The Epitranscriptome

The RNA code

What is the epitrascriptome?

The epitranscriptome is the world of RNA modification

- Cellular RNAs contain more than 100 chemical modifications.
- The RNA modifications can occur both on the atomes of the base or the sugar and on the phosfate backbone
- Rna modifications represent a new layer of gene expression regulation

What is the epitrascriptome?

What do we mean for RNA modification?

There are three kinds of RNA modifications:
- addition of nucleotides
- changes of nucleotides
- **chemical modifications**

What is the epitrascriptome?

What's the effect of a mRNA modification?

It depends on both the molecular consequences and the percentage of transcripts that are modified

1. Addition of nucletides

A polyU-tail can be added to the RNA

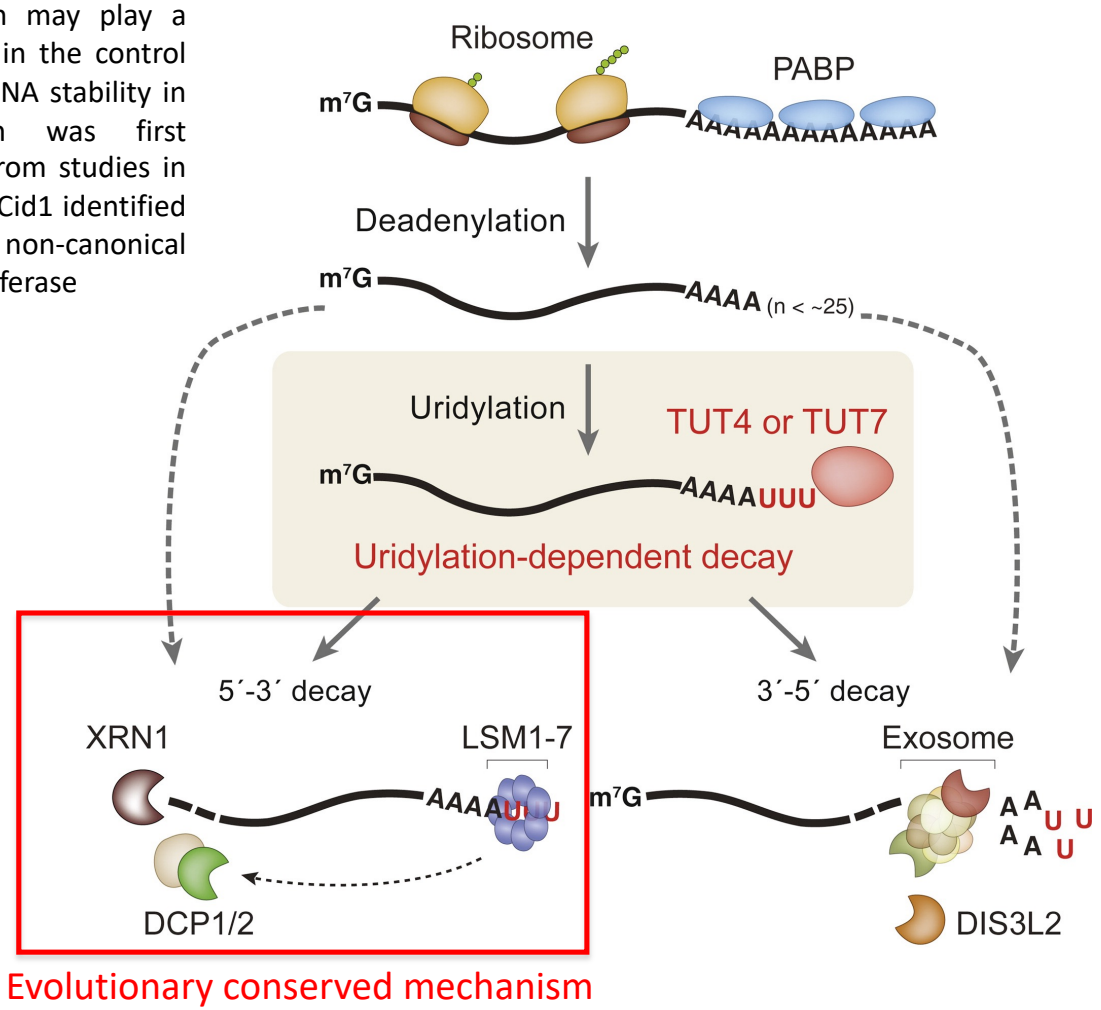
The 25% of mRNA is uridylated.

Histons mRNAs are uridylated by enzymes named TUT or ZCCHC11.

The polyU-tail recruits decapping enzymes and exonucleases

Uridylation-dependent mRNA decapping

That uridylation may play a prominent role in the control of poly(A)+ mRNA stability in the cytoplasm was first demonstrated from studies in S. pombe with, Cid1 identified as the first non-canonical nucleotidyltransferase



Evolutionary conserved mechanism

1. Addition of nucletides

A polyU-tail can be added to the RNA

The 25% of mRNA is uridylated.

Histons mRNAs are uridylated by enzymes named TUT or ZCCHC11.

The polyU-tail recruits decapping enzymes and exonucleases

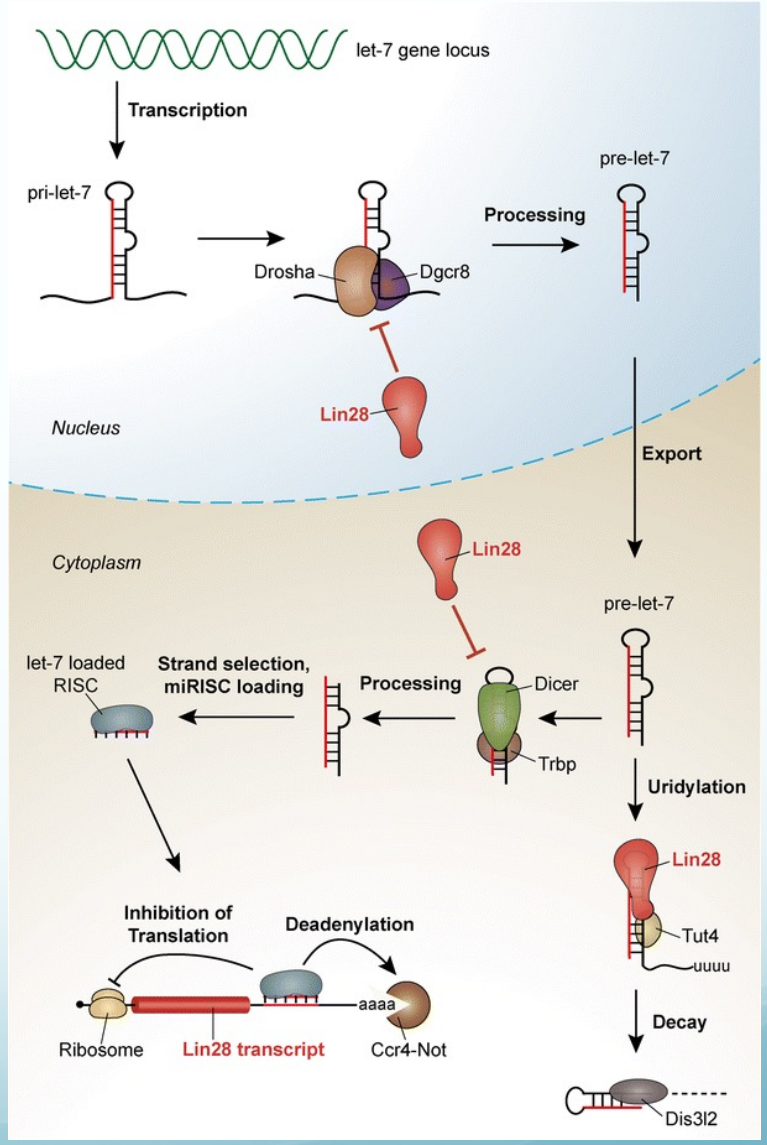
1. Addition of nucletides

A polyU-tail can be added to the RNA

LIN28 binds Let7 and recruits TUT4 to promote its decay

Not only decay:

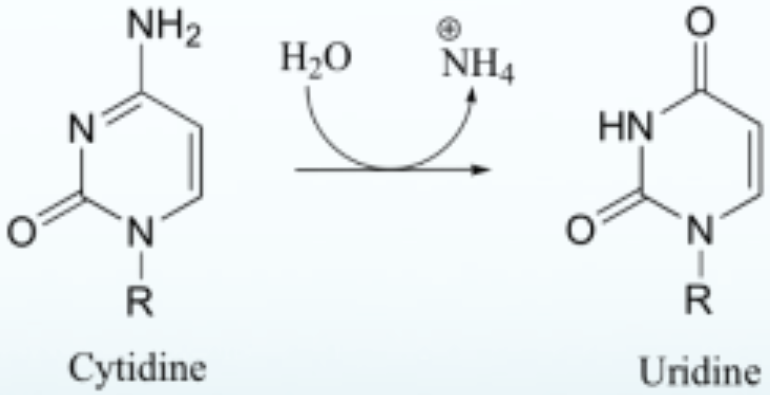
The group II of microRNA are poliuridylated by TUT7 to become DICER better substrate



