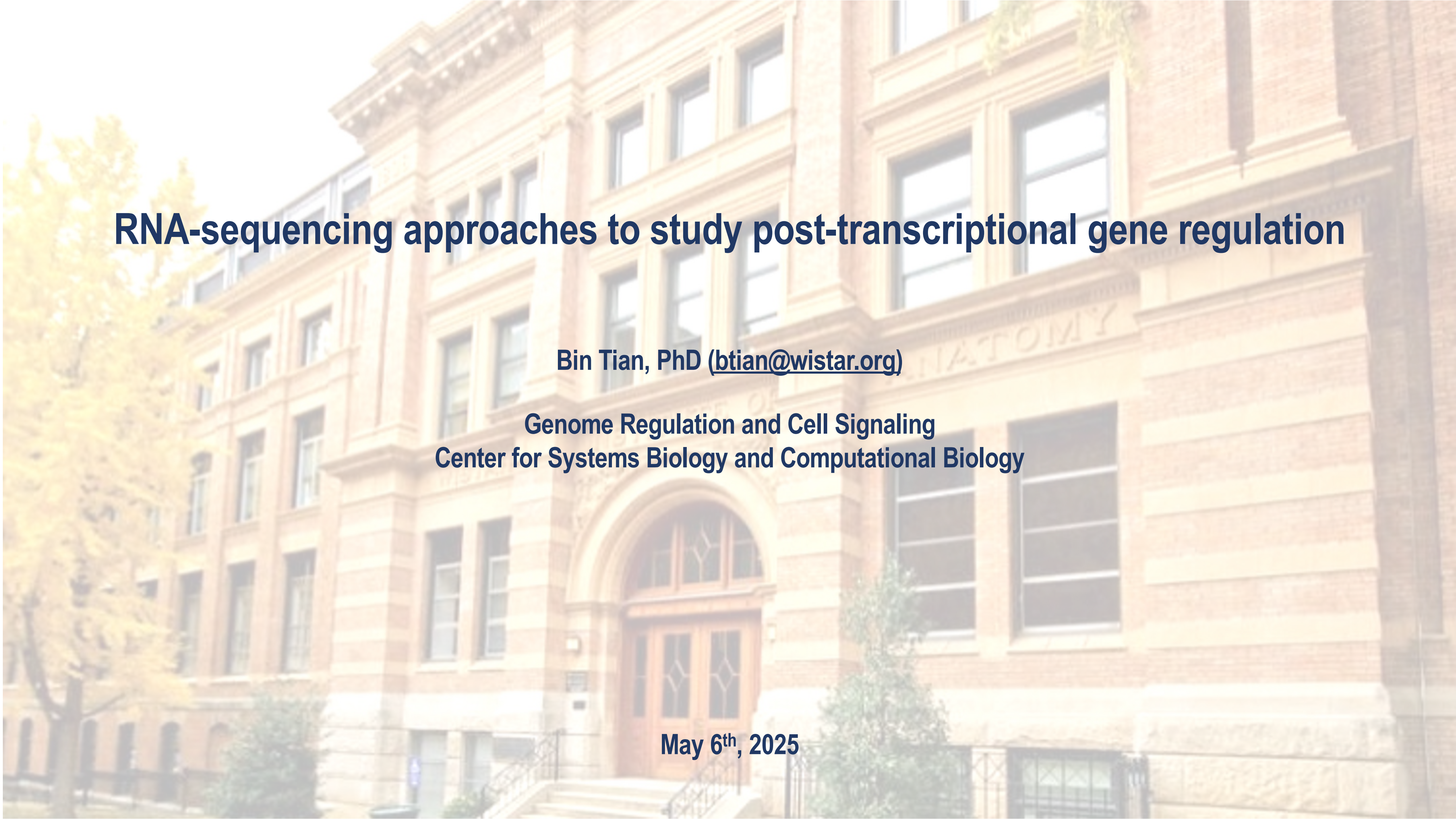


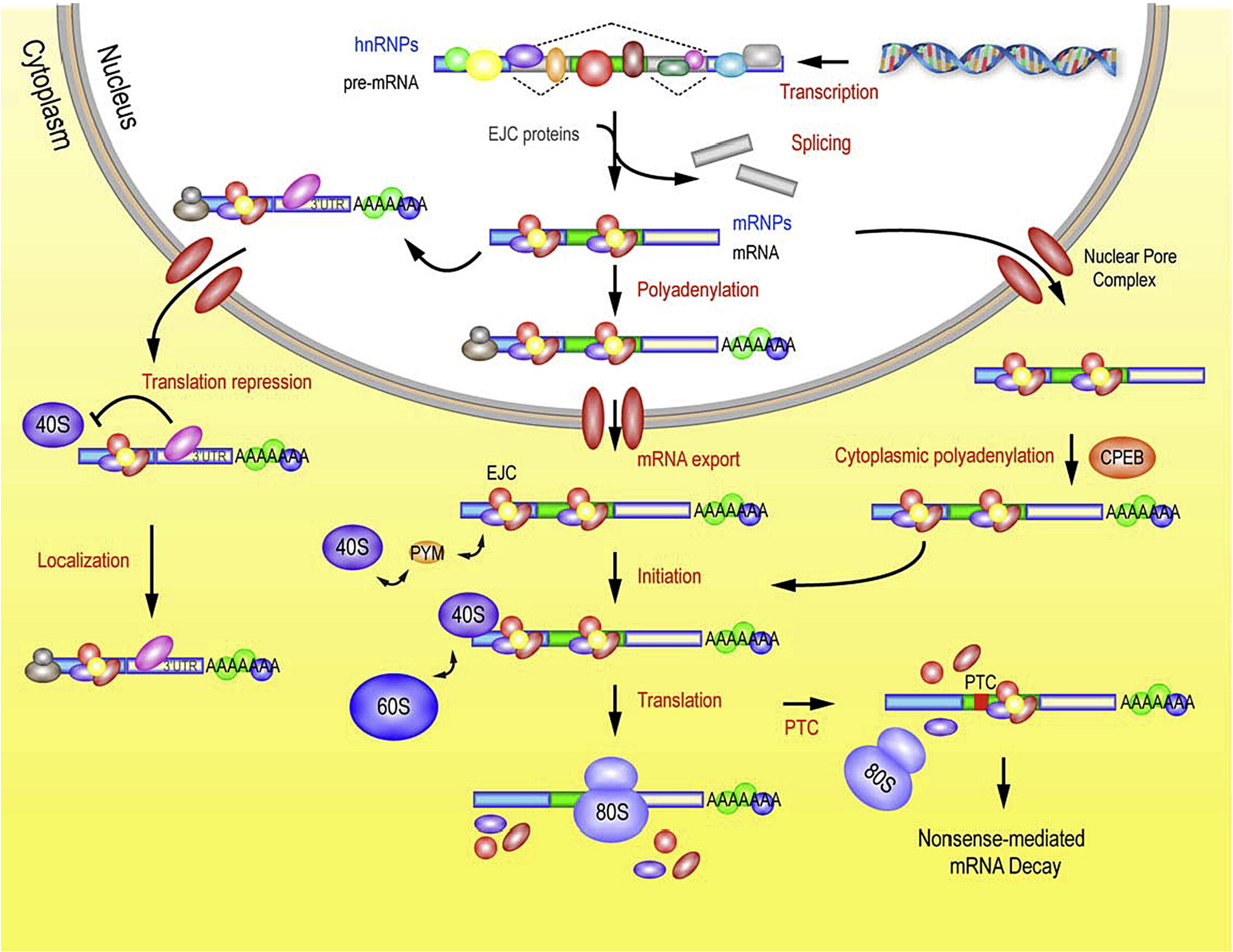
RNA-sequencing approaches to study post-transcriptional gene regulation

Bin Tian, PhD (btian@wistar.org)

Genome Regulation and Cell Signaling
Center for Systems Biology and Computational Biology

May 6th, 2025

Post-transcriptional Gene Regulation Pathways in Eukaryotes



Lecture outline:

