Nucleic Acids Research* (Vol. 23, Issue 24).

Davis, D. R., & Poulter, C. D. (1991). Proton-nitrogen-15 NMR studies of Escherichia coli tRNAPhe from HisT mutants: a structural role for pseudouridine. *Biochemistry*, *30*(17), 4223–4231. https://doi.org/10.1021/bi00231a017

DAVIS, F. F., & ALLEN, F. W. (1957). Ribonucleic acids from yeast which contain a fifth nucleotide. *The Journal of Biological Chemistry*, *227*(2), 907–915. https://doi.org/10.1016/s0021-9258(18)70770-9

De Vlugt, C., Sikora, D., & Pelchat, M. (2018). Insight into influenza: A virus cap-snatching. In *Viruses* (Vol. 10, Issue 11). MDPI AG. https://doi.org/10.3390/v10110641

Deddouche, S., Matt, N., Budd, A., Mueller, S., Kemp, C., Galiana-Arnoux, D., Dostert, C., Antoniewski, C., Hoffmann, J. A., & Imler, J.-L. (2008). The DExD/H-box helicase Dicer-2 mediates the induction of antiviral activity in drosophila A R T I C L E S. *NATURE IMMUNOLOGY*, *9*. https://doi.org/10.1038/ni.1664

Delorme-Axford, E., Donker, R. B., Mouillet, J. F., Chu, T., Bayer, A., Ouyang, Y., Wang, T., Stolz, D. B., Sarkar, S. N., Morelli, A. E., Sadovsky, Y., & Coyne, C. B. (2013). Human placental trophoblasts confer viral resistance to recipient cells. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(29), 12048–12053. https://doi.org/10.1073/pnas.1304718110

Deng, P., Khan, A., Jacobson, D., Sambrani, N., McGurk, L., Li, X., Jayasree, A., Hejatko, J., Shohat-Ophir, G., O'Connell, M. A., Li, J. B., & Keegan, L. P. (2020a). Adar RNA editing-dependent and -independent effects are required for brain and innate immune functions in Drosophila. *Nature Communications*, *11*(1), 1580. https://doi.org/10.1038/s41467-020-15435-1

Deng, P., Khan, A., Jacobson, D., Sambrani, N., McGurk, L., Li, X., Jayasree, A., Hejatko, J., Shohat-Ophir, G., O'Connell, M. A., Li, J. B., & Keegan, L. P. (2020b). Adar RNA editing-dependent and -independent effects are required for brain and innate immune functions in Drosophila. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-15435-1

Despic, V., & Jaffrey, S. R. (2023). mRNA ageing shapes the Cap2 methylome in mammalian mRNA. *Nature*, *614*(7947), 358–366. https://doi.org/10.1038/s41586-022-05668-z

Desrosiers, R., Friderici, K., & Rottman, F. (1974). Identification of Methylated Nucleosides in Messenger RNA from Novikoff Hepatoma Cells. In *Proceedings of the National Academy of Sciences of the United States of America* (Vol. 71, Issue 10, pp. 3971–3975).

Devarkar, S. C., Wang, C., Miller, M. T., Ramanathan, A., Jiang, F., Khan, A. G., Patel, S. S., & Marcotrigiano, J. (2016). Structural basis for m7G recognition and 2^-O-methyl discrimination in capped RNAs by the innate immune receptor RIG-I. *Proceedings of the National Academy of Sciences*, *113*(3), 596–601. https://doi.org/10.1073/pnas.1515152113

Dever, T. E., Dinman, J. D., & Green, R. (2018). Translation elongation and recoding in eukaryotes. *Cold Spring Harbor Perspectives in Biology*, *10*(8). https://doi.org/10.1101/cshperspect.a032649

Dix, T. C., Haussmann, I. U., Brivio, S., Nallasivan, M. P., HadzHiev, Y., Müller, F., Müller, B., Pettitt, J., & Soller, M. (2022). CMTr mediated 2′- *O* -ribose methylation status of cap-adjacent nucleotides across animals. *RNA*, *28*(10), 1377–1390. https://doi.org/10.1261/rna.079317.122

Dobson, C. M. (2003). Protein folding and misfolding. *Nature*, *426*(6968), 884–890. https://doi.org/10.1038/nature02261

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., & Rechavi, G. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature*, *485*(7397), 201–206. https://doi.org/10.1038/nature11112

Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Shay Ben-Haim, M., Dai, Q., Di Segni, A., Salmon-Divon, M., Clark, W. C., Zheng, G., Pan, T., Solomon, O., Eyal, E., Hershkovitz, V., Han, D., Doré, L. C., Amariglio, N., Rechavi, G., & He, C. (2016). *The dynamic N 1-methyladenosine methylome in eukaryotic messenger RNA*. https://doi.org/10.1038/nature16998

Donati, G., Peddigari, S., Mercer, C. A., & Thomas, G. (2013). 5S Ribosomal RNA Is an Essential Component of a Nascent Ribosomal Precursor Complex that Regulates the Hdm2-p53 Checkpoint. *Cell Reports*, *4*(1), 87–98. https://doi.org/10.1016/j.celrep.2013.05.045

Dong, H., Chang, D. C., Hua, M. H. C., Lim, S. P., Chionh, Y. H., Hia, F., Lee, Y. H., Kukkaro, P., Lok, S. M., Dedon, P. C., & Shi, P. Y. (2012). 2′-O methylation of internal adenosine by flavivirus NS5 methyltransferase. *PLoS Pathogens*, *8*(4). https://doi.org/10.1371/journal.ppat.1002642

Dong, H., Chang, D. C., Xie, X., Toh, Y. X., Chung, K. Y., Zou, G., Lescar, J., Lim, S. P., & Shi, P. Y. (2010). Biochemical and genetic characterization of dengue virus methyltransferase. *Virology*, *405*(2), 568–578. https://doi.org/10.1016/j.virol.2010.06.039

Dong, Z.-W., Shao, P., Diao, L.-T., Zhou, H., Yu, C.-H., & Qu, L.-H. (2012). RTL-P: a sensitive approach for detecting sites of 2′-O-methylation in RNA molecules. *Nucleic Acids Research*, *40*(20), e157–e157. https://doi.org/10.1093/nar/gks698

Dönmez, G., Hartmuth, K., & Lührmann, R. (2004). Modified nucleotides at the 5′ end of human U2 snRNA are required for spliceosomal E-complex formation. *RNA*, *10*(12), 1925–1933. https://doi.org/10.1261/rna.7186504

Drazkowska, K., Tomecki, R., Warminski, M., Baran, N., Cysewski, D., Depaix, A., Kasprzyk, R., Kowalska, J., Jemielity, J., & Sikorski, P. J. (2022). 2′-O-Methylation of the second transcribed nucleotide within the mRNA 5′ cap impacts the protein production level in a cell-specific manner and contributes to RNA immune evasion. *Nucleic Acids Research*, *50*(16), 9051–9071. https://doi.org/10.1093/nar/gkac722

DRs FRANCOIS GROS, B., Hiatt, H., Kurland, G., Risebrough, R. W., & D Watson, D. J. (1959). The Nudeic Acids. In *Proc. U.S. Nat. Acad. Sci* (Vol. 1). Academic Press.

Duncan-Lowey, B., Tal, N., Johnson, A. G., Rawson, S., Mayer, M. L., Doron, S., Millman, A., Melamed, S., Fedorenko, T., Kacen, A., Brandis, A., Mehlman, T., Amitai, G., Sorek, R., & Kranzusch, P. J. (2023). Cryo-EM structure of the RADAR supramolecular anti-phage defense complex. *Cell*, *186*(5), 987-998.e15. https://doi.org/10.1016/j.cell.2023.01.012

Dunn, D. B. (1959). Additional components in ribonucleic acid of rat-liver fractions. *Biochimica et Biophysica Acta*, *34*(C), 286–288. https://doi.org/10.1016/0006-3002(59)90274-4

Durbin, A. F., Wang, C., Marcotrigiano, J., & Gehrke, L. (2016). RNAs containing modified nucleotides fail to trigger RIG-I conformational changes for innate immune signaling. *MBio*, *7*(5). https://doi.org/10.1128/mBio.00833-16

Eberle, F., Gießler, K., Deck, C., Heeg, K., Peter, M., Richert, C., & Dalpke, A. H. (2008). *Modifications in Small Interfering RNA That Separate Immunostimulation from RNA Interference 1*. http://journals.aai.org/jimmunol/article-pdf/180/5/3229/1272127/zim00508003229.pdf

Edery, I., & Sonenberg, N. (1985). Cap-dependent RNA splicing in a HeLa nuclear extract (cap analogues). In *Biochemistry* (Vol. 82).

Elisa Izaurralde, Joe Lewis, Chiara Gamberi, Artur Jarmolowski, Caroline McGuigan, & Iain W. Mattaj. (1995). *A cap-binding protein complex mediating U snRNA export*.

Ellgaard, L., McCaul, N., Chatsisvili, A., & Braakman, I. (2016). Co- and Post-Translational Protein Folding in the ER. *Traffic*, *17*(6), 615–638. https://doi.org/10.1111/tra.12392

Elliott, B. A., Ho, H.-T., Ranganathan, S. V., Vangaveti, S., Ilkayeva, O., Abou Assi, H., Choi, A. K., Agris, P. F., & Holley, C. L. (2019). Modification of messenger RNA by 2′-O-methylation regulates gene expression in vivo. *Nature Communications*, *10*(1), 3401. https://doi.org/10.1038/s41467-019-11375-7

Elsaid, M. F., Chalhoub, N., Ben-Omran, T., Kumar, P., Kamel, H., Ibrahim, K., Mohamoud, Y., Al-Dous, E., Al-Azwani, I., Malek, J. A., Suhre, K., Ross, M. E., & Aleem, A. A. (2017). Mutation in

noncoding RNA RNU12 causes early onset cerebellar ataxia. *Annals of Neurology*, *81*(1), 68–78. https://doi.org/10.1002/ana.24826

Elvin, J. (1998). Mouse models of ovarian failure. *Reviews of Reproduction*, *3*(3), 183–195. https://doi.org/10.1530/ror.0.0030183

Ezkurdia, I., Rodriguez, J. M., Carrillo-De Santa Pau, E., Vázquez, J., Valencia, A., & Tress, M. L. (2015). Most highly expressed protein-coding genes have a single dominant isoform. *Journal of Proteome Research*, *14*(4), 1880–1887. https://doi.org/10.1021/pr501286b

Farlow, J., Ichou, M. A., Huggins, J., & Ibrahim, S. (2010). Open Access RESEARCH Comparative whole genome sequence analysis of wild-type and cidofovir-resistant monkeypoxvirus. In *Virology Journal* (Vol. 7). http://www.virologyj.com/content/7/1/110

Fayomi, A. P., & Orwig, K. E. (2018). Spermatogonial stem cells and spermatogenesis in mice, monkeys and men. *Stem Cell Research*, *29*, 207–214. https://doi.org/10.1016/J.SCR.2018.04.009

Flamand, M. N., & Meyer, K. D. (2022). M6A and YTHDF proteins contribute to the localization of select neuronal mRNAs. *Nucleic Acids Research*, *50*(8), 4464–4483. https://doi.org/10.1093/nar/gkac251

Fleith, R. C., Mears, H. V, Leong, X. Y., Sanford, T. J., Emmott, E., Graham, S. C., Mansur, D. S., & Sweeney, T. R. (2018). IFIT3 and IFIT2/3 promote IFIT1-mediated translation inhibition by enhancing binding to non-self RNA. *Nucleic Acids Research*, *46*(10), 5269–5285. https://doi.org/10.1093/nar/gky191

FURUICHI, Y. (2015). Discovery of m⁷G-cap in eukaryotic mRNAs. *Proceedings of the Japan Academy, Series B*, *91*(8), 394–409. https://doi.org/10.2183/pjab.91.394

Furuichi, Y., Morgan, M., Shatkin, A. J., Jelinek, W., Salditt-Georgieff, M., & Darnell, J. E. (1975). Methylated, blocked 5 termini in HeLa cell mRNA. *Proceedings of the National Academy of Sciences*, *72*(5), 1904–1908. https://doi.org/10.1073/pnas.72.5.1904

Fustin, J. M., Doi, M., Yamaguchi, Y., Hida, H., Nishimura, S., Yoshida, M., Isagawa, T., Morioka, M. S., Kakeya, H., Manabe, I., & Okamura, H. (2013). XRNA-methylation-dependent RNA processing controls the speed of the circadian clock. *Cell*, *155*(4), 793. https://doi.org/10.1016/j.cell.2013.10.026

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., & Okamura, H. (2018). Two *Ck1δ* transcripts regulated by m6A methylation code for two antagonistic kinases in the control of the circadian clock. *Proceedings of the National Academy of Sciences*, *115*(23), 5980–5985. https://doi.org/10.1073/pnas.1721371115

Gale, M., Blakely, C. M., Kwieciszewski, B., Tan, S.-L., Dossett, M., Tang, N. M., Korth, M. J., Polyak, S. J., Gretch, D. R., & Katze, M. G. (1998). Control of PKR Protein Kinase by Hepatitis C Virus Nonstructural 5A Protein: Molecular Mechanisms of Kinase Regulation. In *MOLECULAR AND CELLULAR BIOLOGY* (Vol. 18, Issue 9).