2. Changes of nucleotides

C-to-U editing

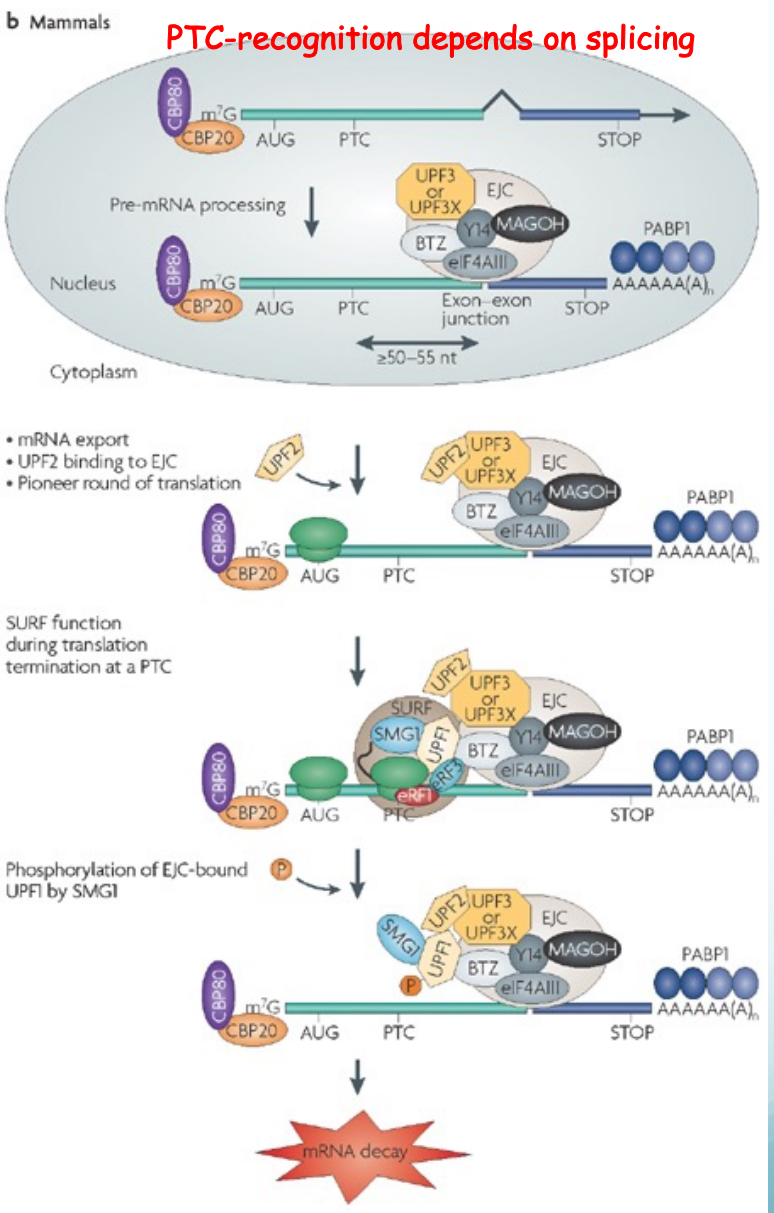
A cytidine base is deaminated into a uridine base.
The apolipoprotein B gene in humans: Apo B100 is expressed in the liver and Apo B48 is expressed in the intestines. In the intestines, the mRNA has a CAA sequence edited to be UAA, a stop codon, thus producing the shorter B48 form.



NMD (mammals)

•In mammalian cells, a **large exon junction protein complex (EJC)** deposited about 20–24 nucleotide (nt) upstream of exon–exon junctions during RNA splicing, is widely considered to **be <u>a mark that triggers NMD</u>**

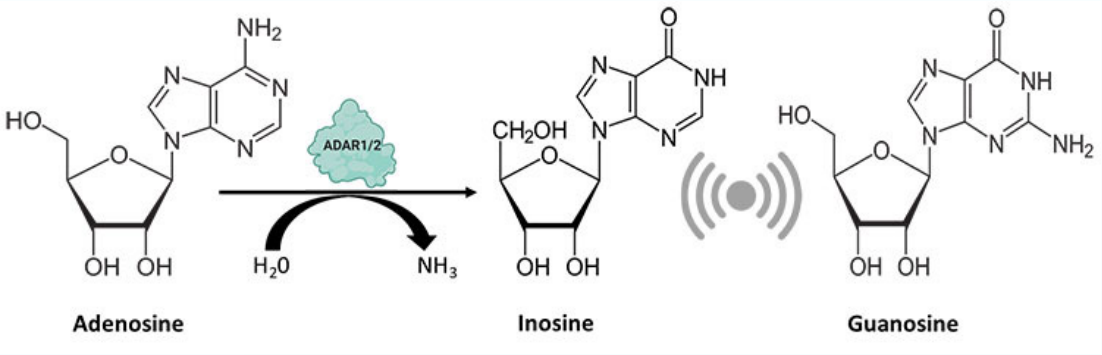
•**Nonsense codons more than 55-nt upstream of the last intron generally trigger NMD**



2. Changes of nucleotides

A-to-I editing

A-to-I editing is the main form of RNA editing in mammals and occurs in dsRNA. Adenosine deaminases acting on RNA (ADARs) are the RNA-editing enzymes involved in the hydrolytic deamination of Adenosine to Inosine.

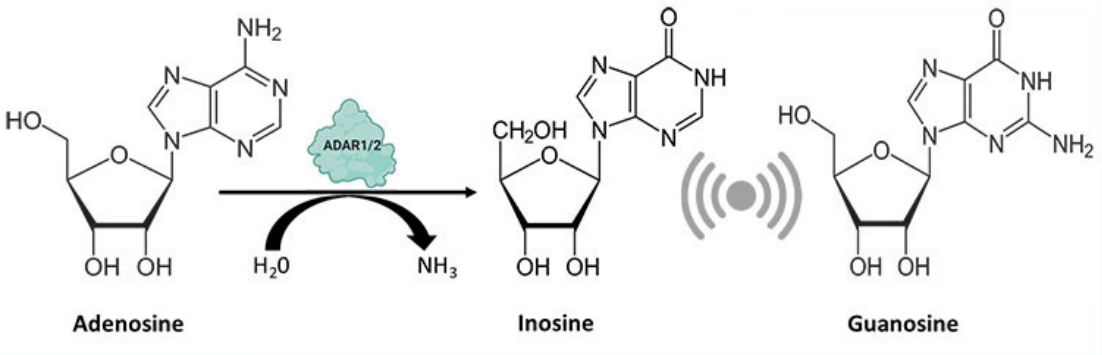


A-to-I editing can be specific (a single adenosine is edited within the stretch of dsRNA) or **promiscuous** (up to 50% of the adenosines are edited). **I behaves as G** both in translation and when forming secondary structures.

2. Changes of nucleotides

A-to-I editing

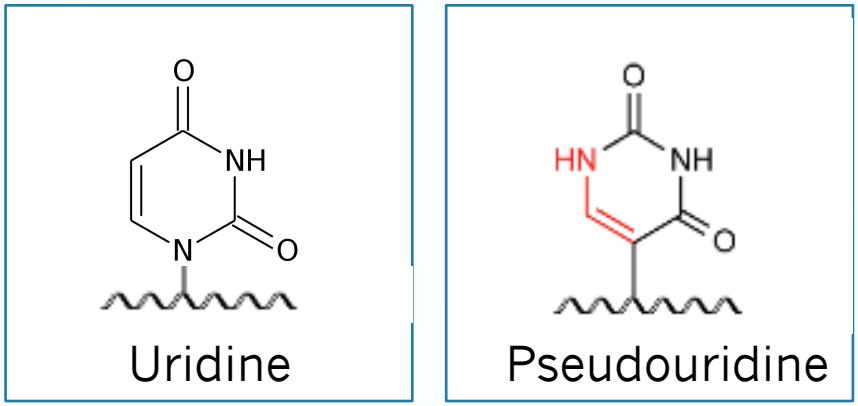
A-to-I editing is the main form of RNA editing in mammals and occurs in dsRNA. Adenosine deaminases acting on RNA (ADARs) are the RNA-editing enzymes involved in the hydrolytic deamination of Adenosine to Inosine.