I. mRNA processing

II. mRNA decay

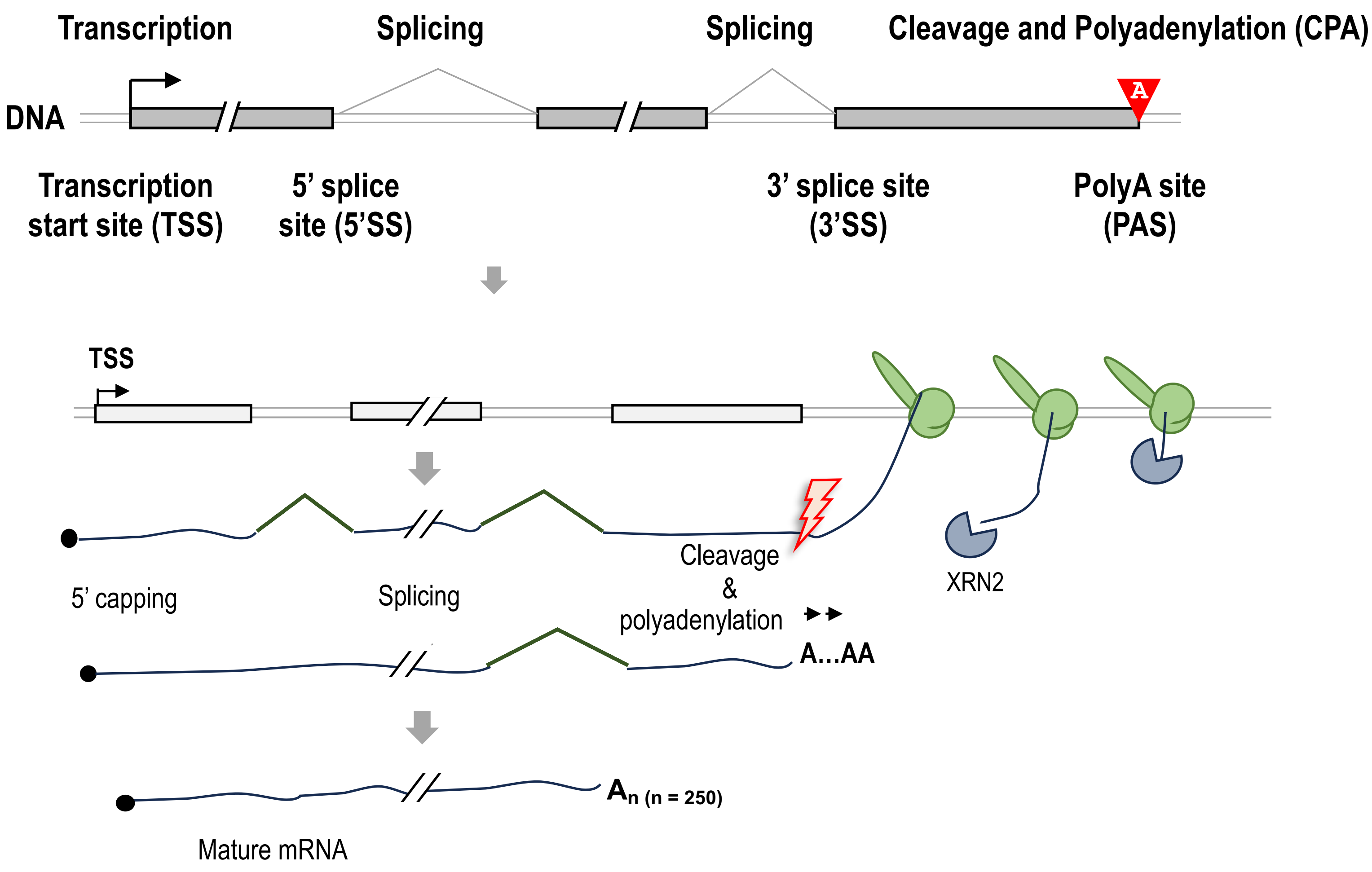
III. mRNA translation.

IV. RNA modifications.

V. RNA-binding proteins (RBPs).

Glisovic et al. *FEBS Letters* 2008.

Pre-mRNA processing



Transcription **Splicing** **Splicing** **Cleavage and Polyadenylation (CPA)**

DNA

Transcription start site (TSS)

5’ splice site (5’SS)

3’ splice site (3’SS)

PolyA site (PAS)