Galipon, J., Ishii, R., Suzuki, Y., Tomita, M., & Ui-Tei, K. (2017). Differential Binding of Three Major Human ADAR Isoforms to Coding and Long Non-Coding Transcripts. *Genes*, *8*(2), 68. https://doi.org/10.3390/genes8020068

Gallie, D. R. (1991). The cap and poly(A) tail function synergistically to regulate mRNA translational efficiency. *Genes & Development*, *5*(11), 2108–2116. https://doi.org/10.1101/gad.5.11.2108

Galloway, A., Atrih, A., Grzela, R., Darzynkiewicz, E., Ferguson, M. A. J., & Cowling, V. H. (2020). CAP-MAP: cap analysis protocol with minimal analyte processing, a rapid and sensitive approach to analysing mRNA cap structures. *Open Biology*, *10*(2). https://doi.org/10.1098/rsob.190306

Galloway, A., Kaskar, A., Ditsova, D., Atrih, A., Yoshikawa, H., Gomez-Moreira, C., Suska, O., Warminski, M., Grzela, R., Lamond, A. I., Darzynkiewicz, E., Jemielity, J., & Cowling, V. H. (2021). Upregulation of RNA cap methyltransferase RNMT drives ribosome biogenesis during T cell activation. *Nucleic Acids Research*, *49*(12), 6722–6738. https://doi.org/10.1093/nar/gkab465

Garneau, N. L., Wilusz, J., & Wilusz, C. J. (2007). The highways and byways of mRNA decay. In *Nature Reviews Molecular Cell Biology* (Vol. 8, Issue 2, pp. 113–126). https://doi.org/10.1038/nrm2104

Gebhardt, A., Habjan, M., Benda, C., Meiler, A., Haas, D. A., Hein, M. Y., Mann, A., Mann, M., Habermann, B., & Pichlmair, A. (2015). ARTICLE mRNA export through an additional cap-binding complex consisting of NCBP1 and NCBP3. *Nature Communications*. https://doi.org/10.1038/ncomms9192

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., … Hanna, J. H. (2015). m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science*, *347*(6225), 1002–1006. https://doi.org/10.1126/science.1261417

Gonatopoulos-Pournatzis, T., & Cowling, V. H. (2014). Cap-binding complex (CBC). *Biochem. J*, *457*, 231–242. https://doi.org/10.1042/BJ20131214

Goubau, D., Schlee, M., Deddouche, S., Pruijssers, A. J., Zillinger, T., Goldeck, M., Schuberth, C., Van Der Veen, A. G., Fujimura, T., Rehwinkel, J., Iskarpatyoti, J. A., Barchet, W., Ludwig, J., Dermody, T. S., Hartmann, G., & Reis E Sousa, C. (2014). Antiviral immunity via RIG-I-mediated recognition of RNA bearing 59-diphosphates. *Nature*. https://doi.org/10.1038/nature13590

Gromadzka, A. M., Steckelberg, A.-L., Singh, K. K., Hofmann, K., & Gehring, N. H. (2016). A short conserved motif in ALYREF directs cap-and EJC-dependent assembly of export complexes on spliced mRNAs. *Nucleic Acids Research*, *44*(5), 2348–2361. https://doi.org/10.1093/nar/gkw009

Groza, T., Gomez, F. L., Mashhadi, H. H., Muñoz-Fuentes, V., Gunes, O., Wilson, R., Cacheiro, P., Frost, A., Keskivali-Bond, P., Vardal, B., McCoy, A., Cheng, T. K., Santos, L., Wells, S., Smedley, D., Mallon, A.-M., & Parkinson, H. (2023). The International Mouse Phenotyping Consortium: comprehensive knockout phenotyping underpinning the study of human disease. *Nucleic Acids Research*, *51*(D1), D1038–D1045. https://doi.org/10.1093/nar/gkac972

Habjan, M., Hubel, P., Lacerda, L., Benda, C., Holze, C., Eberl, C. H., Mann, A., Kindler, E., Gil-Cruz, C., Ziebuhr, J., Thiel, V., & Pichlmair, A. (2013). Sequestration by IFIT1 Impairs Translation of 2'O-unmethylated Capped RNA. *PLoS Pathogens*, *9*(10). https://doi.org/10.1371/journal.ppat.1003663

Hartner, J. C., Walkley, C. R., Lu, J., & Orkin, S. H. (2008). ADAR1 is essential for the maintenance of hematopoiesis and suppression of interferon signaling. *Nature Immunology*, *10*(1), 109–115. https://doi.org/10.1038/ni.1680

Hassan, D., Ariyur, A., Vidhur Daulatabad, S., Mir, Q., & Chandra Janga, S. (2022). Nm-Nano: Predicting 2′-O-Methylation (Nm) Sites in Nanopore RNA Sequencing Data. *BioRxiv*. https://doi.org/10.1101/2022.01.03.473214

Haussmann, I. U., Bodi, Z., Sanchez-Moran, E., Mongan, N. P., Archer, N., Fray, R. G., & Soller, M. (2016). M6 A potentiates Sxl alternative pre-mRNA splicing for robust Drosophila sex determination. *Nature*, *540*(7632), 301–304. https://doi.org/10.1038/nature20577

Haussmann, I. U., Wu, Y., Nallasivan, M. P., Archer, N., Bodi, Z., Hebenstreit, D., Waddell, S., Fray, R., & Soller, M. (2022). CMTr cap-adjacent 2′-O-ribose mRNA methyltransferases are required for reward learning and mRNA localization to synapses. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-28549-5

He, P. C., Wei, J., Dou, X., Harada, B. T., Zhang, Z., Ge, R., Liu, C., Zhang, L.-S., Yu, X., Wang, S., Lyu, R., Zou, Z., Chen, M., & He, C. (2023). Exon architecture controls mRNA m6A suppression and gene expression. *Science*, *379*(6633), 677–682. https://doi.org/10.1126/science.abj9090

He, P., Zhang, C., Ji, Y., Ge, M.-K., Yu, Y., Zhang, N., Yang, S., Yu, J.-X., Shen, S.-M., & Chen, G.-Q. (2022). Epithelial cells-enriched lncRNA SNHG8 regulates chromatin condensation by binding to Histone H1s. *Cell Death & Differentiation*, *29*(8), 1569–1581. https://doi.org/10.1038/s41418-022-00944-x

Heidemann, M., Hintermair, C., Voß, K., & Eick, D. (2013). Dynamic phosphorylation patterns of RNA polymerase II CTD during transcription. In *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* (Vol. 1829, Issue 1, pp. 55–62). https://doi.org/10.1016/j.bbagrm.2012.08.013

Heisenberg, C. P., & Solnica-Krezel, L. (2008). Back and forth between cell fate specification and movement during vertebrate gastrulation. In *Current Opinion in Genetics and Development* (Vol. 18, Issue 4, pp. 311–316). https://doi.org/10.1016/j.gde.2008.07.011

Hertzog, J., Dias Junior, A. G., Rigby, R. E., Donald, C. L., Mayer, A., Sezgin, E., Song, C., Jin, B., Hublitz, P., Eggeling, C., Kohl, A., & Rehwinkel, J. (2018). Infection with a Brazilian isolate of Zika virus generates RIG-I stimulatory RNA and the viral NS5 protein blocks type I IFN induction and signaling. *European Journal of Immunology*, *48*(7), 1120–1136. https://doi.org/10.1002/eji.201847483

Higuchi, M., Maas, S., Single, F. N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., & Seeburg, P. H. (2000). Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. *Nature*, *406*(6791), 78–81. https://doi.org/10.1038/35017558

Hogan, A. K., & Foltz, D. R. (2021). Reduce, Retain, Recycle: Mechanisms for Promoting Histone Protein Degradation versus Stability and Retention. *Molecular and Cellular Biology*, *41*(6). https://doi.org/10.1128/mcb.00007-21

Hopkins, K. C., Tartell, M. A., Herrmann, C., Hackett, B. A., Taschuk, F., Panda, D., Menghani, S. V., Sabin, L. R., & Cherry, S. (2015). Virus-induced translational arrest through 4EBP1/2-dependent decay of 5′-TOP mRNAs restricts viral infection. *Proceedings of the National Academy of*

Sciences of the United States of America, *112*(22), E2920–E2929. https://doi.org/10.1073/pnas.1418805112

Hornung, V., Ellegast, J., Kim, S., Brzozka, K., Jung, A., Kato, H., Poeck, H., Akira, S., Conzelmann, K.-K., Schlee, M., Endres, S., & Hartmann, G. (2006). 5-Triphosphate RNA Is the Ligand for RIG-I. *Science*, *314*(5801), 994–997. https://doi.org/10.1126/science.1132505

Hoshi, T., Zagotta, W. N., Aldrich, R. W., Lsacoff, E. Y., Papazian, D. M., Timpe, L. C., Jan, Y. N., Jan, L. Y., Greenblatt, R. E., & Noda, M. (1984). Internal initiation of translation mediated by the 5' leader of a cellular mRNA. In *12. Catterall, W. A. A. Rev. Biochem* (Vol. 70, Issue 11).