A-to-I editing can be specific (a single adenosine is edited within the stretch of dsRNA) or **promiscuous** (up to 50% of the adenosines are edited). **I behaves as G** both in translation and when forming secondary structures. These effects include alteration of coding capacity, altered miRNA or siRNA target populations, heterochromatin formation, nuclear sequestration, cytoplasmic sequestration, endonucleolytic cleavage by Tudor-SN, inhibition of miRNA and siRNA processing, and altered splicing...

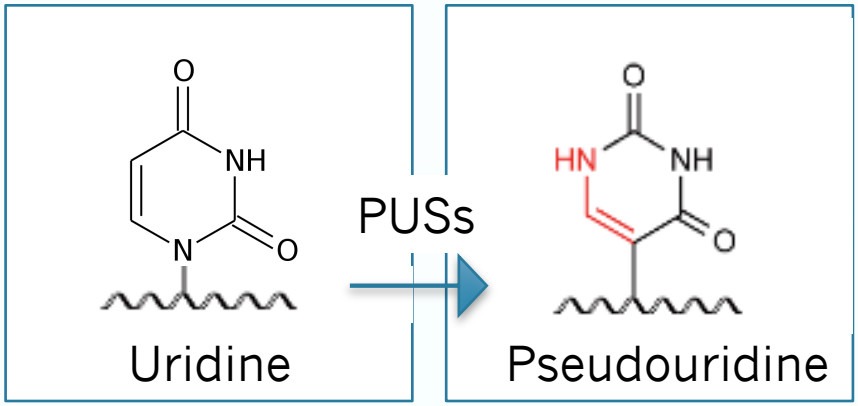
3. Chemical modifications

Pseudouridine: the fifth nucleotide



- Pseudouridine is an isomer of uridine
- Ψ has the same molecular weight and base pairing pattern with regular U, An additional H-bond donor in the N1H of Ψ can bind a water molecule to bridge the interactions of this N1H and the preceding phosphate groups and stabilize the RNA structure.
- Ψ is the most abundant RNA modification, present in tRNA, rRNA, and snRNA; ~~until recently, it had not been found in mRNA and lncRNA.~~ The function of Ψs in mRNA remains essentially unknown.

Pseudourydilation

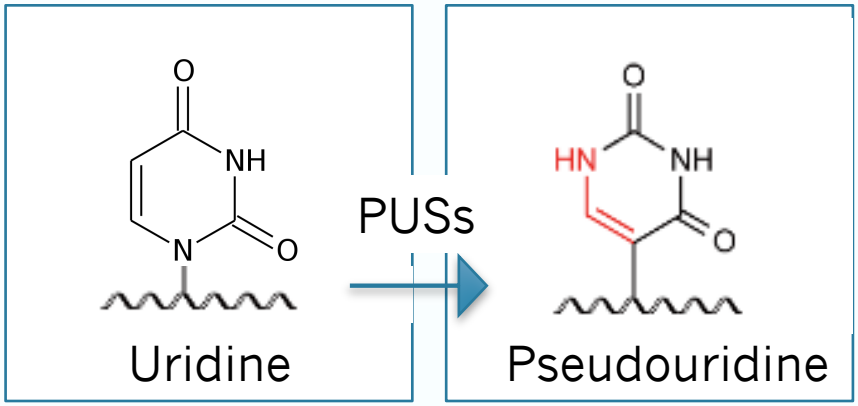


two pathways in eukaryotes:

1. one is catalyzed by the RNA-dependent pseudouridine synthases (PUSs) that require the cofactor box H/ACA ribonucleoproteins as guides to recognize different substrates;
2. another is catalyzed by the PUSs that require no cofactor.

Ten PUSs were identified in yeast: Cbf5 is an RNA-dependent PUS, and the rest (Pus1–9) are RNA independent.

Pseudourydilation



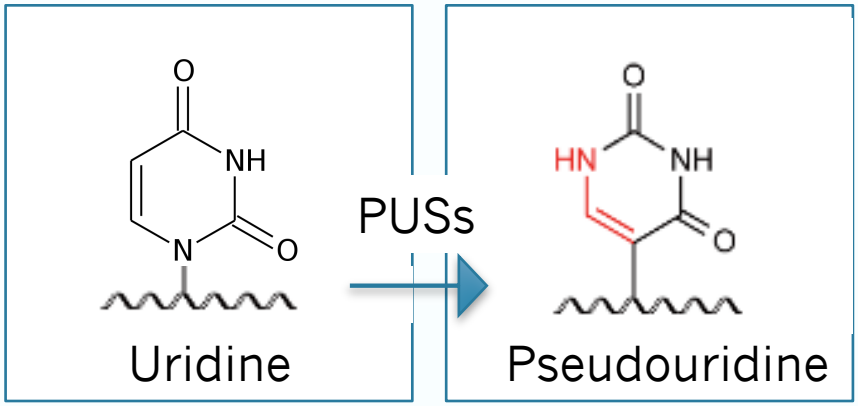
YEAST: Ten PUSs were identified Cbf5 is an RNA-dependent PUS, the rest (Pus1–9) are RNA independent.

HUMAN: DKC1 (a homologue of Cbf5) is an RNA-dependent PUS.

The study of Ψ is complicated by multiple PUSs and potential functional redundancy.

The inert C–C glycosidic bond of Ψ makes it difficult to be reversible. Interestingly, Ψ can be further methylated to 1-methylpseu- douridine (m1Ψ) by EMG1,65 potentially offering a mechanism to reduce the level of Ψ.

Why is pseudourydilation important?

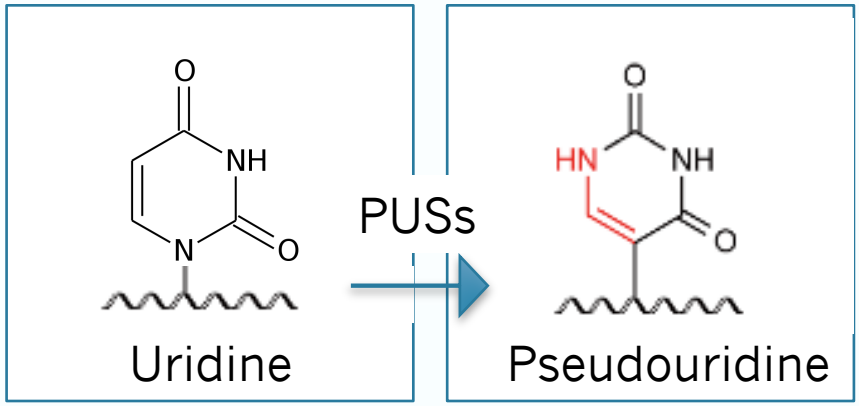


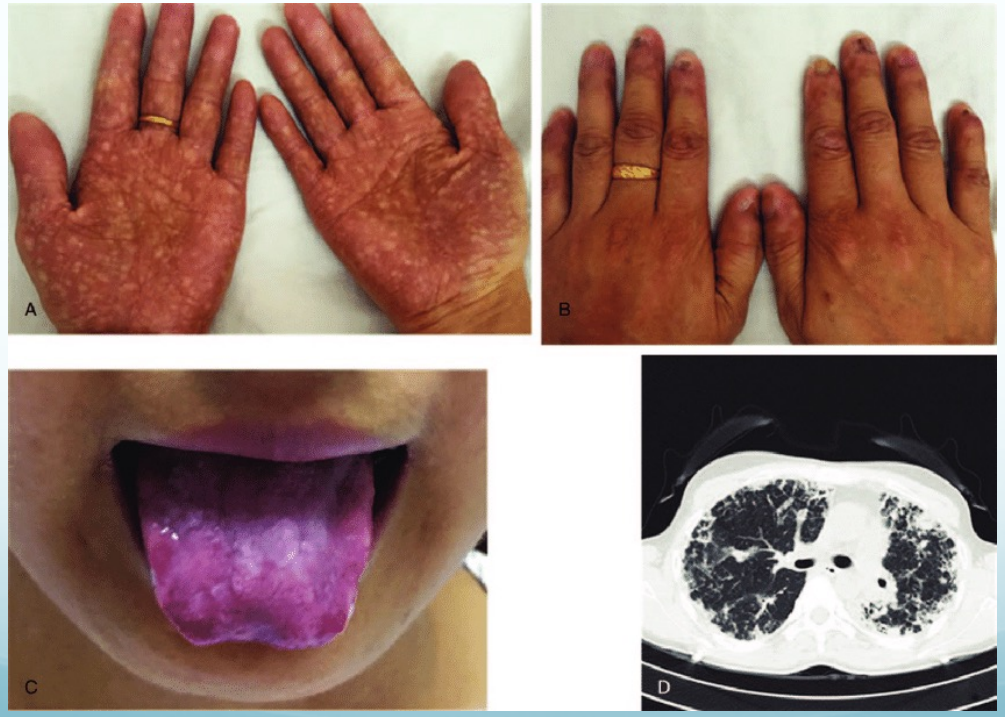
DKC1 is associated with dyskeratosis congenita, characterized by increase risk of cancer and reduced cell proliferation.