TSS

5’ capping

Splicing

Cleavage & polyadenylation

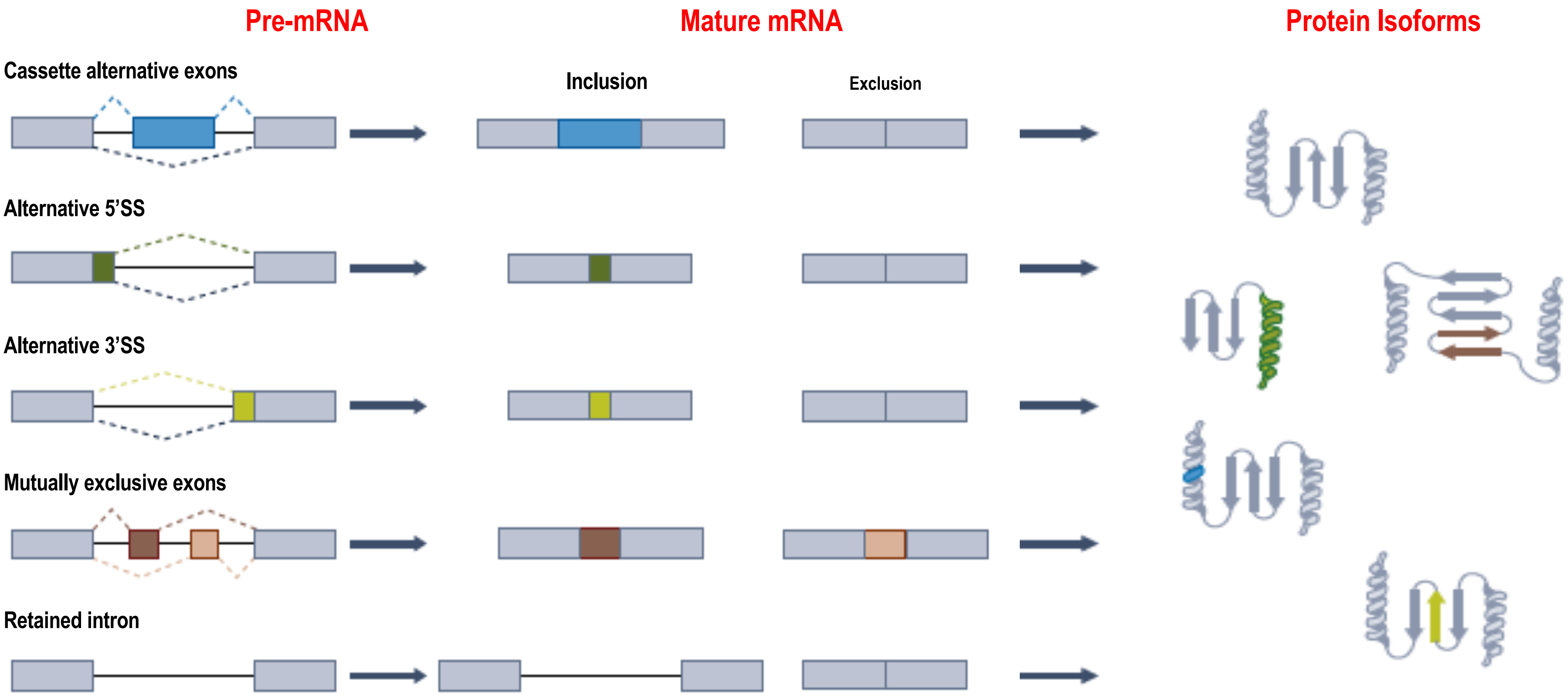
XRN2

A...AA

A_n (n = 250)

Mature mRNA

Alternative splicing



Pre-mRNA

Mature mRNA

Protein Isoforms

Cassette alternative exons

Inclusion

Exclusion

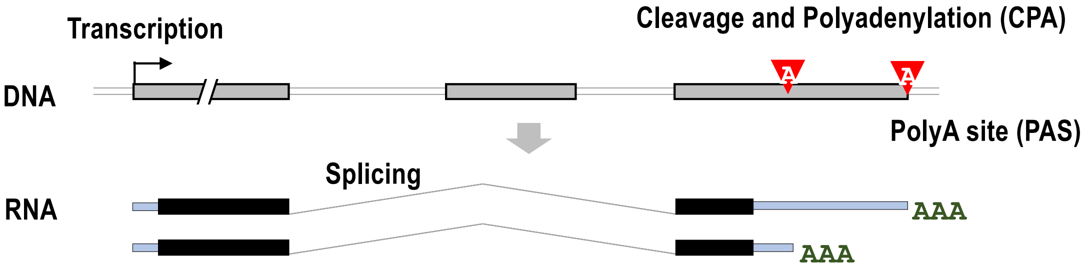
Alternative 5’SS

Alternative 3’SS

Mutually exclusive exons

Retained intron

Alternative polyadenylation



Transcription

Cleavage and Polyadenylation (CPA)

DNA

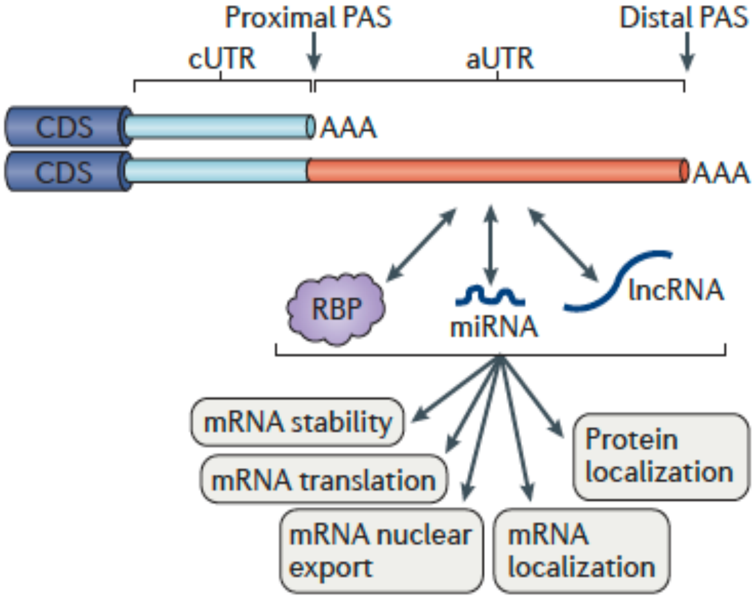
PolyA site (PAS)