Huang, T., Chen, W., Liu, J., Gu, N., & Zhang, R. (2019). Genome-wide identification of mRNA 5-methylcytosine in mammals. *Nature Structural & Molecular Biology*, *26*(5), 380–388. https://doi.org/10.1038/s41594-019-0218-x

Inesta-Vaquera, F., Chaugule, V. K., Galloway, A., Chandler, L., Rojas-Fernandez, A., Weidlich, S., Peggie, M., & Cowling, V. H. (2018). DHX15 regulates CMTR1-dependent gene expression and cell proliferation. *Life Science Alliance*, *1*(3). https://doi.org/10.26508/lsa.201800092

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S., & Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science*, *324*(5924), 218–223. https://doi.org/10.1126/science.1168978

Jain, M., Olsen, H. E., Paten, B., & Akeson, M. (2016). The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology*, *17*(1). https://doi.org/10.1186/s13059-016-1103-0

Jenjaroenpun, P., Wongsurawat, T., Wadley, T. D., Wassenaar, T. M., Liu, J., Dai, Q., Wanchai, V., Akel, N. S., Jamshidi-Parsian, A., Franco, A. T., Boysen, G., Jennings, M. L., Ussery, D. W., He, C., & Nookaew, I. (2021). Decoding the epitranscriptional landscape from native RNA sequences. *Nucleic Acids Research*, *49*(2), 7. https://doi.org/10.1093/nar/gkaa620

Juan-Mateu, J., & Valcárcel, J. (2023). Minority report: The minor spliceosome as a novel cancer vulnerability factor. *Molecular Cell*, *83*(12), 1958–1960. https://doi.org/10.1016/J.MOLCEL.2023.05.027

Junqueira, S. C., dos Santos Coelho, I., Lieberknecht, V., Cunha, M. P., Calixto, J. B., Rodrigues, A. L. S., Santos, A. R. S., & Dutra, R. C. (2016). Inosine, an Endogenous Purine Nucleoside, Suppresses Immune Responses and Protects Mice from Experimental Autoimmune Encephalomyelitis: a Role for A2A Adenosine Receptor. *Molecular Neurobiology*, *54*(5), 3271–3285. https://doi.org/10.1007/s12035-016-9893-3

Jureka, A. S., Kleinpeter, A. B., Tipper, J. L., Harrod, K. S., & Petit, C. M. (2020). The influenza NS1 protein modulates RIG-I activation via a strain-specific direct interaction with the second CARD of RIG-I. *Journal of Biological Chemistry*, *295*(4), 1153–1164. https://doi.org/10.1074/jbc.RA119.011410

Kampen, K. R., Sulima, S. O., Vereecke, S., & de Keersmaecker, K. (2021). Hallmarks of ribosomopathies. In *Nucleic Acids Research* (Vol. 48, Issue 3, pp. 1013–1028). Oxford University Press. https://doi.org/10.1093/NAR/GKZ637

Karikó, K., Buckstein, M., Ni, H., & Weissman, D. (2005). Suppression of RNA Recognition by Toll-like Receptors: The Impact of Nucleoside Modification and the Evolutionary Origin of RNA. *Immunity*, *23*(2), 165–175. https://doi.org/10.1016/j.immuni.2005.06.008

Karikó, K., Muramatsu, H., Welsh, F. A., Ludwig, J., Kato, H., Akira, S., & Weissman, D. (2008). Incorporation of pseudouridine into mRNA yields superior nonimmunogenic vector with increased translational capacity and biological stability. *Molecular Therapy*, *16*(11), 1833–1840. https://doi.org/10.1038/mt.2008.200

Kastern, W. H., & Berry, S. J. (1976). *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS NON-METHYLATED GUANOSINE AS THE 5′ TERMINUS OF CAPPED mRNA FROM INSECT OOCYTES* (Vol. 71, Issue 1).

Kastern, W. H., Swindlehurst, M., Aaron, C., Hooper, J., & Berry, S. J. (1982). *Control of mRNA Translation in Oocytes and Developing Embryos of Giant Moths I. Function of the 5′ Terminal "Cap" in the Tobacco Hornworm, Manduca sexta*.

Kazzi, P. El, Rabah, N., Chamontin, C., Poulain, L., Ois Ferron, F., Oise Debart, F., Canard, B., Doroth´, D., Missé, D., Missé, M., Coutard, B., Nisole, S., & Decroly, E. (2023). NAR Breakthrough Article Internal RNA 2 O-methylation in the HIV-1 genome counteracts ISG20 nuclease-mediated antiviral effect. *Nucleic Acids Research*, *51*(6), 2501–2515. https://doi.org/10.1093/nar/gkac996

Keenan, E. K., Zachman, D. K., & Hirschey, M. D. (2021). Discovering the landscape of protein modifications. *Molecular Cell*, *81*(9), 1868–1878. https://doi.org/10.1016/J.MOLCEL.2021.03.015

Kersey, P. J., Allen, J. E., Allot, A., Barba, M., Boddu, S., Bolt, B. J., Carvalho-Silva, D., Christensen, M., Davis, P., Grabmueller, C., Kumar, N., Liu, Z., Maurel, T., Moore, B., McDowall, M. D., Maheswari, U., Naamati, G., Newman, V., Ong, C. K., … Yates, A. (2018). Ensembl Genomes 2018: An integrated omics infrastructure for non-vertebrate species. *Nucleic Acids Research*, *46*(D1), D802–D808. https://doi.org/10.1093/nar/gkx1011

Kierzek, E., Malgowska, M., Lisowiec, J., Turner, D. H., Gdaniec, Z., & Kierzek, R. (2014). The contribution of pseudouridine to stabilities and structure of RNAs. *Nucleic Acids Research*, *42*(5), 3492–3501. https://doi.org/10.1093/nar/gkt1330

Kojima, Y., Tam, O. H., & Tam, P. P. L. (2014). Timing of developmental events in the early mouse embryo. *Seminars in Cell & Developmental Biology*, *34*, 65–75. https://doi.org/10.1016/J.SEMCDB.2014.06.010

Konarska, M. M., Padgett, R. A., & Sharp, P. A. (1984). Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell*, *38*(3), 731–736. https://doi.org/10.1016/0092-8674(84)90268-x

Kozak, M. (1981). *Nucleic Acids Research Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic rnbosomes* (Vol. 9).

Krieg, A. M. (2002). CpG Motifs in Bacterial DNA and Their Immune Effects. *Annual Review of Immunology*, *20*(1), 709–760. https://doi.org/10.1146/annurev.immunol.20.100301.064842

Krogh, N., Kongsbak-Wismann, M., Geisler, C., & Nielsen, H. (2017). Substoichiometric ribose methylations in spliceosomal snRNAs. *Organic and Biomolecular Chemistry*, *15*(42), 8872–8876. https://doi.org/10.1039/c7ob02317k

Krug, R. M., Morgan, M. A., & Shatkin, A. J. (1976). Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. *Journal of Virology*, *20*(1), 45–53. https://doi.org/10.1128/jvi.20.1.45-53.1976

Kuge, H., Brownlee, G. G., Gershon, P. D., & Richter, J. D. (1998). Cap ribose methylation of c-mos mRNA stimulates translation and oocyte maturation in Xenopus laevis. In *Nucleic Acids Research* (Vol. 26, Issue 13).

Kumar, P., Sweeney, T. R., Skabkin, M. A., Skabkina, O. V, Hellen, C. U. T., & Pestova, T. V. (2013). Inhibition of translation by IFIT family members is determined by their ability to interact selectively with the 5^-terminal regions of cap0-, cap1- and 5^ppp- mRNAs. *Nucleic Acids Research*, *42*(5), 3228–3245. https://doi.org/10.1093/nar/gkt1321

Kumar, S., & Mohapatra, T. (2021). Deciphering Epitranscriptome: Modification of mRNA Bases Provides a New Perspective for Post-transcriptional Regulation of Gene Expression. *Frontiers in Cell and Developmental Biology*, *9*. https://doi.org/10.3389/fcell.2021.628415

Labeit, S., & Kolmerer, B. (1995). Titins: Giant Proteins in Charge of Muscle Ultrastructure and Elasticity. *Science*, *270*(5234), 293–296. https://doi.org/10.1126/science.270.5234.293

Lan, J., Rajan, N., Bizet, M., Penning, A., Singh, N. K., Guallar, D., Calonne, E., Li Greci, A., Bonvin, E., Deplus, R., Hsu, P. J., Nachtergaele, S., Ma, C., Song, R., Fuentes-Iglesias, A., Hassabi, B., Putmans, P., Mies, F., Menschaert, G., … Fuks, F. (2020). Functional role of Tet-mediated RNA hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature Communications*, *11*(1), 4956. https://doi.org/10.1038/s41467-020-18729-6

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., … Hanna, J. H. (2020). Context-dependent functional compensation between Ythdf m 6 A reader proteins. *Genesde*. https://doi.org/10.1101/gad.340695.120

Laudenbach, B. T., Krey, K., Emslander, Q., Andersen, L. L., Reim, A., Scaturro, P., Mundigl, S., Dächert, C., Manske, K., Moser, M., Ludwig, J., Wohlleber, D., Kröger, A., Binder, M., & Pichlmair, A. (2021). NUDT2 initiates viral RNA degradation by removal of 5'-phosphates. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-27239-y

Lee, H.-C., Chathuranga, K., & Lee, J.-S. (2019). Intracellular sensing of viral genomes and viral evasion. *Experimental & Molecular Medicine*, *51*(12), 1–13. https://doi.org/10.1038/s12276-019-0299-y

Lee, Y. L., Kung, F. C., Lin, C. H., & Huang, Y. S. (2020). CMTR1-Catalyzed 2'-O-Ribose Methylation Controls Neuronal Development by Regulating Camk2α Expression Independent of RIG-I Signaling. *Cell Reports*, *33*(3). https://doi.org/10.1016/j.celrep.2020.108269

Lei, X., Cao, Y., Ma, B., Zhang, Y., Ning, L., Qian, J., Zhang, L., Qu, Y., Zhang, T., Li, D., Chen, Q., Shi, J., Zhang, X., Ma, C., Zhang, Y., & Duan, E. (2020). *Development of mouse preimplantation embryos in space*. https://doi.org/10.1093/nsr/nwaa062

Lemaire, P. A., Anderson, E., Lary, J., & Cole, J. L. (2008). Mechanism of PKR Activation by dsRNA. *Journal of Molecular Biology*, *381*(2), 351–360. https://doi.org/10.1016/j.jmb.2008.05.056

Lesbirel, S., Viphakone, N., Parker, M., Parker, J., Heath, C., Sudbery, I., & Wilson, S. A. (2018). The m6A-methylase complex recruits TREX and regulates mRNA export. *Scientific Reports*, *8*(1). https://doi.org/10.1038/s41598-018-32310-8

Lewis, J. D., Izaurralde, E., Jarmolowski, A., Mcguigan, C., & Mattaj^, I. W. (1996). *A nuclear cap-binding complex facilitates association of Ul snRNP with the cap-proximal 5' splice site*.