Dyskerin is required for hematopoietic stem cell differentiation

rRNA is pseudouridylated: defects in pseudourydilation affects IRES-dependent translation of important genes such as P53, P27, XIAP and BCL-XL

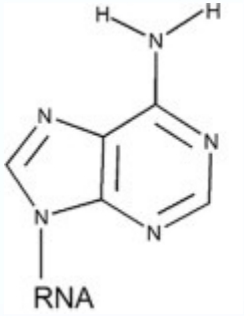
Why is pseudourydilation important?

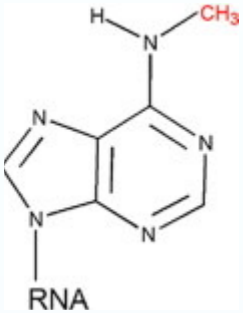




*N*6-methyladenosine (m6A)

m6A is widely **conserved** across plants and vertebrates, viruses, archaea, bacteria and yeast

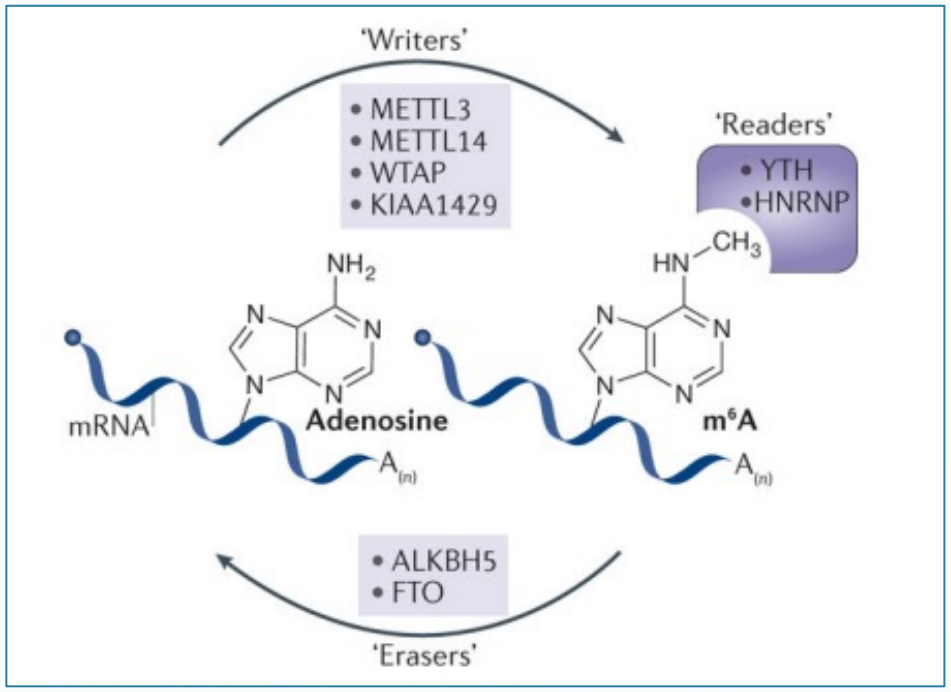




0.1–0.4% of all adenosine nucleotides in mammals are m6A
0.7–0.9% of adenine nucleotides (all within GA dinucleotides) in meiotic Saccharomyces cerevisiae
1–15 sites per virion RNA molecule in various viruses.

Mutation and in vitro enzymatic studies have identified a consensus motif of RRm6ACH ([G/A/U][G>A]m6AC[U>A>C]).

Who is the m6A writer?

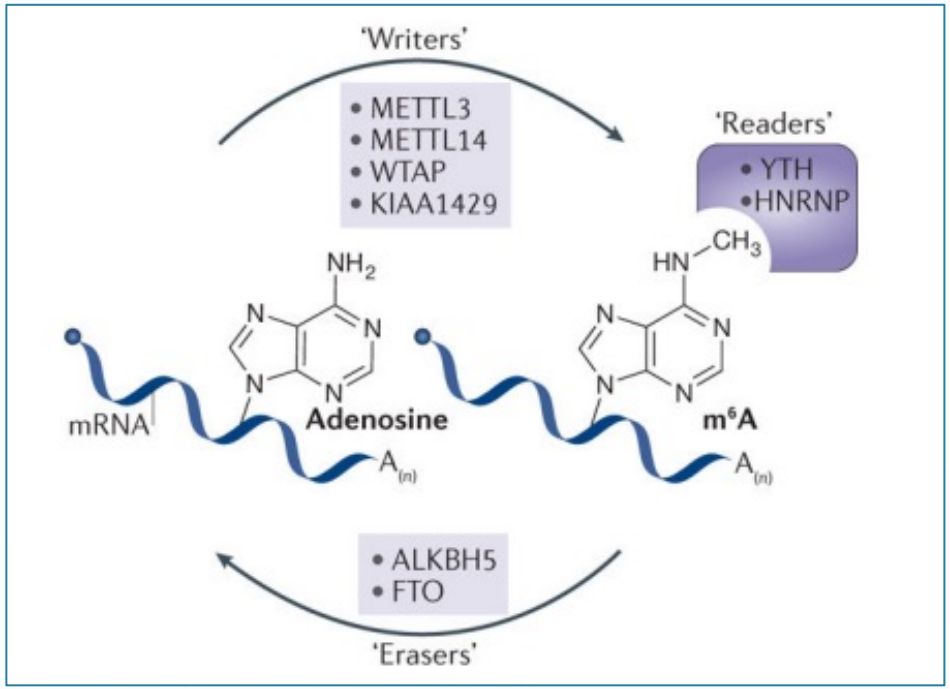


The deposition of m6A is carried out by a multicomponent methyltransferase complex that was first reported in 1994.

A key protein, methyltransferase-like 3 (**METTL3**), was subsequently identified as an S- adenosyl methionine-binding protein with methyltransferase capacity.

The m6A methyltransferase complex in mammals includes **METTL14**, Wilms tumour 1-associated protein (**WTAP**) and **KIAA1429**.

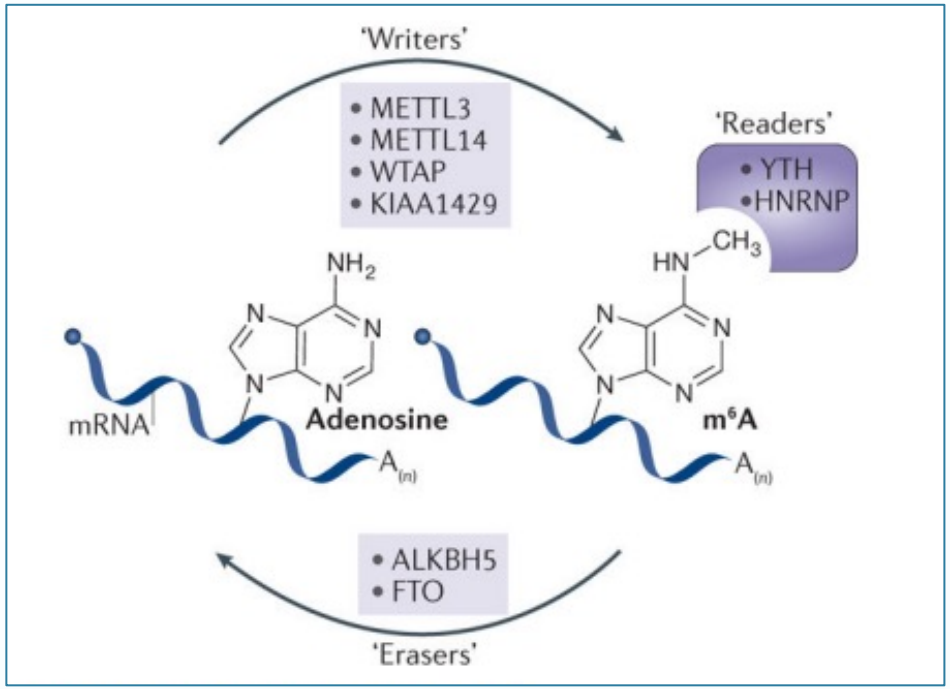
Who are the erasers?