Splicing

RNA

AAA

AAA



Proximal PAS

Distal PAS

cUTR

aUTR

CDS

AAA

CDS

AAA

RBP

miRNA

lncRNA

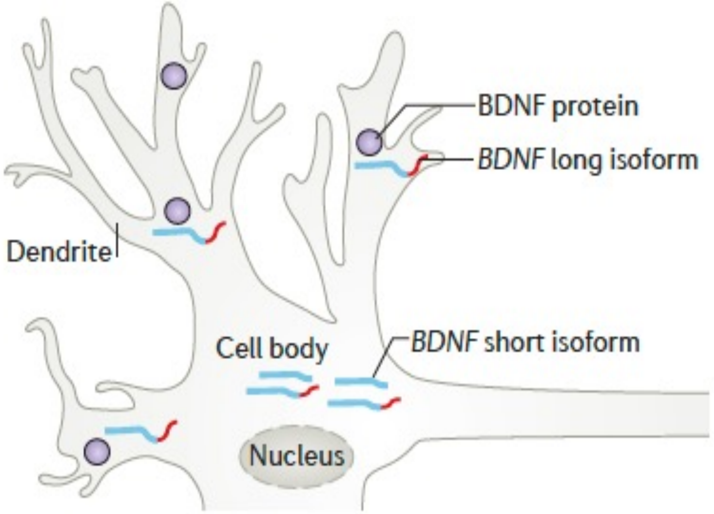
mRNA stability

mRNA translation

mRNA nuclear export

mRNA localization

Protein localization



BDNF protein

BDNF long isoform

Dendrite

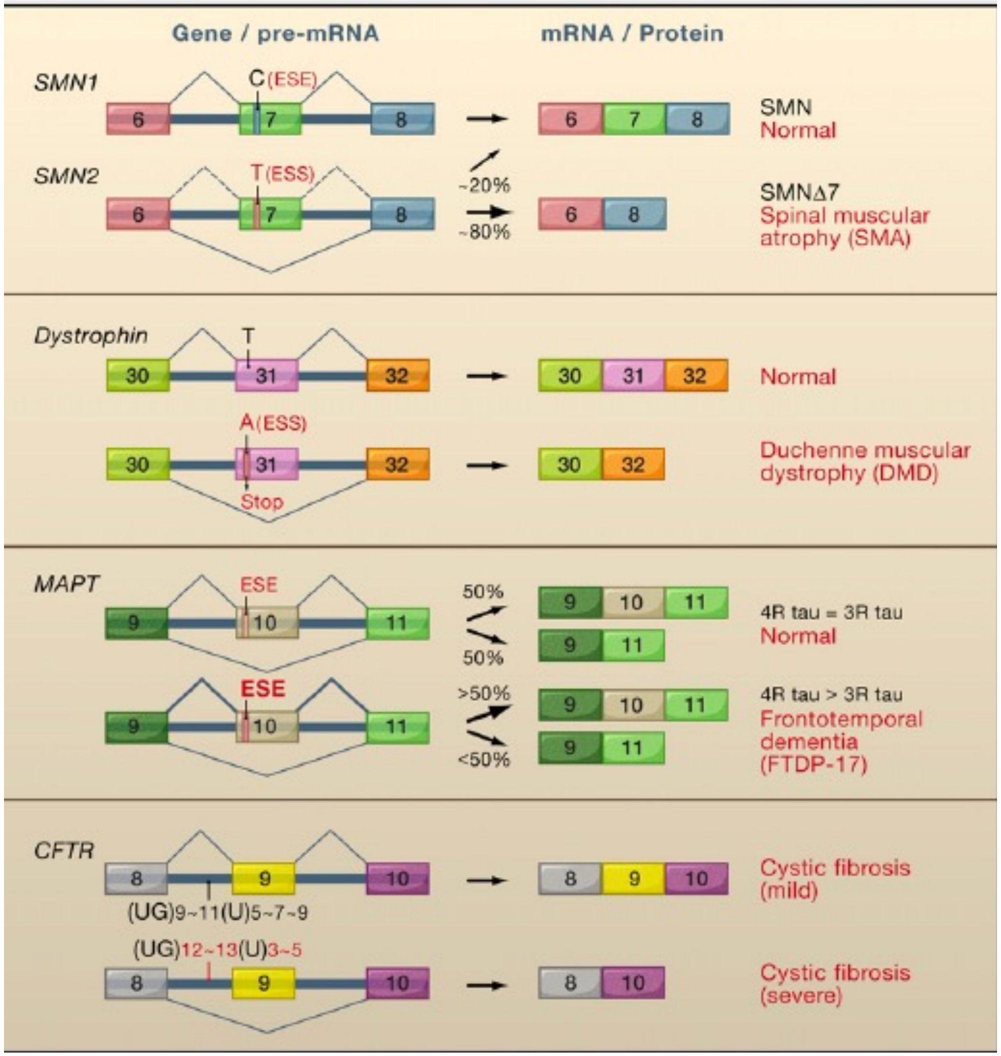
Cell body

BDNF short isoform

Nucleus

Tian & Manley . *Nat Rev Mol Cell Biol*. 2017.