Li, B., Clohisey, S. M., Chia, B. S., Wang, B., Cui, A., Eisenhaure, T., Schweitzer, L. D., Hoover, P., Parkinson, N. J., Nachshon, A., Smith, N., Regan, T., Farr, D., Gutmann, M. U., Bukhari, S. I., Law, A., Sangesland, M., Gat-Viks, I., Digard, P., … Hacohen, N. (2020). Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-019-13965-x

Li, J., & Sun, C. K. (2018). *SNHG5 is a potential prognostic biomarker in acute myeloid leukemia*. https://doi.org/10.26355/eurrev_201806_15154

Li, K., Foy, E., Ferreon, J. C., Nakamura, M., Ferreon, A. C. M., Ikeda, M., Ray, S. C., Gale, M. ‡, & Lemon, S. M. (2005). *Immune evasion by hepatitis C virus NS34A protease-mediated cleavage of the Toll-like receptor 3 adaptor protein TRIF*. www.pnas.orgcgidoi10.1073pnas.0408824102

Li, X., Xiong, X., Wang, K., Wang, L., Shu, X., Ma, S., & Yi, C. (2016). Transcriptome-wide mapping reveals reversible and dynamic N 1-methyladenosine methylome. *NATURE CHEMICAL BIOLOGY |*, *12*. https://doi.org/10.1038/NCHEMBIO.2040

Li, X., Zhu, P., Ma, S., Song, J., Bai, J., Sun, F., & Yi, C. (2015). *Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome*. https://doi.org/10.1038/nchembio.1836

Liang, S., Almohammed, R., & Cowling, V. H. (2023). The RNA cap methyltransferases RNMT and CMTR1 co-ordinate gene expression during neural differentiation. *Biochemical Society Transactions*. https://doi.org/10.1042/BST20221154

Liang, S., Silva, J. C., Suska, O., Lukoszek, R., Almohammed, R., & Cowling, V. H. (2022). CMTR1 is recruited to transcription start sites and promotes ribosomal protein and histone gene expression in embryonic stem cells. *Nucleic Acids Research*, *50*(5), 2905–2922. https://doi.org/10.1093/nar/gkac122

Liao, K. C., & Garcia-Blanco, M. A. (2021). Role of alternative splicing in regulating host response to viral infection. In *Cells* (Vol. 10, Issue 7). MDPI. https://doi.org/10.3390/cells10071720

Liddicoat, B. J., Piskol, R., Chalk, A. M., Ramaswami, G., Higuchi, M., Hartner, J. C., Li, J. B., Seeburg, P. H., & Walkley, C. R. (2015a). RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science*, *349*(6252), 1115–1120. https://doi.org/10.1126/science.aac7049

Liddicoat, B. J., Piskol, R., Chalk, A. M., Ramaswami, G., Higuchi, M., Hartner, J. C., Li, J. B., Seeburg, P. H., & Walkley, C. R. (2015b). RNA editing by ADAR1 prevents MDA5 sensing of endogenous dsRNA as nonself. *Science*, *349*(6252), 1115–1120. https://doi.org/10.1126/science.aac7049

Littlefield, J. W., & Dunn, D. B. (1958). The occurrence and distribution of thymine and three methylated-adenine bases in ribonucleic acids from several sources. *Biochemical Journal*, *70*(4), 642–651. https://doi.org/10.1042/bj0700642

Liu, J., Huang, T., Chen, W., Ding, C., Zhao, T., Zhao, X., Cai, B., Zhang, Y., Li, S., Zhang, L., Xue, M., He, X., Ge, W., Zhou, C., Xu, Y., & Zhang, R. (2022). Developmental mRNA m5C landscape and

regulatory innovations of massive m5C modification of maternal mRNAs in animals. *Nature Communications*, *13*(1), 2484. https://doi.org/10.1038/s41467-022-30210-0

Liu, L., Dong, H., Chen, H., Zhang, J., Ling, H., Li, Z., Shi, P.-Y., & Li, H. (2010). Flavivirus RNA cap methyltransferase: structure, function, and inhibition. *Frontiers in Biology*, *5*(4), 286–303. https://doi.org/10.1007/s11515-010-0660-y

Lu, L., Yi, C., Jian, X., Zheng, G., & He, C. (2010). Structure determination of DNA methylation lesions N 1 -meA and N 3 -meC in duplex DNA using a cross-linked protein–DNA system. *Nucleic Acids Research*, *38*(13), 4415–4425. https://doi.org/10.1093/nar/gkq129

Lu, W. T., Wilczynska, A., Smith, E., & Bushell, M. (2014). The diverse roles of the eIF4A family: You are the company you keep. *Biochemical Society Transactions*, *42*(1), 166–172. https://doi.org/10.1042/BST20130161

Mahmoud, M. E., Ui, F., Salman, D., Nishimura, M., & Nishikawa, Y. (2015). Mechanisms of interferon-beta-induced inhibition of Toxoplasma gondii growth in murine macrophages and embryonic fibroblasts: Role of immunity-related GTPase M1. *Cellular Microbiology*, *17*(7), 1069–1083. https://doi.org/10.1111/cmi.12423

Mannion, N., Greenwood, S. M., Young, R., Cox, S., Brindle, J., Read, D., Nellåker, C., Vesely, C., Ponting, C., McLaughlin, P., Jantsch, M., Dorin, J., Adams, I., Scadden, A. D. J., Öhman, M., Keegan, L., & O'Connell, M. (2014). The RNA-Editing Enzyme ADAR1 Controls Innate Immune Responses to RNA. *Cell Reports*, *9*(4), 1482–1494. https://doi.org/10.1016/j.celrep.2014.10.041

Mao, Y., Dong, L., Liu, X.-M., Guo, J., Ma, H., Shen, B., & Qian, S.-B. (2019). m6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nature Communications*, *10*(1), 5332. https://doi.org/10.1038/s41467-019-13317-9

Martin, B., Coutard, B., Guez, T., Paesen, G. C., Canard, B., Debart, F., Vasseur, J. J., Grimes, J. M., & Decroly, E. (2018). The methyltransferase domain of the Sudan ebolavirus L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure. *Nucleic Acids Research*, *46*(15), 7902–7912. https://doi.org/10.1093/nar/gky637

Martin, B., Coutard, B., Guez, T., Paesen, G. C., Canard, B., Oise Debart, F., Vasseur, J.-J., Grimes, J. M., & Decroly, E. (2018). The methyltransferase domain of the Sudan ebolavirus L protein specifically targets internal adenosines of RNA substrates, in addition to the cap structure. *Nucleic Acids Research*, *46*(15), 7902–7912. https://doi.org/10.1093/nar/gky637

Martin, H., Rupkey, J., Asthana, S., Yoon, J., Patel, S., Mott, J., Pei, Z., & Mao, Y. (2022). Diverse Roles of the Exon Junction Complex Factors in the Cell Cycle, Cancer, and Neurodevelopmental Disorders-Potential for Therapeutic Targeting. In *International Journal of Molecular Sciences* (Vol. 23, Issue 18). MDPI. https://doi.org/10.3390/ijms231810375

Martinez, N. M., Su, A., Burns, M. C., Nussbacher, J. K., Schaening, C., Sathe, S., Yeo, G. W., & Gilbert, W. V. (2022). Pseudouridine synthases modify human pre-mRNA co-transcriptionally and affect pre-mRNA processing. *Molecular Cell*, *82*(3), 645-659.e9. https://doi.org/10.1016/j.molcel.2021.12.023

Martínez-Espinoza, I., & Guerrero-Plata, A. (2022). The Relevance of TLR8 in Viral Infections. In *Pathogens* (Vol. 11, Issue 2). MDPI. https://doi.org/10.3390/pathogens11020134

Matera, A. G., & Wang, Z. (2014). A day in the life of the spliceosome. In *Nature Reviews Molecular Cell Biology* (Vol. 15, Issue 2, pp. 108–121). https://doi.org/10.1038/nrm3742

McCann, K. L., Kavari, S. L., Burkholder, A. B., Phillips, B. T., & Tanaka Hall, T. M. (2020). H/ACA snoRNA levels are regulated during stem cell differentiation. *Nucleic Acids Research*, *48*(15), 8686–8703. https://doi.org/10.1093/nar/gkaa612

McKendrick, L., Thompson, E., Ferreira, J., Morley, S. J., & Lewis, J. D. (2001). Interaction of Eukaryotic Translation Initiation Factor 4G with the Nuclear Cap-Binding Complex Provides a Link between Nuclear and Cytoplasmic Functions of the m 7 Guanosine Cap . *Molecular and Cellular Biology*, *21*(11), 3632–3641. https://doi.org/10.1128/mcb.21.11.3632-3641.2001

Melcher, T., Maas, S., Herb, A., Sprengel, R., Seeburg, P. H., & Higuchi, M. (1996). A mammalian RNA editing enzyme. *Nature*, *379*(6564), 460–464. https://doi.org/10.1038/379460a0

Mendel, M., Chen, K. M., Homolka, D., Gos, P., Pandey, R. R., McCarthy, A. A., & Pillai, R. S. (2018). Methylation of Structured RNA by the m6A Writer METTL16 Is Essential for Mouse Embryonic Development. *Molecular Cell*, *71*(6), 986-1000.e11. https://doi.org/10.1016/j.molcel.2018.08.004

Mendel, M., Delaney, K., Pandey, R. R., Chen, K. M., Wenda, J. M., Vågbø, C. B., Steiner, F. A., Homolka, D., & Pillai, R. S. (2021). Splice site m6A methylation prevents binding of U2AF35 to inhibit RNA splicing. *Cell*, *184*(12), 3125-3142.e25. https://doi.org/10.1016/j.cell.2021.03.062

Meng, Z. (2006). Mass spectrometry of RNA: linking the genome to the proteome. *Briefings in Functional Genomics and Proteomics*, *5*(1), 87–95. https://doi.org/10.1093/bfgp/ell012

Merkin, J., Russell, C., Chen, P., & Burge, C. B. (2012). Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science*, *338*(6114), 1593–1599. https://doi.org/10.1126/science.1228186

Meyer, K. D., Patil, D. P., Zhou, J., Zinoviev, A., Skabkin, M. A., Elemento, O., Pestova, T. V., Qian, S. B., & Jaffrey, S. R. (2015). 5′ UTR m6A Promotes Cap-Independent Translation. *Cell*, *163*(4), 999–1010. https://doi.org/10.1016/j.cell.2015.10.012

Meyer, K. D., Saletore, Y., Zumbo, P., Elemento, O., Mason, C. E., & Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell*, *149*(7), 1635–1646. https://doi.org/10.1016/j.cell.2012.05.003

Miedziak, B., Dobieżyńska, A., Darżynkiewicz, Z. M., Bartkowska, J., Miszkiewicz, J., Kowalska, J., Warminski, M., Tyras, M., Trylska, J., Jemielity, J., Darzynkiewicz, E., & Grzela, R. (2020). Kinetic analysis of IFIT1 and IFIT5 interactions with different native and engineered RNAs and its consequences for designing mRNA-based therapeutics. *RNA*, *26*(1), 58–68. https://doi.org/10.1261/rna.073304.119

Miillner, E. W., & Kiihn, L. C. (1988). A Stem-Loop in the 3′ Untranslated Region Mediates Iron-Dependent Regulation of Transferrin Receptor mRNA Stability in the Cytoplasm. In *Cell* (Vol. 53).

Miles, M. F., Kõks, S., Merienne, K., Mattick jmattick, J. S., Barry, G., Bitar, M., Guennewig, B., Mladenova, D., Konen, L. M., Pineda, S. S., Avesson, L., Zinn, R., Schonrock, N., Jonkhout, N., Crumlish, L., Kaczorowski, D. C., Gong, A., Pinese, M., Franco, G. R., … Mattick, J. S. (2018).

Adar3 Is Involved in Learning and Memory in Mice. *Frontiers in Neuroscience | Www.Frontiersin.Org*, *1*, 243. https://doi.org/10.3389/fnins.2018.00243

Moore, P. B., & Steitz, T. A. (2011). The roles of RNA in the synthesis of protein. In *Cold Spring Harbor Perspectives in Biology* (Vol. 3, Issue 11). https://doi.org/10.1101/cshperspect.a003780

Moreno-Contreras, J., Sánchez-Tacuba, L., Arias, C. F., & López, S. (2022). Mature Rotavirus Particles Contain Equivalent Amounts of 7me GpppG-Capped and Noncapped Viral Positive-Sense RNAs . *Journal of Virology*, *96*(17). https://doi.org/10.1128/jvi.01151-22

Morrison, J., & García-Sastre, A. (2014). STAT2 signaling and dengue virus infection. *JAK-STAT*, *3*(1), e27715. https://doi.org/10.4161/jkst.27715

Motorin, Y., & Marchand, V. (2021). *Analysis of RNA Modifications by Second-and Third-Generation Deep Sequencing: 2020 Update*. https://doi.org/10.3390/genes12020278

Motorin, Y., Quinternet, M., Rhalloussi, W., & Marchand, V. (2021). *Constitutive and variable 2'-O-methylation (Nm) in human ribosomal RNA*. https://doi.org/10.1080/15476286.2021.1974750

Muthukrishnan, S., Filipowicz, W., Sierra, J. M., Both, G. W., Shatkin, A. J., & Ochoa, S. (1975). mRNA methylation and protein synthesis in extracts from embryos of brine shrimp, Artemia salina. *Journal of Biological Chemistry*, *250*(24), 9336–9341. https://doi.org/10.1016/s0021-9258(19)40648-0

Muthukrishnan, S., Moss, B., Cooper, J. A., & Maxwell, E. S. (1978). Influence of 5'-terminal cap structure on the initiation of translation of vaccinia virus mRNA. *Journal of Biological Chemistry*, *253*(5), 1710–1715. http://www.jbc.org/content/253/5/1710.abstract

Nakatsuji, N. (1992). Development Development of Postimplantation Mouse Embryos: Unexplored Field Rich in Unanswered Questions. In *Develop. Growth & Differ* (Vol. 34, Issue 5).