The removal of m6A is facilitated by fat mass and obesity-associated protein (**FTO?**) and alkB homologue 5 (**ALKBH5**)

m6A is the first evidence of reversible post-transcriptional modification in RNA transcribed by RNA polymerase II

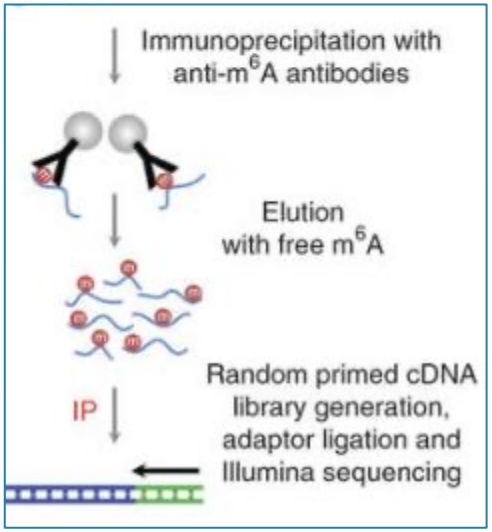
Who are the readers?

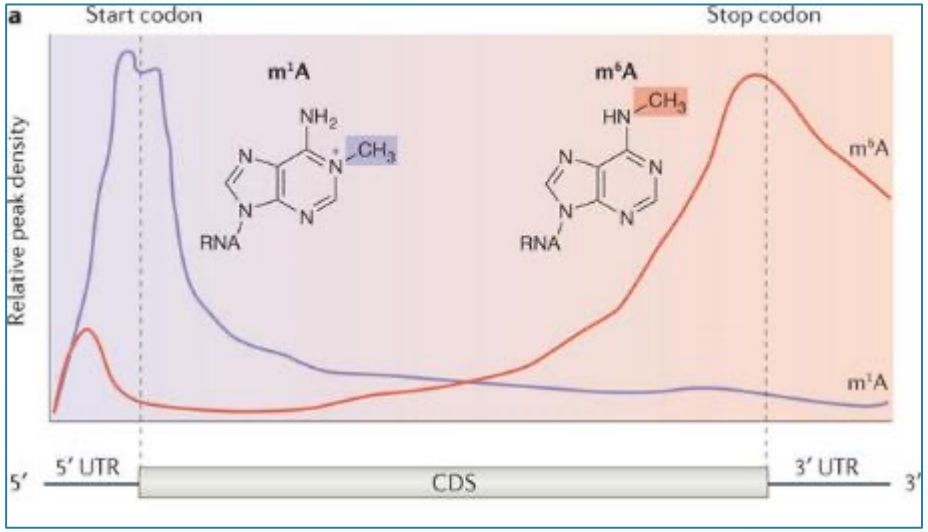


The YT521-B homology (YTH) domain family of proteins (**YTHDF1, YTHDF2, YTHDF3 and YTHDC1**) are direct readers of m6A and have a conserved m6A-binding pocket. In addition, the heterogeneous nuclear ribonucleoprotein (HNRNP) proteins HNRNPA2B1 and HNRNPC selectively bind m6A-containing mRNAs.

Where are m6As located?

Where are m6As located?

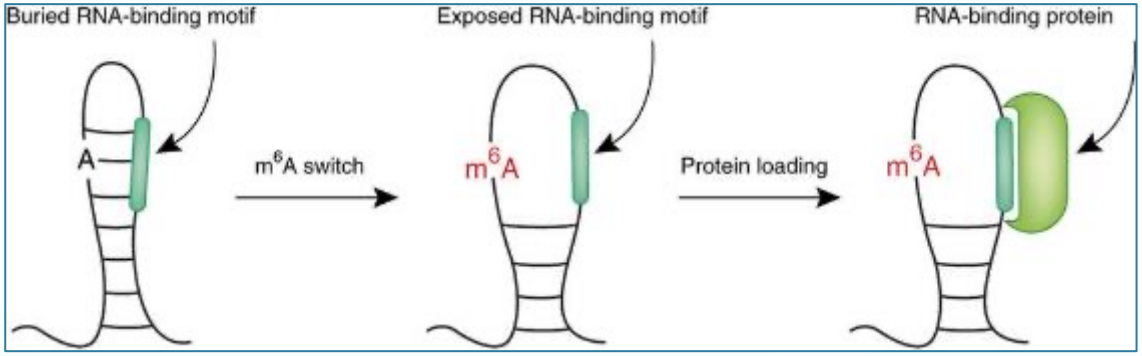




m6A-specific antibodies were used for immunoprecipitation followed by high-throughput sequencing to generate transcriptome-wide maps of m6A, charting the m6A epitranscriptome.
These studies have uncovered the presence of more than 10,000 m6A sites in over 25% of human transcripts, with enrichment in long exons, near stop codons and in 3′ untranslated regions (3′ UTRs).

m6A modulates mRNA metabolism

m6A alters RNA folding and structure



m6A does not change Watson–Crick A•U base pairing but weakens duplex RNA.

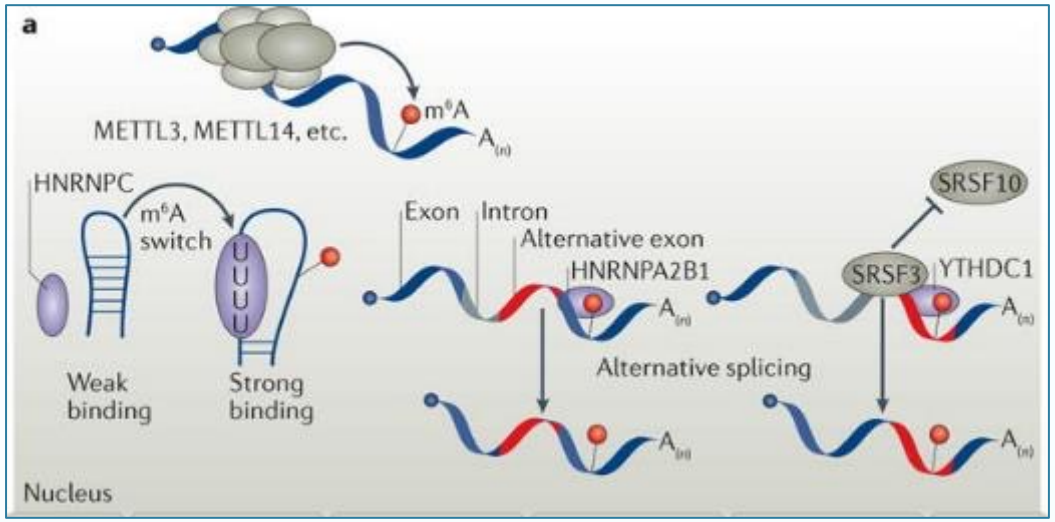
m6A stacks better than an unmodified base, stabilizing surrounding RNA structures or promoting the folding of adjacent RNA sequences.

m6A within coding regions could induce steric constraints that destabilize pairing between codons and tRNA anticodons, thus affecting translation dynamics.

m6A-triggered structural remodelling may change the accessibility of RBP interaction motifs to RNA: the **m6A switch**.

m6A modulates mRNA metabolism

m6A affects mRNA maturation



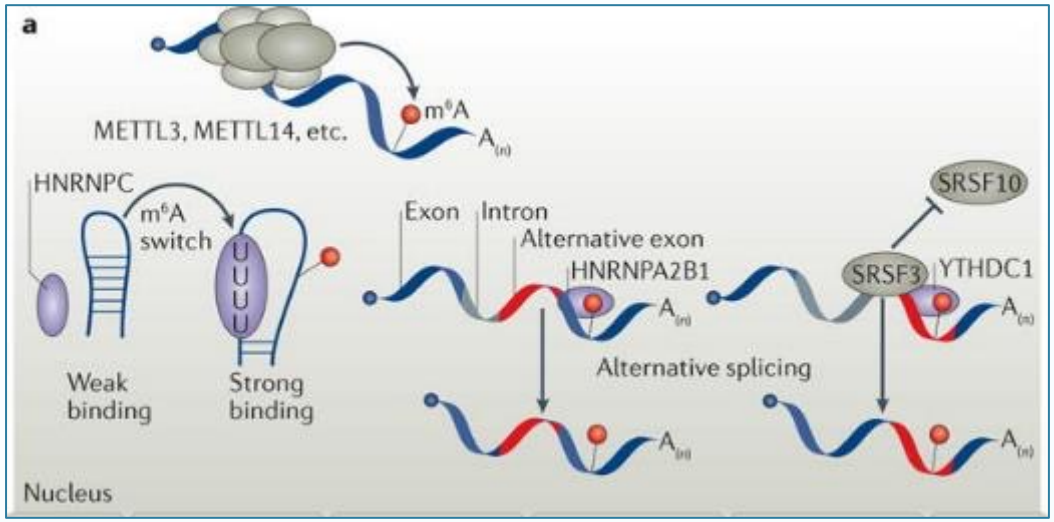
m6A is more abundant in pre-mRNA than in mature.

mRNAs that undergo alternative splicing have more METTL3-binding sites and more N6-adenosine methylation sites.