Lost-of-function germline mutations



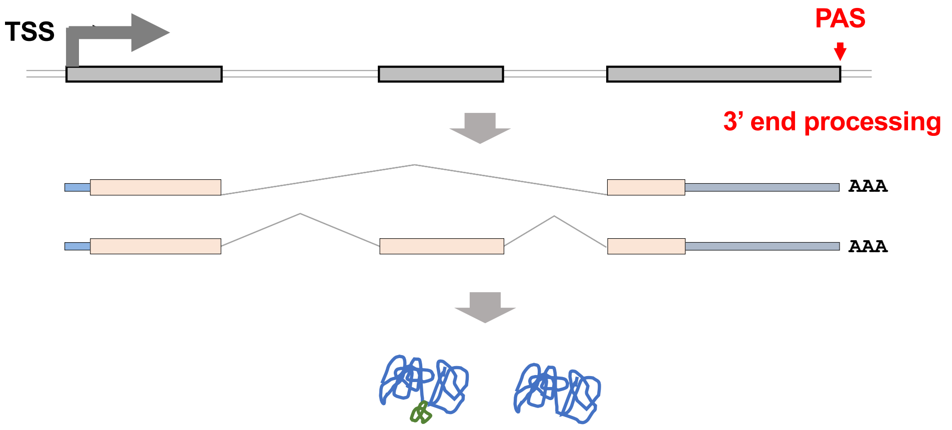
ESE: exonic splicing enhancer
ESS: exonic splicing silencer

Up to 25% of synonymous (in terms of amino acid coding) substitutions can disrupt normal splicing, as can nonsynonomous and termination codons.

For some genes, up to 50% of point mutations within exons affect splicing, and it has been hypothesized that more than half of known disease-causing mutations disrupt splicing.

Cooper et al. *Cell*. 2009.

Germline mutations in the 3’ End Processing region



TSS: transcription start site

PAS: polyadenylation site

▭ Coding sequence

▬ Untranslated region

Loss-of-function (weakened PAS): Thalassemia: HBA (Higgs et al. Nature 1983); HBB (Orkin et al. EMBO J 1985).

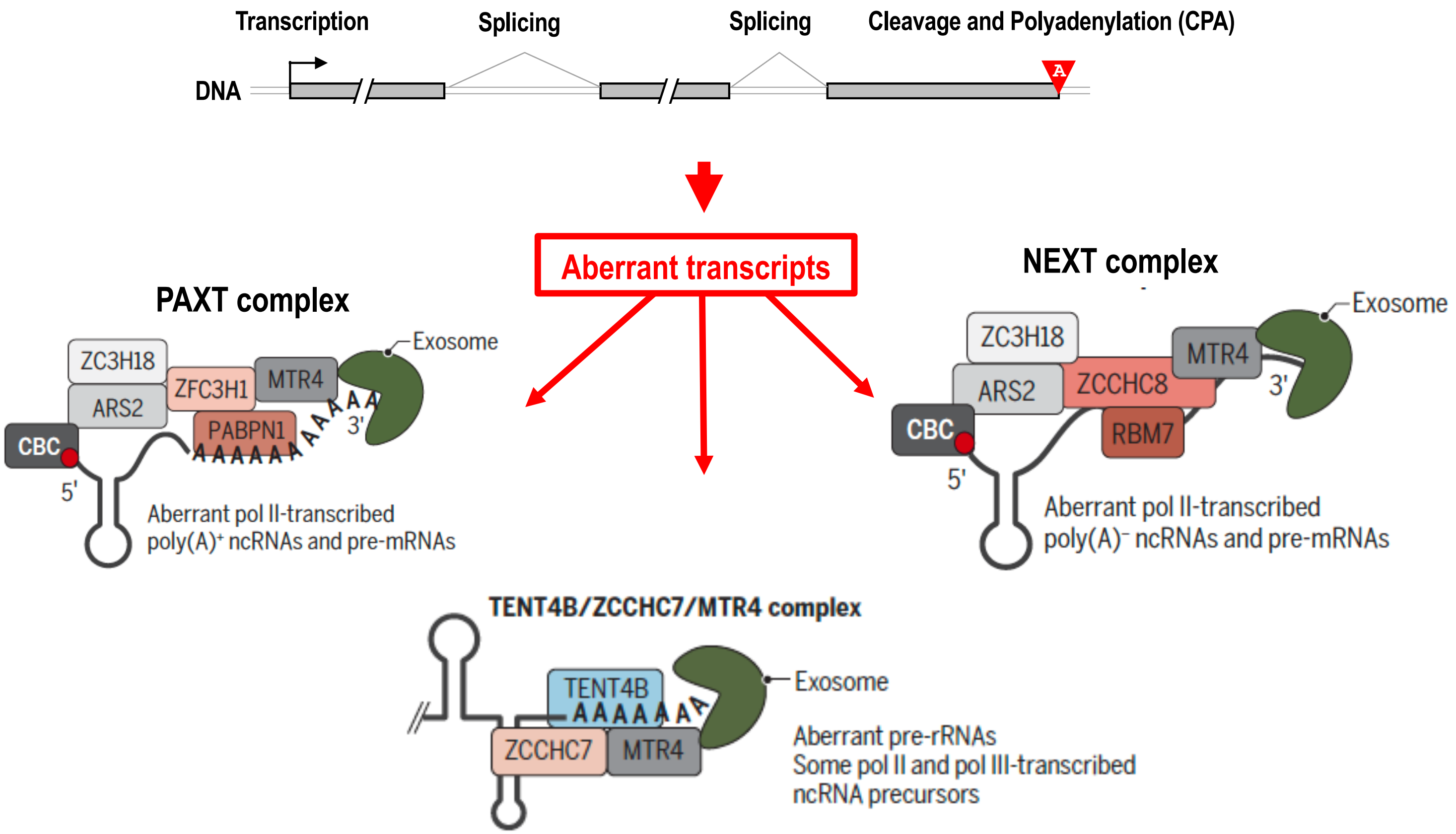
Glanzmann Thrombasthenia, Esophageal squamous cell carcinoma, Lynch syndrome, Microphthalmia, Osteogenesis imperfecta, Neonatal diabetes, IPEX syndrome, Severe combined immunodeficiency, Wiskott-Aldrich syndrome, Fabry disease, Metachromatic leukodystrophy, Zellweger syndrome, ...