Nakayama, E., Kawai, Y., Taniguchi, S., Hazlewood, J. E., Shibasaki, K. I., Takahashi, K., Sato, Y., Tang, B., Yan, K., Katsuta, N., Tajima, S., Lim, C. K., Suzuki, T., Suhrbier, A., & Saijo, M. (2021). Embryonic stage of congenital zika virus infection determines fetal and postnatal outcomes in mice. *Viruses*, *13*(9). https://doi.org/10.3390/v13091807

Nallagatla, S. R., & Bevilacqua, P. C. (2008). Nucleoside modifications modulate activation of the protein kinase PKR in an RNA structure-specific manner. *RNA*, *14*(6), 1201–1213. https://doi.org/10.1261/rna.1007408

Nallagatla, S. R., Hwang, J., Toroney, R., Zheng, X., Cameron, C. E., & Bevilacqua, P. C. (2007). 5'-triphosphate-dependent activation of PKR by RNAs with short stem-loops. *Science*, *318*(5855), 1455–1458. https://doi.org/10.1126/science.1147347

Nees, G., Kaufmann, A., & Bauer, S. (2014). Detection of rna modifi cations by hplc analysis and competitive elisa. *Methods in Molecular Biology*, *1169*, 3–14. https://doi.org/10.1007/978-1-4939-0882-0_1

Newby, M. I., & Greenbaum, N. L. (2002). *Investigation of Overhauser effects between pseudouridine and water protons in RNA helices* (Vol. 99). www.rcsb.org

Nguyen, H., Kerimoglu, C., Pirouz, M., Pham, L., Kiszka, K. A., Sokpor, G., Sakib, M. S., Rosenbusch, J., Teichmann, U., Seong, R. H., Stoykova, A., Fischer, A., Staiger, J. F., & Tuoc, T. (2018). Epigenetic Regulation by BAF Complexes Limits Neural Stem Cell Proliferation by Suppressing Wnt

Signaling in Late Embryonic Development. *Stem Cell Reports*, *10*(6), 1734–1750. https://doi.org/10.1016/j.stemcr.2018.04.014

Niescierowicz, K., Pryszcz, L., Navarrete, C., Tralle, E., Sulej, A., Abu Nahia, K., Kasprzyk, M. E., Misztal, K., Pateria, A., Pakuła, A., Bochtler, M., & Winata, C. (2022). Adar-mediated A-to-I editing is required for embryonic patterning and innate immune response regulation in zebrafish. *Nature Communications*, *13*(1), 5520. https://doi.org/10.1038/s41467-022-33260-6

Nobuaki Kono, & Kazuharu Arakawa. (2019). *Nanopore sequencing: Review of potential applications in functional genomics*. https://doi.org/10.1111/dgd.12608

Nomoto, Aki., Fon Lee, Y., & Wimmer, E. (1976). The 5' end of poliovirus mRNA is not capped with m7G(5')ppp(5')Np (polyribosomes/fingerprints). In *Biochemistry* (Vol. 73, Issue 2). https://www.pnas.org

Obata, Y., Ono, Y., Akuzawa, H., Kwon, O. Y., Yoshizawa, M., & Kono, T. (2000). Post-implantation development of mouse androgenetic embryos produced by in-vitro fertilization of enucleated oocytes. In *Human Reproduction* (Vol. 15, Issue 4).

Ogden, K. M., Snyder, M. J., Dennis, A. F., & Patton, J. T. (2014). Predicted Structure and Domain Organization of Rotavirus Capping Enzyme and Innate Immune Antagonist VP3. *Journal of Virology*, *88*(16), 9072–9085. https://doi.org/10.1128/jvi.00923-14

Ogino, T., & Green, T. J. (2019). Transcriptional Control and mRNA Capping by the GDP Polyribonucleotidyltransferase Domain of the Rabies Virus Large Protein. *Viruses*, *11*(6), 504. https://doi.org/10.3390/v11060504

Olschewski, S., Cusack, S., & Rosenthal, M. (2020). The Cap-Snatching Mechanism of Bunyaviruses. In *Trends in Microbiology* (Vol. 28, Issue 4, pp. 293–303). Elsevier Ltd. https://doi.org/10.1016/j.tim.2019.12.006

Orellana, E. A., Siegal, E., & Gregory, R. I. (2022). tRNA dysregulation and disease. In *Nature Reviews Genetics* (Vol. 23, Issue 11, pp. 651–664). Nature Research. https://doi.org/10.1038/s41576-022-00501-9

Palladino, M. J., Keegan, L. P., O'connell, M. A., & Reenan, R. A. (2000). A-to-I Pre-mRNA Editing in Drosophila Is Primarily Involved in Adult Nervous System Function and Integrity and RED2. They contain two or three dsRNA binding domains and a catalytic deaminase domain. The deaminase domain is distantly related to cytidine deaminases, and in particular to another editing enzyme, APOBEC-1, that catalyzes the conversion of C to U in transcripts of. In *Cell* (Vol. 102).

Pan, R., Kindler, E., Cao, L., Zhou, Y., Zhang, Z., Liu, Q., Ebert, N., Züst, R., Sun, Y., Gorbalenya, A. E., Perlman, S., Thiel, V., Chen, Y., & Guo, D. (2022). N7-Methylation of the Coronavirus RNA Cap Is Required for Maximal Virulence by Preventing Innate Immune Recognition. *MBio*, *13*(1). https://doi.org/10.1128/mbio.03662-21

Pandey, R. R., Delfino, E., Homolka, D., Roithova, A., Chen, K. M., Li, L., Franco, G., Vågbø, C. B., Taillebourg, E., Fauvarque, M. O., & Pillai, R. S. (2020). The Mammalian Cap-Specific m6Am RNA Methyltransferase PCIF1 Regulates Transcript Levels in Mouse Tissues. *Cell Reports*, *32*(7). https://doi.org/10.1016/j.celrep.2020.108038

Parker, M. T., Knop, K., Sherwood, A. V, Schurch, N. J., Mackinnon, K., Gould, P. D., Hall, A. J., Barton, G. J., & Simpson, G. G. (2020). Nanopore direct RNA sequencing maps the complexity of

Arabidopsis mRNA processing and m6A modification. *ELife*, *9*. https://doi.org/10.7554/eLife.49658

Patil, D. P., Chen, C. K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., & Jaffrey, S. R. (2016). M6 A RNA methylation promotes XIST-mediated transcriptional repression. *Nature*, *537*(7620), 369–373. https://doi.org/10.1038/nature19342

Patzelt, E., Thalmann, E., Hartmuth, K., Blass, D., & Kuechler, E. (1987). Assembly of pre-mRNA splicing complex is cap dependent. *Nucleic Acids Research*, *15*(4), 1387–1399. https://doi.org/10.1093/nar/15.4.1387

Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., & Conrad, N. K. (2017). The U6 snRNA m6A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. *Cell*, *169*(5), 824-835.e14. https://doi.org/10.1016/j.cell.2017.05.003

Pestal, K., Funk, C. C., Snyder, J. M., Price, N. D., Treuting, P. M., & Stetson, D. B. (2015). Isoforms of RNA-Editing Enzyme ADAR1 Independently Control Nucleic Acid Sensor MDA5-Driven Autoimmunity and Multi-organ Development. *Immunity*, *43*(5), 933–944. https://doi.org/10.1016/j.immuni.2015.11.001

Petes, C., Odoardi, N., & Gee, K. (2017). The Toll for trafficking: Toll-like receptor 7 delivery to the endosome. In *Frontiers in Immunology* (Vol. 8, Issue SEP). Frontiers Media S.A. https://doi.org/10.3389/fimmu.2017.01075

Philippe, L., Van Den Elzen, A. M. G., Watson, M. J., & Thoreen, C. C. (2020). Global analysis of LARP1 translation targets reveals tunable and dynamic features of 5′ TOP motifs. *PNAS*, *117*(10). https://doi.org/10.1073/pnas.1912864117/-/DCSupplemental

Phizicky, E. M., & Hopper, A. K. (2010). tRNA biology charges to the front. In *Genes and Development* (Vol. 24, Issue 17, pp. 1832–1860). https://doi.org/10.1101/gad.1956510

Picard-Jean, F., Brand, C., Tremblay-Létourneau, M., Allaire, A., Beaudoin, M. C., Boudreault, S., Duval, C., Rainville-Sirois, J., Robert, F., Pelletier, J., Geiss, B. J., & Bisaillon, M. (2018). 2-O-methylation of the mRNA cap protects RNAs from decapping and degradation by DXO. *PLOS ONE*, *13*(3), e0193804. https://doi.org/10.1371/journal.pone.0193804

Pichlmair, A., Lassnig, C., Eberle, C. A., Górna, M. W., Baumann, C. L., Burkard, T. R., Búrckstúmmer, T., Stefanovic, A., Krieger, S., Bennett, K. L., Rúlicke, T., Weber, F., Colinge, J., Múller, M., & Superti-Furga, G. (2011). IFIT1 is an antiviral protein that recognizes 5′-triphosphate RNA. *Nature Immunology*, *12*(7), 624–630. https://doi.org/10.1038/ni.2048

Pichlmair, A., Schulz, O., Tan, C. P., Näslund, T. I., Liljeström, P., Weber, F., & Reis e Sousa, C. (2006). RIG-I-Mediated Antiviral Responses to Single-Stranded RNA Bearing 5′-Phosphates. *Science*, *314*(5801), 997–1001. https://doi.org/10.1126/science.1132998

Pichot, F., Marchand, V., Ayadi, L., Bourguignon-Igel, V., Helm, M., & Motorin, Y. (2020). Holistic Optimization of Bioinformatic Analysis Pipeline for Detection and Quantification of 2′-O-Methylations in RNA by RiboMethSeq. *Frontiers in Genetics*, *11*. https://doi.org/10.3389/fgene.2020.00038

Piechotta, M., Naarmann-de Vries, I. S., Wang, Q., Altmüller, J., & Dieterich, C. (2022). RNA modification mapping with JACUSA2. *Genome Biology*, *23*(1), 115. https://doi.org/10.1186/s13059-022-02676-0

Price, A. M., Hayer, K. E., McIntyre, A. B. R., Gokhale, N. S., Abebe, J. S., Della Fera, A. N., Mason, C. E., Horner, S. M., Wilson, A. C., Depledge, D. P., & Weitzman, M. D. (2020). Direct RNA sequencing reveals m6A modifications on adenovirus RNA are necessary for efficient splicing. *Nature Communications*, *11*(1), 6016. https://doi.org/10.1038/s41467-020-19787-6

Qiu, J., Wang, P., Chen, Z., Zhou, Y., Zhang, G., Wang, Z., Wu, J., Zhu, Q., & Jiang, C. (2023). Long noncoding RNA SNHG4 promotes the malignant progression of hepatocellular carcinoma through the miR -211-5p/CREB5 axis. *Cancer Medicine*, *12*(7), 8388–8402. https://doi.org/10.1002/cam4.5559