Writers and erasers of m6A localize predominantly in nuclear speckles

m6A modulates mRNA metabolism

m6A affects mRNA maturation



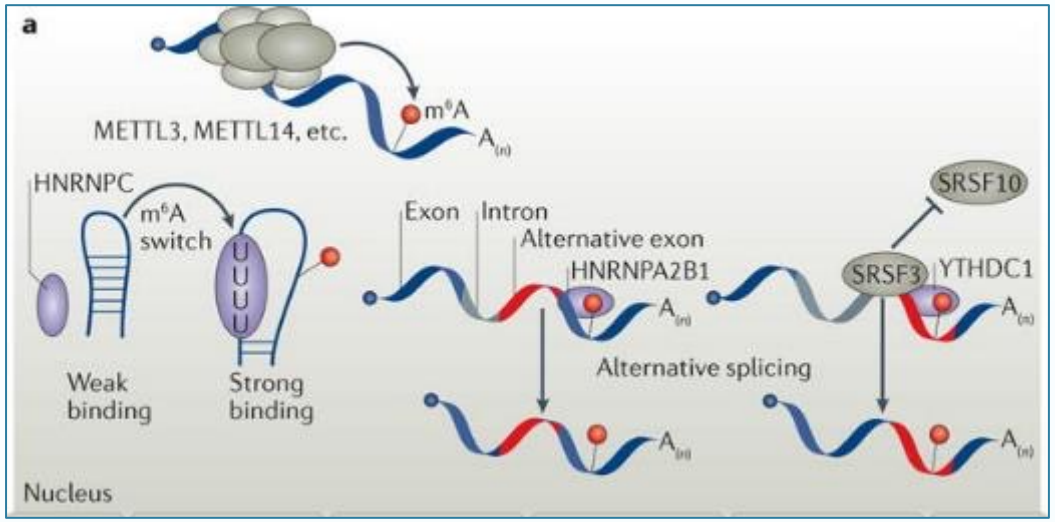
removing m6A around splice sites can prevent binding of serine- and arginine-rich splicing factor 2 (SRSF2).

m6A readers affect splicing:

- YTHDC1 recruits SRSF3 while blocking binding by SRSF10, leading to exon inclusion.
- HNRNPA2B1 regulates alternative splicing events in a similar manner to METTL3, as well as microRNA (miRNA) biogenesis from intronic sequences

m6A modulates mRNA metabolism

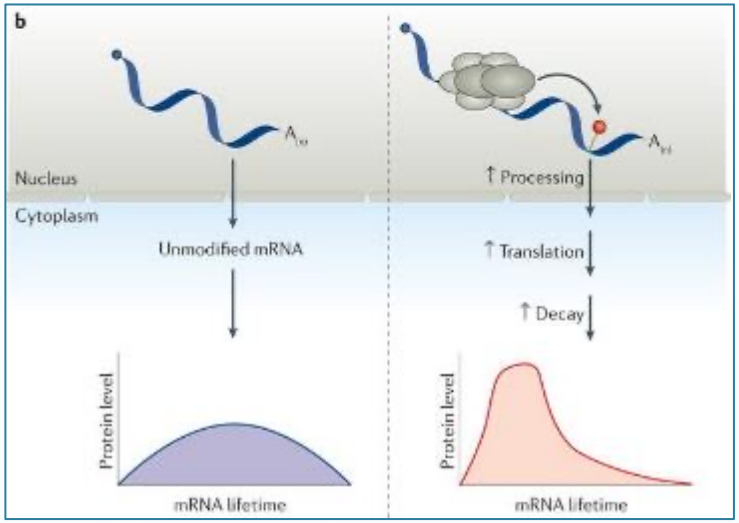
m6A affects mRNA maturation



Alternative polyadenylation (APA) is coupled to splicing of the last intron and associated with mRNA N6-adenosine methylation. Two-thirds of m6A sites in the last exon are found at the 3' UTR and the knockdown of m6A writers can cause APA.

m6A modulates mRNA metabolism

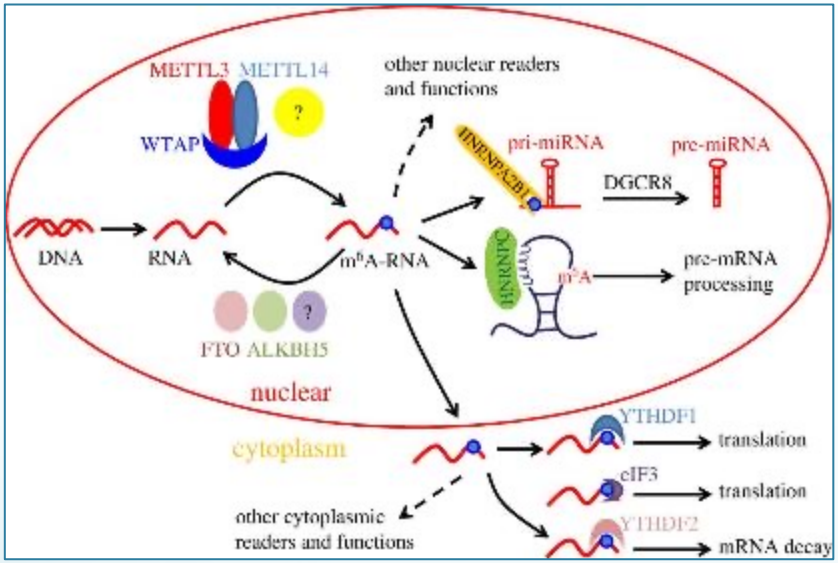
m6A enhances nuclear processing and export of mRNAs

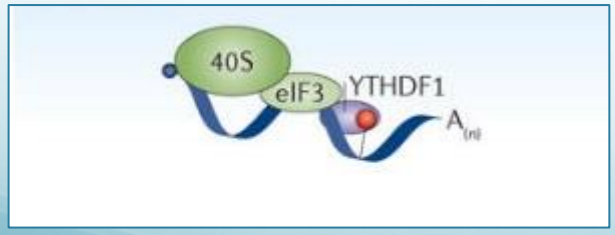


mRNA nuclear export can selectively modulate gene expression. N6-adenosine methylation was suggested to promote mRNA export**: depletion of METTL3 inhibited mRNA export**, whereas depletion of ALKBH5 enhanced mRNA export to the cytoplasm.

m6A modulates mRNA metabolism

m6A promotes mRNA translation



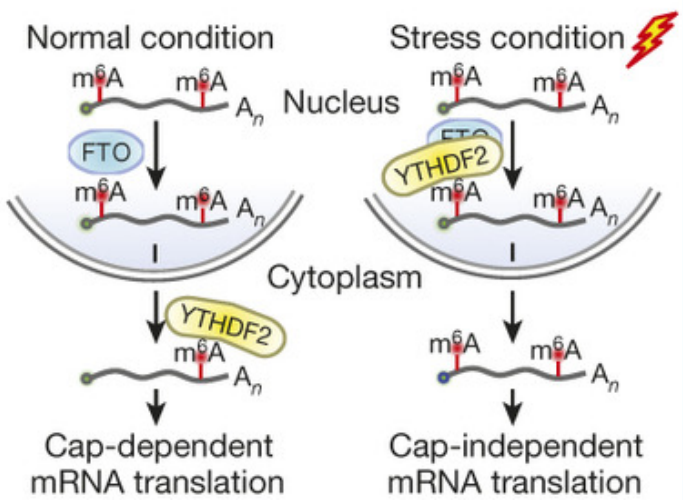


YTHDF1 can recruit translation initiation factors, improving the efficiency of cap-dependent translation.
YTHDF1 recruits the translation initiation factor complex eukaryotic initiation factor 3 (eIF3).
METTL3 functions also as an m6A reader by enhancing eIF4E-dependent translation in a specific subset of mRNAs by recruiting eIF3 during translation initiation.