Gain-of-function (strengthened PAS): Thrombophila: F2 (Poort et al. Blood 1996; Gehring et al. Nat Genet 2021).

Hypertension, Systemic lupus erythematosus, Facioscapulohumeral muscular dystrophy, Amyotrophic lateral sclerosis, Huntington’s disease, ...

Gruber & Zavolan. *Nat Rev Genet.* 2019.

Nuclear mRNA Surveillance Mechanisms

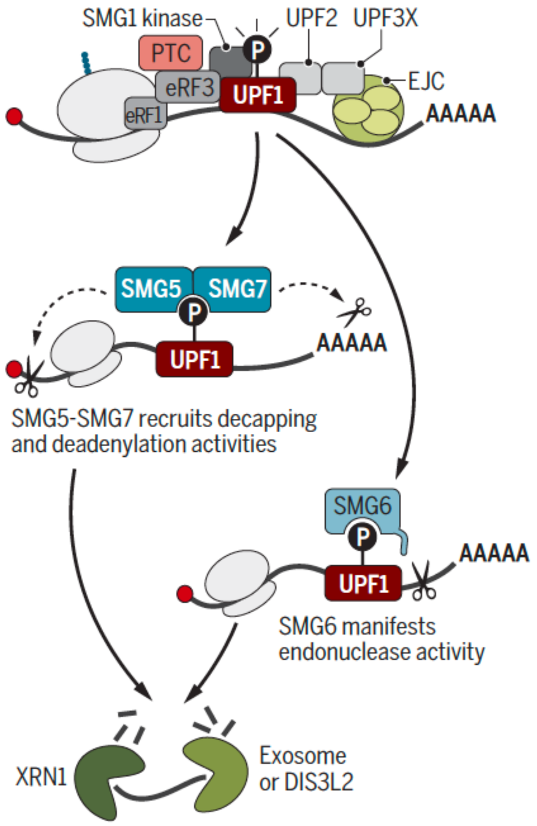


Oligo-adenylation is a signal for nuclear exosome-mediated RNA degradation

Wolin & Maquat. *Science*. 2019.

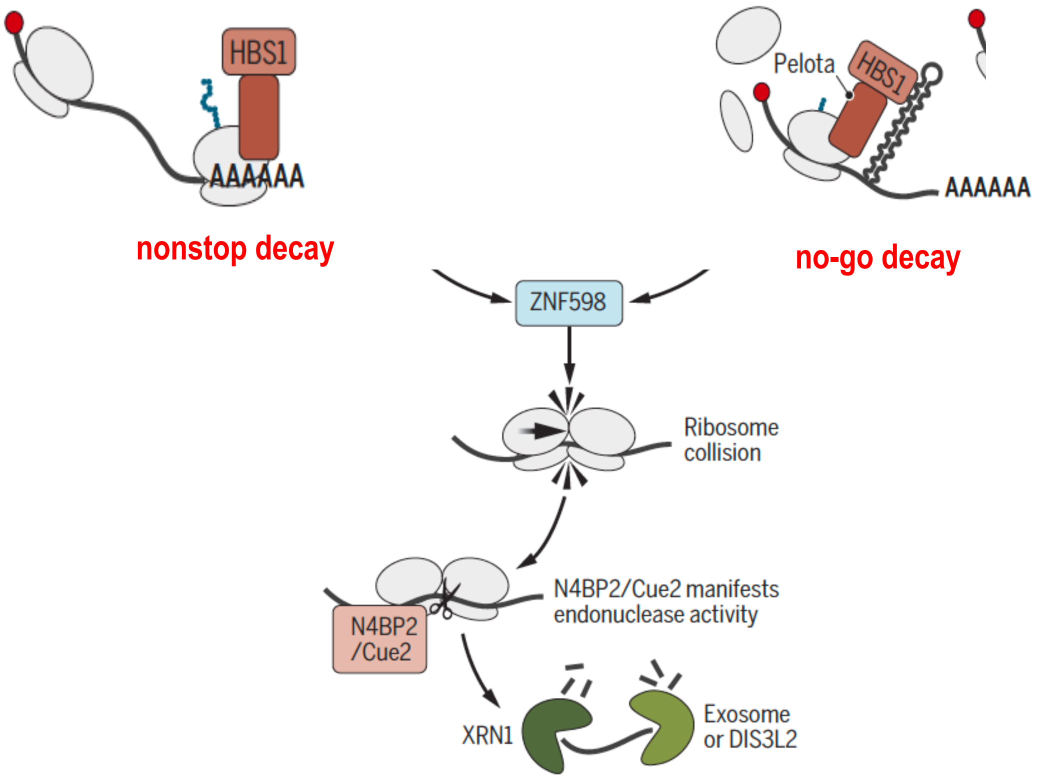
Cytoplasmic mRNA Surveillance Mechanisms