Rand, A. C., Jain, M., Eizenga, J. M., Musselman-Brown, A., Olsen, H. E., Akeson, M., & Paten, B. (2017). *mapping dna methylation with high-throughput nanopore sequencing*. *14*(4). https://doi.org/10.1038/nmeth.4189

Rehwinkel, J., & Gack, M. U. (n.d.). *RIG-I-like receptors: their regulation and roles in RNA sensing*. https://doi.org/10.1038/s41577-020-0288-3

Ren, X., Linehan, M. M., Iwasaki, A., & Pyle, A. M. (2019). RIG-I Recognition of RNA Targets: The Influence of Terminal Base Pair Sequence and Overhangs on Affinity and Signaling. *Cell Reports*, *29*(12), 3807-3815.e3. https://doi.org/10.1016/j.celrep.2019.11.052

Rice, G. I., Kasher, P. R., Forte, G. M. A., Mannion, N. M., Greenwood, S. M., Szynkiewicz, M., Dickerson, J. E., Bhaskar, S. S., Zampini, M., Briggs, T. A., Jenkinson, E. M., Bacino, C. A., Battini, R., Bertini, E., Brogan, P. A., Brueton, L. A., Carpanelli, M., De Laet, C., De Lonlay, P., … Crow, J. (2012). Mutations in ADAR1 cause Aicardi-Goutières syndrome associated with a type I interferon signature. *Nature Genetics*, *44*(11), 22–24. https://doi.org/10.1038/ng.2414

Ringeard, M., Marchand, V., Decroly, E., Motorin, Y., & Bennasser, Y. (2019). FTSJ3 is an RNA 2′-O-methyltransferase recruited by HIV to avoid innate immune sensing. *Nature*, *565*(7740), 500–504. https://doi.org/10.1038/s41586-018-0841-4

Robbins, M., Judge, adam, Liang, L., McClintock, K., Yaworski, E., & MacLachlan, I. (2001). 2′-O-methyl-modified RNAs Act as TLR7 Antagonists Enhanced Reader. *The American Society of Gene Therapy*.

Robinson, E. K., Jagannatha, P., Covarrubias, S., Cattle, M., Smaliy, V., Safavi, R., Shapleigh, B., Abu-Shumays, R., Jain, M., Cloonan, S. M., Akeson, M., Brooks, A. N., & Carpenter, S. (2021). Inflammation drives alternative first exon usage to regulate immune genes including a novel iron regulated isoform of aim2. *ELife*, *10*. https://doi.org/10.7554/eLife.69431

Robinson, J. E., Paluch, J., Dickman, D. K., & Joiner, W. J. (2015). ARTICLE ADAR-mediated RNA editing suppresses sleep by acting as a brake on glutamatergic synaptic plasticity. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms10512

S Brenner, B. D., & Meselson, D. M. (1957). AN UNSTABLE INTERMEDIATE CARRYING INFORMATION FROM GENES TO RIBOSOMES FOR PROTEIN SYNTHESIS. In *Defence Atomic Support Agency, Rep. DASA-532 B* (Vol. 186). Harwell.

Safra, modi, Sas-Chen, A., nir, ronit, Winkler, roni, nachshon, A., bar-Yaacov, D., erlacher, matthias, rossmanith, W., Stern-Ginossar, noam, & Schwartz, S. (2017). *The m 1 A landscape on cytosolic and mitochondrial mRNA at single-base resolution*. https://doi.org/10.1038/nature24456

Saitou, M., Barton, S. C., & Surani, M. A. (2002). *A molecular programme for the specification of germ cell fate in mice*. www.nature.com/nature

Saitou, M., & Yamaji, M. (2012). Primordial germ cells in mice. *Cold Spring Harbor Perspectives in Biology*, *4*(11). https://doi.org/10.1101/cshperspect.a008375

Sato, H., & Maquat, L. E. (2009). Remodeling of the pioneer translation initiation complex involves translation and the karyopherin importin β. *Genes & Development*, *23*(21), 2537–2550. https://doi.org/10.1101/gad.1817109

Schibler, U., Kelley, D. E., & Perry, R. P. (1977). Comparison of methylated sequences in messenger RNA and heterogeneous nuclear RNA from mouse L cells. *Journal of Molecular Biology*, *115*(4), 695–714. https://doi.org/10.1016/0022-2836(77)90110-3

Schlautmann, L. P., & Gehring, N. H. (2020). A Day in the Life of the Exon Junction Complex. *Biomolecules*, *10*(6), 866. https://doi.org/10.3390/biom10060866

Schoggins, J. W., & Rice, C. M. (2011). Interferon-stimulated genes and their antiviral effector functions. In *Current Opinion in Virology* (Vol. 1, Issue 6, pp. 519–525). Elsevier B.V. https://doi.org/10.1016/j.coviro.2011.10.008

Schoggins, J. W., Wilson, S. J., Panis, M., Murphy, M. Y., Jones, C. T., Bieniasz, P., & Rice, C. M. (2011). A diverse range of gene products are effectors of the type i interferon antiviral response. *Nature*, *472*(7344), 481–485. https://doi.org/10.1038/nature09907

Schuberth-Wagner, C., Ludwig, J., Bruder, A., Herzner, A.-M., Zillinger, T., Goldeck, M., Schmidt, T., Schmid-Burgk, J., Kerber, R., Wolter, S., Stümpel, J.-P., Roth, A., Bartok, E., Drosten, C., Coch, C., Hornung, V., Barchet, W., Kümmerer, B., Hartmann, G., & Schlee, M. (2015). A Conserved Histidine in the RNA Sensor RIG-I Controls Immune Tolerance to N1-2^O-Methylated Self RNA. *Immunity*, *43*(1), 41–51. https://doi.org/10.1016/j.immuni.2015.06.015

Schumann, U., Zhang, H.-N., Sibbritt, T., Pan, A., Horvath, A., Gross, S., Clark, S. J., Yang, L., & Preiss, T. (2020). Multiple links between 5-methylcytosine content of mRNA and translation. *BMC Biology*, *18*(1), 40. https://doi.org/10.1186/s12915-020-00769-5

Schwartz, S. (2018). *m 1 A within cytoplasmic mRNAs at single nucleotide resolution: a reconciled transcriptome-wide map*. https://doi.org/10.1261/rna

Schwartz, S., Agarwala, S. D., Mumbach, M. R., Jovanovic, M., Mertins, P., Shishkin, A., Tabach, Y., Mikkelsen, T. S., Satija, R., Ruvkun, G., Carr, S. A., Lander, E. S., Fink, G. R., & Regev, A. (2013). High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. *Cell*, *155*(6). https://doi.org/10.1016/j.cell.2013.10.047

Seeburg, P. H., Higuchi, M., & Sprengel, R. (1998). RNA editing of brain glutamate receptor channels: mechanism and physiology. *Brain Research Reviews*, *26*(2–3), 217–229. https://doi.org/10.1016/S0165-0173(97)00062-3

Selmi, T., Hussain, S., Dietmann, S., Heiß, M., Borland, K., Flad, S., Carter, J.-M., Dennison, R., Huang, Y.-L., Kellner, S., Borneï Ov, S., & Frye, M. (2021). Sequence-and structure-specific cytosine-5 mRNA methylation by NSUN6. *Nucleic Acids Research*, *49*(2), 1006–1022. https://doi.org/10.1093/nar/gkaa1193

Sendinc, E., Valle-Garcia, D., Dhall, A., Chen, H., Sheng, W., Adelman, K., & Shi, Y. (2019). PCIF1 catalyzes m6Am mRNA methylation to regulate gene expression 1 2. *Molecular Cell*, *75*(3), 620-630.e9. https://doi.org/10.1101/484931

Sertznig, H., Roesmann, F., Wilhelm, A., Heininger, D., Bleekmann, B., Elsner, C., Santiago, M., Schuhenn, J., Karakoese, Z., Benatzy, Y., Snodgrass, R., Esser, S., Sutter, K., Dittmer, U., & Widera, M. (2022). SRSF1 acts as an IFN-I-regulated cellular dependency factor decisively affecting HIV-1 post-integration steps. *Frontiers in Immunology*, *13*. https://doi.org/10.3389/fimmu.2022.935800

Shatkin, A. J., & Manley, J. L. (2000). The ends of the affair: Capping and polyadenylation. *Nat Struct Biol* . https://doi.org/10.1038/79583

Shaw, A. E., Hughes, J., Gu, Q., Behdenna, A., Singer, J. B., Dennis, T., Orton, R. J., Varela, M., Gifford, R. J., Wilson, S. J., & Palmarini, M. (2017). Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. *PLoS Biology*, *15*(12). https://doi.org/10.1371/journal.pbio.2004086

Shih, D. S., Dasgupta, R., & Kaesberg, P. (1976). 7-Methyl-Guanosine and Efficiency of RNA Translation. In *JOURNAL OF VIROLOGY* (Vol. 19, Issue 2). https://journals.asm.org/journal/jvi

Shiono, Y., Mun, H. S., He, N., Nakazaki, Y., Fang, H., Furuya, M., Aosai, F., & Yano, A. (2007). Maternal-fetal transmission of Toxoplasma gondii in interferon-γ deficient pregnant mice. *Parasitology International*, *56*(2), 141–148. https://doi.org/10.1016/j.parint.2007.01.008

Sidrat, T., Rehman, Z. U., Joo, M. D., Lee, K. L., & Kong, I. K. (2021a). Wnt/β-catenin pathway-mediated pparδ expression during embryonic development differentiation and disease. In *International Journal of Molecular Sciences* (Vol. 22, Issue 4, pp. 1–10). MDPI AG. https://doi.org/10.3390/ijms22041854

Sidrat, T., Rehman, Z.-U., Joo, M.-D., Lee, K.-L., & Kong, I.-K. (2021b). Wnt/β-catenin Pathway-Mediated PPARδ Expression during Embryonic Development Differentiation and Disease. *International Journal of Molecular Sciences*, *22*(4), 1854. https://doi.org/10.3390/ijms22041854

Simmons, R. A., Willberg, C. B., & Paul, K. (2013). Immune Evasion by Viruses. In *eLS*. Wiley. https://doi.org/10.1002/9780470015902.a0024790

Simonović, M., & Steitz, T. A. (2009). A structural view on the mechanism of the ribosome-catalyzed peptide bond formation. In *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms* (Vol. 1789, Issues 9–10, pp. 612–623). https://doi.org/10.1016/j.bbagrm.2009.06.006

Sloan, K. E., Bohnsack, M. T., & Watkins, N. J. (2013). The 5S RNP Couples p53 Homeostasis to Ribosome Biogenesis and Nucleolar Stress. *Cell Reports*, *5*(1), 237–247. https://doi.org/10.1016/j.celrep.2013.08.049

Smietanski, M., Werner, M., Purta, E., Kaminska, K. H., Stepinski, J., Darzynkiewicz, E., Nowotny, M., & Bujnicki, J. M. (2014). Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms4004

Sokol, S. Y. (2015). Spatial and temporal aspects of Wnt signaling and planar cell polarity during vertebrate embryonic development. In *Seminars in Cell and Developmental Biology* (Vol. 42, pp. 78–85). Academic Press. https://doi.org/10.1016/j.semcdb.2015.05.002

Song, Y., Feng, N., Sanchez-Tacuba, L., Yasukawa, L. L., Ren, L., Silverman, R. H., Ding, S., & Greenberg, H. B. (2020). *Reverse Genetics Reveals a Role of Rotavirus VP3 Phosphodiesterase Activity in Inhibiting RNase L Signaling and Contributing to Intestinal Viral Replication In Vivo*. https://doi.org/10