Two recent studies have further indicated that the presence of m6A at the 5′ UTR improves cap-independent translation and eIF3 was proposed to interact with m6A and facilitate ribosome loading.

m6A modulates mRNA metabolism

m6A promotes mRNA translation

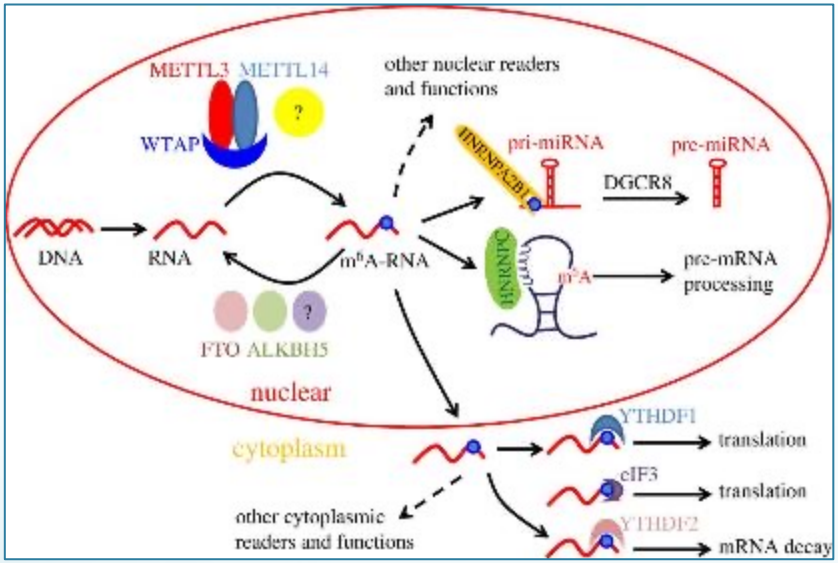


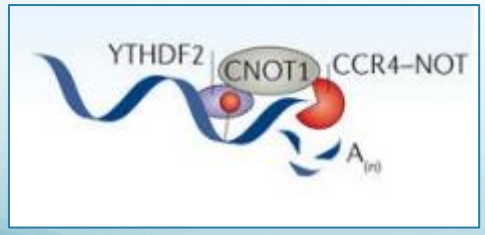
Dynamic m(6)A mRNA methylation directs translational control of heat shock response.

Upon heat shock stress, YTHDF2 translocates in the nucleus and preserves 5'UTR methylation of stress-induced transcripts by limiting the m(6)A 'eraser' FTO from demethylation. Remarkably, the increased 5'UTR methylation in the form of m(6)A promotes cap-independent translation initiation, providing a mechanism for selective mRNA translation under heat shock stress.

m6A modulates mRNA metabolism

m6A marks mRNA for decay





m6A has been linked to reduced mRNA stability, as knockdown of METTL3 and METTL14 in human and mouse cells has been shown to lead to increases in the expression of their respective target mRNAs.

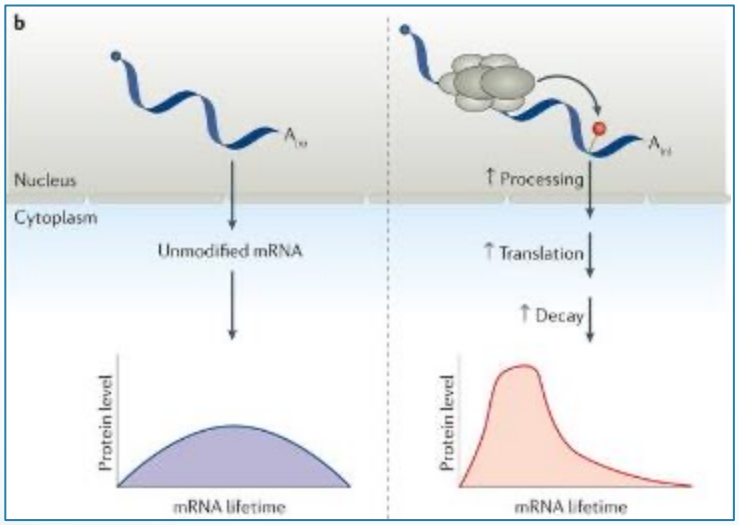
Many mRNAs encoding transcription repressors are target substrates of the methyltransferase complex; reduced methylation of these transcripts could cause transcription repression.

Knockdown of YTHDF2 prolonged the stability of its target mRNAs. **YTHDF2 induces accelerated deadenylation** of YTHDF2-bound mRNA.

m6A has a role in miRNA biogenesis

m6A modulates mRNA metabolism

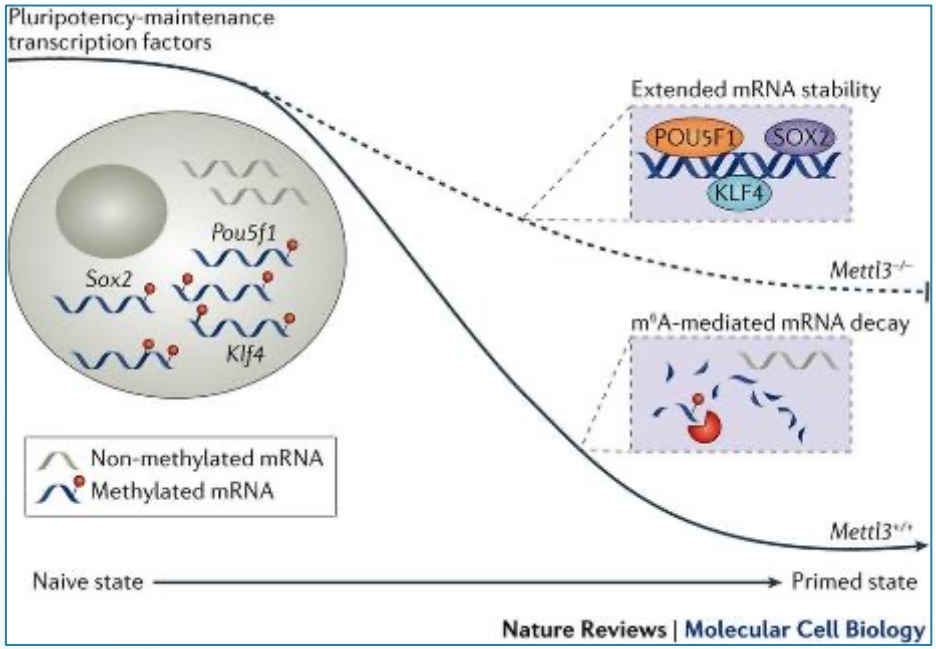
m6A can sort transcripts groups



Readers decode these messages and may functionally sort methylated mRNAs into distinct functional groups.

Methylation could be particularly beneficial in grouping and synchronizing the expression of hundreds to thousands of mRNAs that otherwise may possess markedly different properties with varied stabilities and translation efficiencies.

m6A shapes cell function and identity



m6A functions in circadian rhythm maintenance and cell cycle regulation

m6A functions in cell differentiation and reprogramming

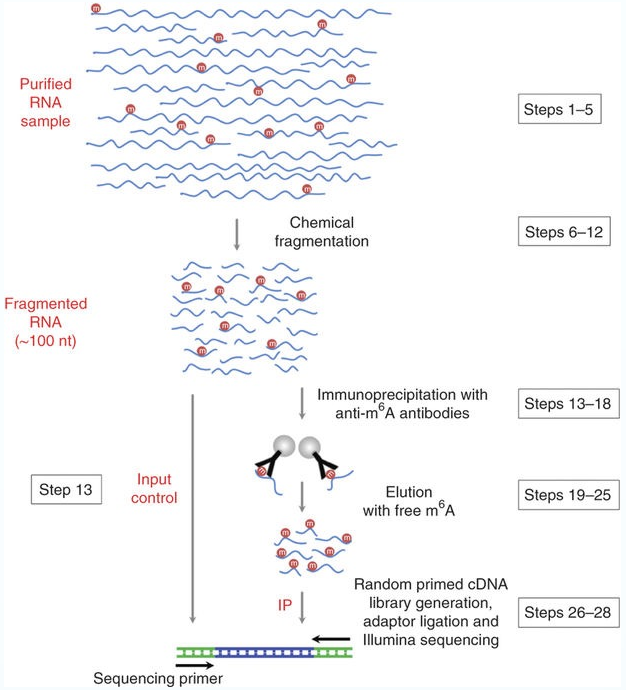
m6A facilitates cell state transitions

m6A functions in stress responses

m6A-seq and miCLIP

Transcriptome-wide mapping of N^6-methyladenosine by m^6A-seq based on immunocapturing and massively parallel sequencing

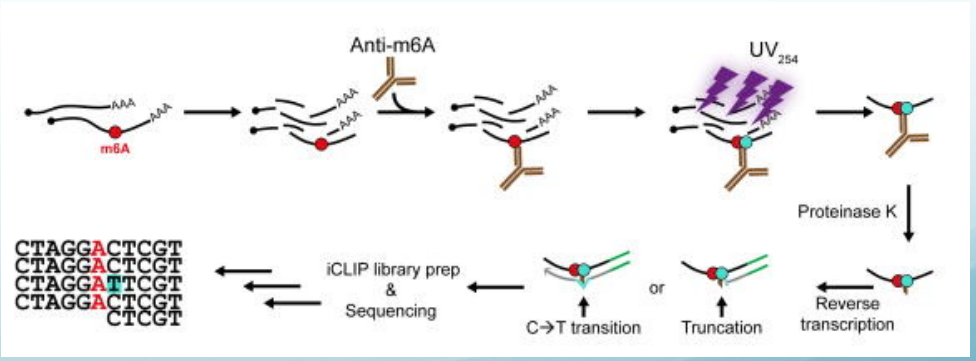
Dan Dominissini, Sharon Moshitch-Moshkovitz, Mali Salmon-Divon, Ninette Amariglio & Gideon Rechavi