Steitz, T. A. (2008). A structural understanding of the dynamic ribosome machine. In *Nature Reviews Molecular Cell Biology* (Vol. 9, Issue 3, pp. 242–253). https://doi.org/10.1038/nrm2352

Stephenson, W., Razaghi, R., Busan, S., Weeks, K. M., Timp, W., & Smibert, P. (2022). Direct detection of RNA modifications and structure using single-molecule nanopore sequencing. *Cell Genomics*, *2*(2), 100097. https://doi.org/10.1016/j.xgen.2022.100097

Stoiber1, M., Quick2, J., Egan3, R., Lee3, J. E., & Celniker1, S. (2017). De novo Identification of DNA Modifications Enabled by Genome-Guided Nanopore Signal Processing. *BioRxiv*. https://doi.org/10.1101/094672

Stok, J. E., Oosenbrug, T., Haar, L. R., Gravekamp, D., Bromley, C. P., Zelenay, S., Reis e Sousa, C., & Veen, A. G. (2022). RNA sensing via the RIG-I-like receptor LGP2 is essential for the induction of a type I IFN response in ADAR1 deficiency. *The EMBO Journal*, *41*(6). https://doi.org/10.15252/embj.2021109760

Sun, H., Li, K., Liu, C., & Yi, C. (2023). Regulation and functions of non-m6A mRNA modifications. *Nature Reviews Molecular Cell Biology*. https://doi.org/10.1038/s41580-023-00622-x

Sun, H., Zhang, M., Li, K., Bai, D., & Yi, C. (2019). Cap-specific, terminal N 6-methylation by a mammalian m 6 Am methyltransferase. *Cell Research*, *29*(1), 80–82. https://doi.org/10.1038/s41422-018

Tam, P. P. L., & Loebel, D. A. F. (2007). Gene function in mouse embryogenesis: Get set for gastrulation. In *Nature Reviews Genetics* (Vol. 8, Issue 5, pp. 368–381). https://doi.org/10.1038/nrg2084

Tamarkin-Ben-Harush, A., Vasseur, J.-J., oise Debart, F., Ulitsky, I., & Dikstein, R. (2017). Cap-proximal nucleotides via differential eIF4E binding and alternative promoter usage mediate translational response to energy stress. *ELife*. https://doi.org/10.7554/eLife.21907.001

Tan, X., Zheng, C., Zhuang, Y., Jin, P., & Wang, F. (2023). The m6A reader PRRC2A is essential for meiosis I completion during spermatogenesis. *Nature Communications*, *14*(1). https://doi.org/10.1038/s41467-023-37252-y

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., & 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 Research*, *27*(10), 1759–1768. https://doi.org/10.1101/gr.220962.117

Tarn, W. Y., & Steitz, J. A. (1996). A Novel Spliceosome Containing U11, U12, and U5 snRNPs Excises a Minor Class (AT–AC) Intron In Vitro. *Cell*, *84*(5), 801–811. https://doi.org/10.1016/S0092-8674(00)81057-0

Tartell, M. A., Boulias, K., Brunsting Hoffmann, G., Bloyet, L.-M., Lieberman Greer, E., & J Whelan, S. P. (2021). Methylation of viral mRNA cap structures by PCIF1 attenuates the antiviral activity of interferon-β. *PNAS*, *118*, 2021. https://doi.org/10.1073/pnas.2025769118/-/DCSupplemental

Tesson, F., Hervé, A., Mordret, E., Touchon, M., d'Humières, C., Cury, J., & Bernheim, A. (2022). Systematic and quantitative view of the antiviral arsenal of prokaryotes. *Nature Communications*, *13*(1). https://doi.org/10.1038/s41467-022-30269-9

Thalalla Gamage, S., Sas-Chen, A., Schwartz, S., & Meier, J. L. (2021). Quantitative nucleotide resolution profiling of RNA cytidine acetylation by ac4C-seq. *Nature Protocols*, *16*(4), 2286–2307. https://doi.org/10.1038/s41596-021-00501-9

Theil, E. C. (1994). Iron regulatory elements (IREs): a family of mRNA non-coding sequences INTRODUCTION: THE IRON REGULATORY ELEMENTS FAMILY. In *Biochem. J* (Vol. 304).

Toczydlowska-Socha, D., Zielinska, M. M., Kurkowska, M., Astha, Almeida, C. F., Stefaniak, F., Purta, E., & Bujnicki, J. M. (2018). Human RNA cap1 methyltransferase CMTr1 cooperates with RNA helicase DHX15 to modify RNAs with highly structured 50 termini. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1762). https://doi.org/10.1098/rstb.2018.0161

Topisirovic, I., Svitkin, Y. V, Sonenberg, N., & Shatkin, A. J. (2011). Cap and cap-binding proteins in the control of gene expression. *WIREs RNA*, *2*(2), 277–298. https://doi.org/https://doi.org/10.1002/wrna.52

Tress, M. L., Abascal, F., & Valencia, A. (2017). Alternative Splicing May Not Be the Key to Proteome Complexity. In *Trends in Biochemical Sciences* (Vol. 42, Issue 2, pp. 98–110). Elsevier Ltd. https://doi.org/10.1016/j.tibs.2016.08.008

Trixl, L., & Lusser, A. (2019). Getting a hold on cytosine methylation in mRNA. *NATurE STrucTurAl & MolEculAr Biology |*, *26*. https://doi.org/10.1038/s41594-019-0217-y

Tsukamoto, Y., Hiono, T., Yamada, S., Matsuno, K., Faist, A., Claff, T., Hou, J., Namasivayam, V., vom Hemdt, A., Sugimoto, S., Ng, J. Y., Christensen, M. H., Tesfamariam, Y. M., Wolter, S., Juranek, S., Zillinger, T., Bauer, S., Hirokawa, T., Schmidt, F. I., … Kato, H. (2023). Inhibition of cellular RNA methyltransferase abrogates influenza virus capping and replication. *Science*, *379*(6632), 586–591. https://doi.org/10.1126/science.add0875

Tudek, A., Lloret-Llinares, M., & Jensen, T. H. (2018). The multitasking polyA tail: nuclear RNA maturation, degradation and export. *Phil. Trans. R. Soc.* https://doi.org/10.1098/rstb.2018.0169

Uzonyi, A., Dierks, D., Nir, R., Kwon, O. S., Toth, U., Barbosa, I., Burel, C., Brandis, A., Rossmanith, W., Le Hir, H., Slobodin, B., & Schwartz, S. (2023). Exclusion of m6A from splice-site proximal regions by the exon junction complex dictates m6A topologies and mRNA stability. *Molecular Cell*, *83*(2), 237-251.e7. https://doi.org/10.1016/j.molcel.2022.12.026

Uzri, D., & Greenberg, H. B. (2013). Characterization of Rotavirus RNAs That Activate Innate Immune Signaling through the RIG-I-Like Receptors. *PLoS ONE*, *8*(7). https://doi.org/10.1371/journal.pone.0069825

van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., & Thermes, C. (2018). The Third Revolution in Sequencing Technology. *Trends in Genetics : TIG*, *34*(9), 666–681. https://doi.org/10.1016/j.tig.2018.05.008

Varshney, D., Petit, A.-P., Bueren-Calabuig, J. A., Jansen, C., Fletcher, D. A., Peggie, M., Weidlich, S., Scullion, P., Pisliakov, A. V, & Cowling, V. H. (2016). Molecular basis of RNA guanine-7 methyltransferase (RNMT) activation by RAM. *Nucleic Acids Research*, *44*(21), 10423–10436. https://doi.org/10.1093/nar/gkw637

Verma, B., Akinyi, M. V., Norppa, A. J., & Frilander, M. J. (2018). Minor spliceosome and disease. *Seminars in Cell & Developmental Biology*, *79*, 103–112. https://doi.org/10.1016/J.SEMCDB.2017.09.036

Visa, N., Izaurralde, E., Ferreira, J., Daneholt, B., & Mattaj, I. W. (1977). A Nuclear Cap-binding Complex Binds Balbiani Ring Pre-mRNA Cotranscriptionally and Accompanies the Ribonucleoprotein Particle during Nuclear Export. In *Shatkin*. Hamm and Mattaj.

Wang, G., Wang, H., Singh, S., Zhou, P., Yang, S., Wang, Y., Zhu, Z., Zhang, J., Chen, A., Billiar, T., Monga, S. P., & Wang, Q. (2015). ADAR1 prevents liver injury from inflammation and suppresses interferon production in hepatocytes. *American Journal of Pathology*, *185*(12), 3224–3237. https://doi.org/10.1016/j.ajpath.2015.08.002

Wang, J., Alvin Chew, B. L., Lai, Y., Dong, H., Xu, L., Balamkundu, S., Cai, W. M., Cui, L., Liu, C. F., Fu, X.-Y., Lin, Z., Shi, P.-Y., Lu, T. K., Luo, D., Jaffrey, S. R., & Dedon, P. C. (2019). Quantifying the RNA cap epitranscriptome reveals novel caps in cellular and viral RNA. *Nucleic Acids Research*, *47*(20), e130–e130. https://doi.org/10.1093/nar/gkz751

Wang, L., Wang, S., Wu, L., Li, W., Bray, W., Clark, A. E., Gonzalez, G. M., Wang, Y., Carlin, A. F., & Rana, T. M. (2023). PCIF1-mediated deposition of 5′-cap N6,2′-O-dimethyladenosine in ACE2 and TMPRSS2 mRNA regulates susceptibility to SARS-CoV-2 infection. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(5). https://doi.org/10.1073/pnas.2210361120

Wang, R., Wang, J., Acharya, D., Paul, A. M., Bai, F., Huang, F., & Guo, Y. L. (2014). Antiviral responses in mouse embryonic stem cells: Differential development of cellular mechanisms in type I interferon production and response. *Journal of Biological Chemistry*, *289*(36), 25186–25198. https://doi.org/10.1074/jbc.M113.537746

Wang, R., Wang, J., Paul, A. M., Acharya, D., Bai, F., Huang, F., & Guo, Y. L. (2013). Mouse embryonic stem cells are deficient in type i interferon expression in response to viral infections and double-stranded RNA. *Journal of Biological Chemistry*, *288*(22), 15926–15936. https://doi.org/10.1074/jbc.M112.421438

Wang, Y., Wang, A., Liu, Z., Thurman, A. L., Powers, L. S., Zou, M., Zhao, Y., Hefel, A., Li, Y., Zabner, J., & Au, K. F. (2019). *Single-molecule long-read sequencing reveals the chromatin basis of gene expression*. https://doi.org/10.1101/gr.251116.119

Wang, Y., Zhao, Y., & Bollas, A. ✉. (2021). Nanopore sequencing technology, bioinformatics and applications. *Nature Biotechnology*. https://doi.org/10.1038/s41587-021-01108-x

Wang Yanli, Ludwig, J., Schuberth, C., Goldeck, M., Schlee, M., Li, H., Juranek, S., Sheng, G., Micura, R., Tuschl, T., Hartmann, G., & Patel, D. J. (2010). Structural and functional insights into 5′-ppp RNA pattern recognition by the innate immune receptor RIG-I. *Nature Structural and Molecular Biology*, *17*(7), 781–787. https://doi.org/10.1038/nsmb.1863

Wei, C. M., Gershowitz, A., & Moss, B. (1976). 5′-Terminal and Internal Methylated Nucleotide Sequences in Hela Cell mRNA. *Biochemistry*, *15*(2), 397–401. https://doi.org/10.1021/bi00647a024

Wei Cha-Mer and Gershowitz, A. and M. B. (1975). N6, O2′-dimethyladenosine a novel methylated ribonucleoside next to the 5′ terminal of animal cell and virus mRNAs. *Nature*, *257*(5523), 251–253. https://doi.org/10.1038/257251a0

Weichmann, F., Hett, R., Schepers, A., Ito-Kureha, T., Flatley, A., Slama, K., Hastert, F. D., Angstman, N. B., Cardoso, M. C., König, J., Hüttelmaier, S., Dieterich, C., Canzar, S., Helm, M., Heissmeyer, V., Feederle, R., & Meister, G. (2020). *Validation strategies for antibodies targeting modified ribonucleotides*. https://doi.org/10.1261/rna

Werner, M., Purta, E., Kaminska, K. H., Cymerman, I. A., Campbell, D. A., Mittra, B., Zamudio, J. R., Sturm, N. R., Jaworski, J., & Bujnicki, J. M. (2011). 2^-O-ribose methylation of cap2 in human: function and evolution in a horizontally mobile family. *Nucleic Acids Research*, *39*(11), 4756–4768. https://doi.org/10.1093/nar/gkr038

Wiener, D., & Schwartz, S. (2021). The epitranscriptome beyond m6A. *Nature Reviews Genetics*, *22*(2), 119–131. https://doi.org/10.1038/s41576-020-00295-8

Wilamowski, M., Sherrell, D. A., Minasov, G., Kim, Y., Shuvalova, L., Lavens, A., Chard, R., Maltseva, N., Jedrzejczak, R., Rosas-Lemus, M., Saint, N., Foster, I. T., Michalska, K., Satchell, K. J. F., & Joachimiak, A. (2021). 2′-O methylation of RNA cap in SARS-CoV-2 captured by serial crystallography. *Proceedings of the National Academy of Sciences*, *118*(21). https://doi.org/10.1073/pnas.2100170118

Wilbertz, J. H., Voigt, F., Horvathova, I., Roth, G., Zhan, Y., & Chao, J. A. (2019). Single-Molecule Imaging of mRNA Localization and Regulation during the Integrated Stress Response. *Molecular Cell*, *73*(5), 946-958.e7. https://doi.org/10.1016/j.molcel.2018.12.006

Wilczynska, A., Gillen, S. L., Schmidt, T., Meijer, H. A., Jukes-Jones, R., Langlais, C., Kopra, K., Lu, W. T., Godfrey, J. D., Hawley, B. R., Hodge, K., Zanivan, S., Cain, K., Le Quesne, J., & Bushell, M. (2019). EIF4A2 drives repression of translation at initiation by Ccr4-Not through purine-rich motifs in the 5′UTR. *Genome Biology*, *20*(1). https://doi.org/10.1186/s13059-019-1857-2

Will, C. L., & Lührmann, R. (2011). Spliceosome structure and function. *Cold Spring Harbor Perspectives in Biology*, *3*(7), 1–2. https://doi.org/10.1101/cshperspect.a003707

Williams, G. D., Gokhale, N. S., Snider, D. L., & Horner, S. M. (2020). The mRNA Cap 2′- O -Methyltransferase CMTR1 Regulates the Expression of Certain Interferon-Stimulated Genes . *MSphere*, *5*(3). https://doi.org/10.1128/msphere.00202-20

Witteveldt, J., Knol, L. I., & Macias, S. (2019). MicroRNA-deficient mouse embryonic stem cells acquire a functional interferon response. *ELife*, *8*. https://doi.org/10.7554/eLife.44171

Worch, R., Niedzwiecka, A., Stepinski, J., Mazza, C., Jankowska-Anyszka, M., Darzynkiewicz, E., Cusack, S., & Stolarski, R. (2005). Specificity of recognition of mRNA 5′ cap by human nuclear cap-binding complex. *RNA*, *11*(9), 1355–1363. https://doi.org/10.1261/rna.2850705

Wu, S., Zhang, Y., Yao, L., Wang, J., Lu, F., & Liu, Y. (2023). m6A-modified RNAs possess distinct poly(A) tails. *Journal of Genetics and Genomics*, *50*(3), 208–211. https://doi.org/10.1016/J.JGG.2022.10.001

Xia, H., Zhong, C., Wu, X., Chen, J., Tao, B., Xia, X., Shi, M., Zhu, Z., Trudeau, V. L., & Hu, W. (2018). *Mettl3* Mutation Disrupts Gamete Maturation and Reduces Fertility in Zebrafish. *Genetics*, *208*(2), 729–743. https://doi.org/10.1534/genetics.117.300574

Xu, C., Wu, X., Zhang, X., Xie, Q., Fan, C., & Zhang, H. (2018). Embryonic Lethality and Host Immunity of RelA-Deficient Mice Are Mediated by Both Apoptosis and Necroptosis. *The Journal of Immunology*, *200*(1), 271–285. https://doi.org/10.4049/jimmunol.1700859

Xu, L., Yang, L., & Liu, W. (2013). Distinct evolution process among type I interferon in mammals. *Protein and Cell*, *4*(5), 383–392. https://doi.org/10.1007/s13238-013-3021-1

Yang, D.-H., Wang, S., Li, X., Yu, B., Wang, B., Wu, Z., Wu, C., Ling, J., Gao, X., & Zeng, H. (2021). Article 650287 (2021) LncRNA SNHG8 Promotes Proliferation and Inhibits Apoptosis of Diffuse Large B-Cell Lymphoma via Sponging miR-335-5p. *Front. Oncol*, *11*, 650287. https://doi.org/10.3389/fonc.2021.650287

Yang, W., Wang, Q., Howell, K. L., Lee, J. T., Cho, D. S. C., Murray, J. M., & Nishikura, K. (2005). ADAR1 RNA deaminase limits short interfering RNA efficacy in mammalian cells. *Journal of Biological Chemistry*, *280*(5), 3946–3953. https://doi.org/10.1074/jbc.M407876200

Yang, X., Triboulet, R., Liu, Q., Sendinc, E., & Gregory, R. I. (2022). Exon junction complex shapes the m6A epitranscriptome. *Nature Communications*, *13*(1), 7904. https://doi.org/10.1038/s41467-022-35643-1

Yang, X., Yang, Y., Sun, B.-F., Chen, Y.-S., Xu, J.-W., Lai, W.-Y., Li, A., Wang, X., Bhattarai, D. P., Xiao, W., Sun, H.-Y., Zhu, Q., Ma, H.-L., Adhikari, S., Sun, M., Hao, Y.-J., Zhang, B., Huang, C.-M., Huang, N., … Yang, Y.-G. (2017). 5-methylcytosine promotes mRNA export-NSUN2 as the methyltransferase and ALYREF as an m 5 C reader. *Cell Research*, *27*, 606–625. https://doi.org/10.1038/cr.2017.55

Yang, Y., & Wang, Z. (2019). IRES-mediated cap-independent translation, a path leading to hidden proteome. *Journal of Molecular Cell Biology*, *911*(10), 911–919. https://doi.org/10.1093/jmcb/mjz091

Yockey, L. J., Varela, L., Rakib, T., Khoury-Hanold, W., Fink, S. L., Stutz, B., Szigeti-Buck, K., Van den Pol, A., Lindenbach, B. D., Horvath, T. L., & Iwasaki, A. (2016). Vaginal Exposure to Zika Virus during Pregnancy Leads to Fetal Brain Infection. *Cell*, *166*(5), 1247-1256.e4. https://doi.org/10.1016/j.cell.2016.08.004

You, A. Bin, Yang, H., Lai, C. P., Lei, W., Yang, L., Lin, J. L., Liu, S. C., Ding, N., & Ye, F. (2023). CMTR1 promotes colorectal cancer cell growth and immune evasion by transcriptionally regulating STAT3. *Cell Death & Disease*, *14*(4), 245. https://doi.org/10.1038/s41419-023-05767-3

Yuan, X., Yan, Y., & Xue, M. (2021). Small nucleolar RNA host gene 8: A rising star in the targets for cancer therapy. In *Biomedicine and Pharmacotherapy* (Vol. 139). Elsevier Masson s.r.l. https://doi.org/10.1016/j.biopha.2021.111622

Zhang, J., Huang, J., Xu, K., Xing, P., Huang, Y., Liu, Z., Tong, L., & Manley, J. L. (2022). DHX15 is involved in SUGP1-mediated RNA missplicing by mutant SF3B1 in cancer. *Proceedings of the National Academy of Sciences*, *119*(49). https://doi.org/10.1073/pnas.2216712119

Zhang, R., Jha, B. K., Ogden, K. M., Dong, B., Zhao, L., Elliott, R., Patton, J. T., Silverman, R. H., & Weiss, S. R. (2013). Homologous 2?,5?-phosphodiesterases from disparate RNA viruses

antagonize antiviral innate immunity. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(32), 13114–13119. https://doi.org/10.1073/pnas.1306917110

Zhang, Y., Lu, L., & Li, X. (2022). Detection technologies for RNA modifications. *Experimental & Molecular Medicine*, *54*(10), 1601–1616. https://doi.org/10.1038/s12276-022-00821-0

Zhao, J., Qin, C., Liu, Y., Rao, Y., & Feng, P. (2021). Herpes Simplex Virus and Pattern Recognition Receptors: An Arms Race. *Frontiers in Immunology*, *11*. https://doi.org/10.3389/fimmu.2020.613799

Zhao, Y., Ye, X., Dunker, W., Song, Y., & Karijolich, J. (2018). RIG-I like receptor sensing of host RNAs facilitates the cell-intrinsic immune response to KSHV infection. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-07314-7

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., … He, C. (2013). ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. *Molecular Cell*, *49*(1), 18–29. https://doi.org/10.1016/j.molcel.2012.10.015

Zhong, S., Li, H., Bodi, Z., Button, J., Vespa, L., Herzog, M., & 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*(5), 1278–1288. https://doi.org/10.1105/tpc.108.058883

Zhou, X., Liao, W. J., Liao, J. M., Liao, P., & Lu, H. (2015). Ribosomal proteins: Functions beyond the ribosome. *Journal of Molecular Cell Biology*, *7*(2), 92–104. https://doi.org/10.1093/jmcb/mjv014

Zhu, Y., Pirnie, S. P., & Carmichael, G. G. (2017). *High-throughput and site-specific identification of 2′ ′′′′-O-methylation sites using ribose oxidation sequencing (RibOxi-seq)*. https://doi.org/10.1261/rna

Zimta, A.-A., Tigu, A. B., Braicu, C., Stefan, C., Ionescu, C., & Berindan-Neagoe, I. (2020). An Emerging Class of Long Non-coding RNA With Oncogenic Role Arises From the snoRNA Host Genes. *Frontiers in Oncology*, *10*. https://doi.org/10.3389/fonc.2020.00389

Zou, Y. (2006). Navigating the Anterior-Posterior Axis with Wnts. *Neuron*, *49*(6), 787–789. https://doi.org/10.1016/j.neuron.2006.03.004

Züst, R., Cervantes-Barragan, L., Habjan, M., Maier, R., Neuman, B. W., Ziebuhr, J., Szretter, K. J., Baker, S. C., Barchet, W., Diamond, M. S., Siddell, S. G., Ludewig, B., & Thiel, V. (2011). Ribose 2^-O-methylation provides a molecular signature for the distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nature Immunology*, *12*(2), 137–143. https://doi.org/10.1038/ni.1979

Publications from my PhD

Dohnalkova M, Krasnykov K, Mendel M, Li L, Panasenko O, Fleury-Olela F, Vågbø CB, Homolka D, Pillai RS. Essential roles of RNA cap-proximal ribose methylation in mammalian embryonic development and fertility. **Cell Rep**. 2023 Jul 11;42(7):112786. doi: 10.1016/j.celrep.2023.112786. Epub ahead of print. PMID: 37436893.