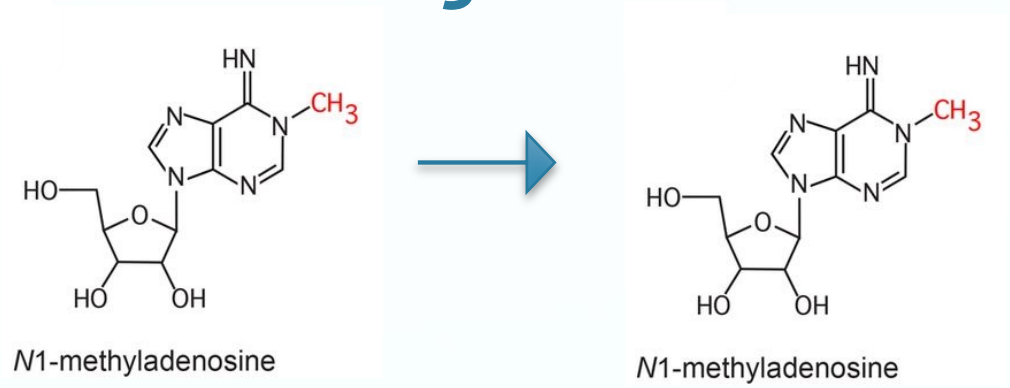
NATURE METHODS | ARTICLE

Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome

Bastian Linder, Anya V Grozhik, Anthony O Olarerin-George, Cem Meydan, Christopher E Mason & Samie R Jaffrey



*N*1-methyladenosine

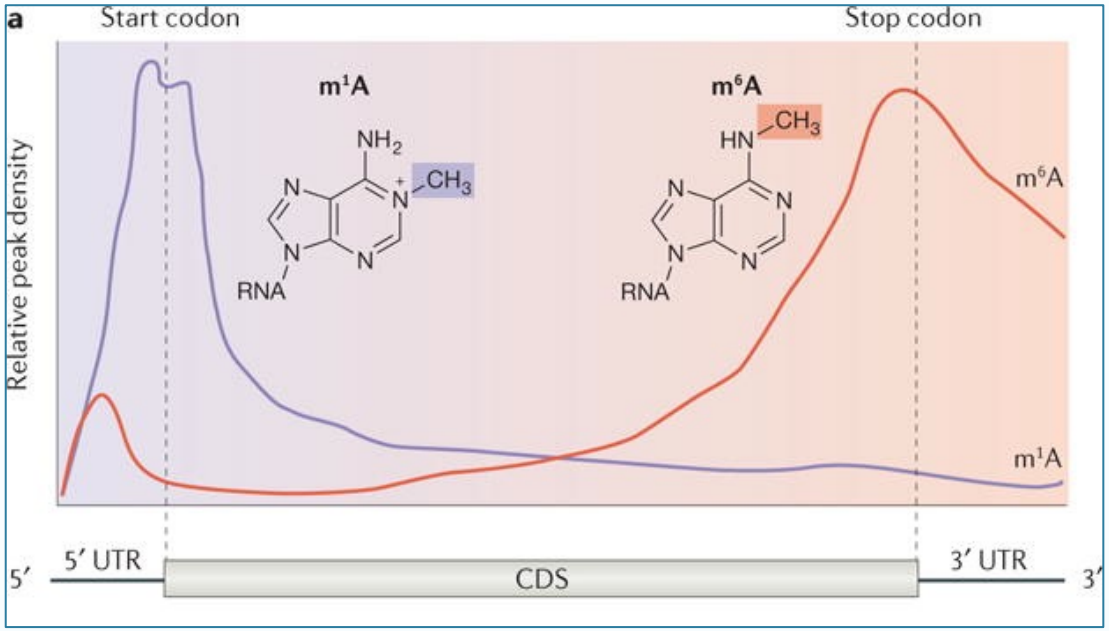


m1A regulates the structure and **stability** of tRNA and rRNA but was revealed also in eukaryo tic mRNA

The positive charge associated with this modification could alter RNA secondary structures.

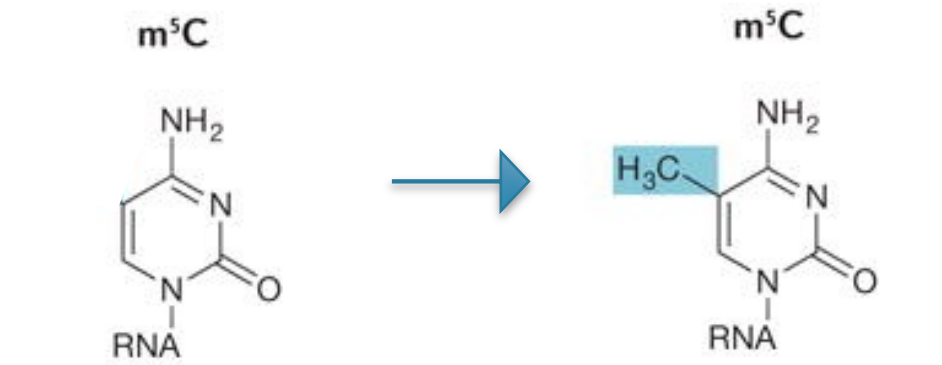
Methylated RNA immunoprecipitation followed by high-throughput sequencing was used to map more than 7,000 *m*1A locations in coding and long non-coding

*N*1-methyladenosine



The distribution of *m*1A in mRNAs is unique in its proximity to translation starting sites and the first splice site

5-Methylcytosine



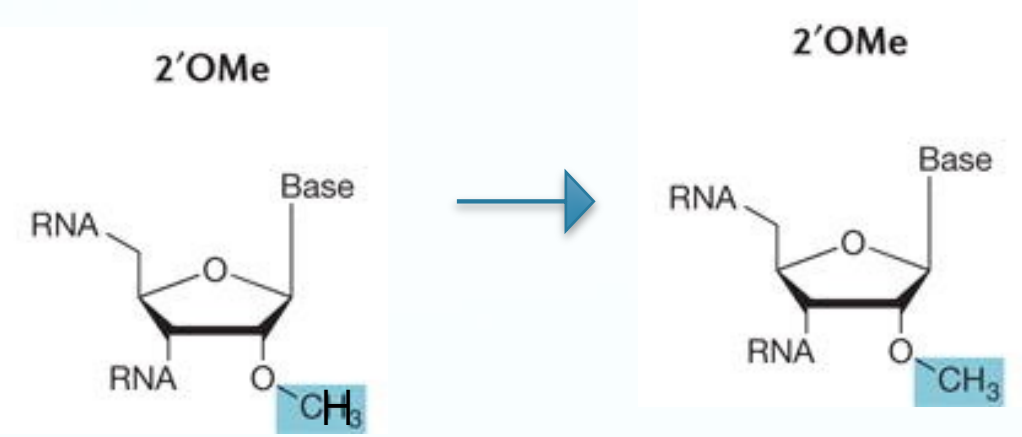
m5C has long been studied as an epigenetic modification in DNA. Adopting the bisulfite treatment several m5C sites in tRNA and rRNA have been characterized

tRNA methyltransferase 4 (Trm4) is a tRNA 5-cytosine methyltransferase in yeast

tRNA aspartic acid methyltransferase 1 (also known as **Dnmt2**) was reported to be a tRNA m5C writer in several eukaryotic species and shown to have protective functions against stress-induced tRNA cleavage.

NOP2/Sun RNA methyltransferase family member 2 (**NSUN2**, a homologue of yeast Trm4), was reported to also methylate 5-cytosine in mRNAs and in various non-coding RNAs.

2’-O-methylnucleosides



2’-O-methylation (2’OMe) is a common RNA modification that resides on the 2’hydroxyl ribose moiety of all ribonucleosides 2’-OMe has been found in all major classes of eukaryotic RNA.
2’-OMe can **inhibit A to I RNA** editing in vitro.
Small nucleolar RNAs (**snoRNAs**) are known to guide 2’OMe on eukaryotic rRNA and may also target other RNA species such as mRNA.

However, the precise sites of 2′OMe in eukaryotic mRNA or its function are currently unclear.

Suggested readings:

https://www.nature.com/articles/nrm.2016.132

https://www.frontiersin.org/articles/10.3389/fcell.2021.628415/full

https://www.nature.com/articles/nature15377

https://genomebiology.biomedcentral.com/articles/10.1186/s13059-017-1347-3

Further readings at:
Collection the Epitranscriptome by Nature:
http://www.nature.com/collections/nrdmjjhwky