Non-sense mediated decay (NMD)



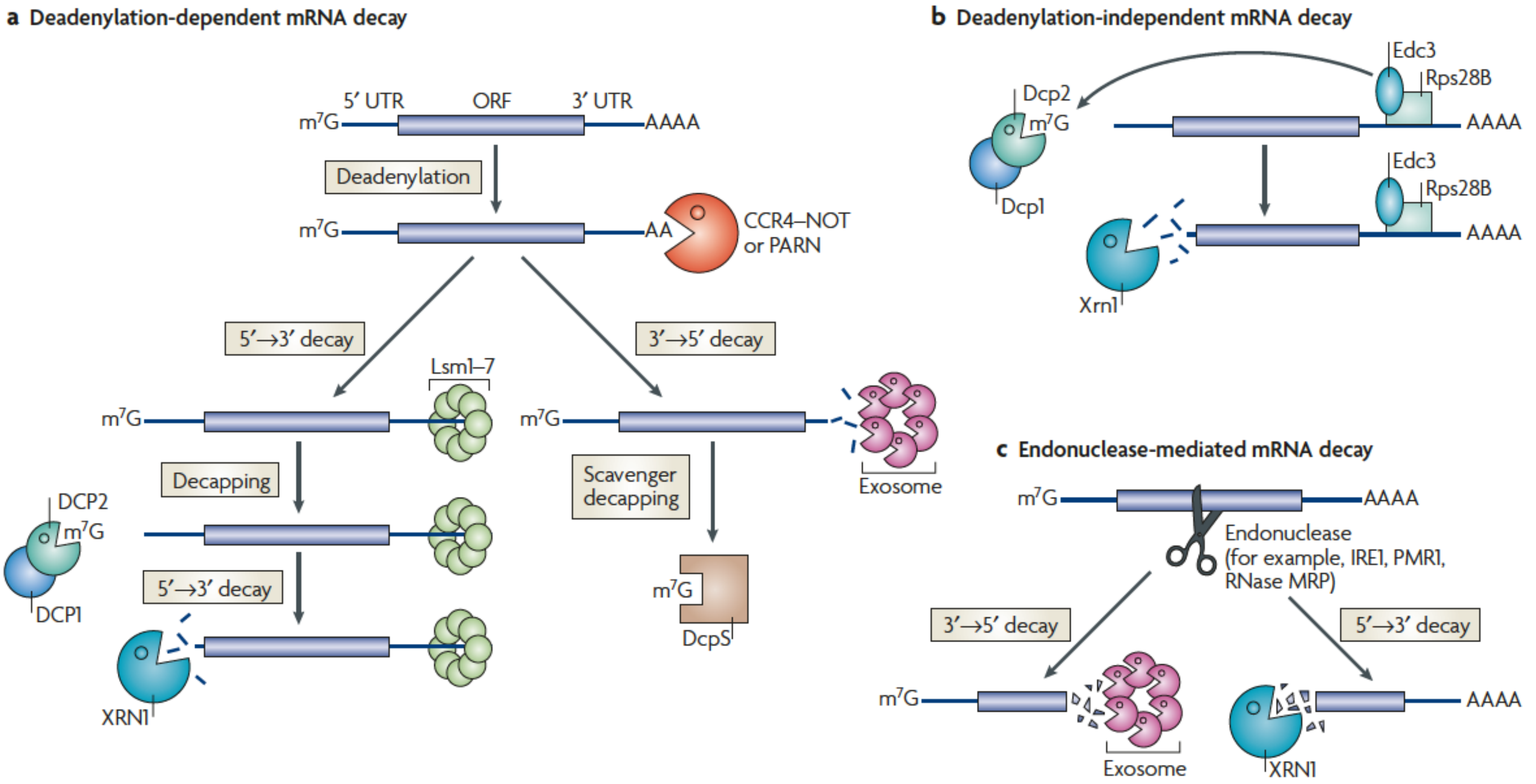
Transcripts without a stop codon

Ribosome stalling due to RNA structure, etc.



Wolin & Maquat. *Science*. 2019.

mRNA decay



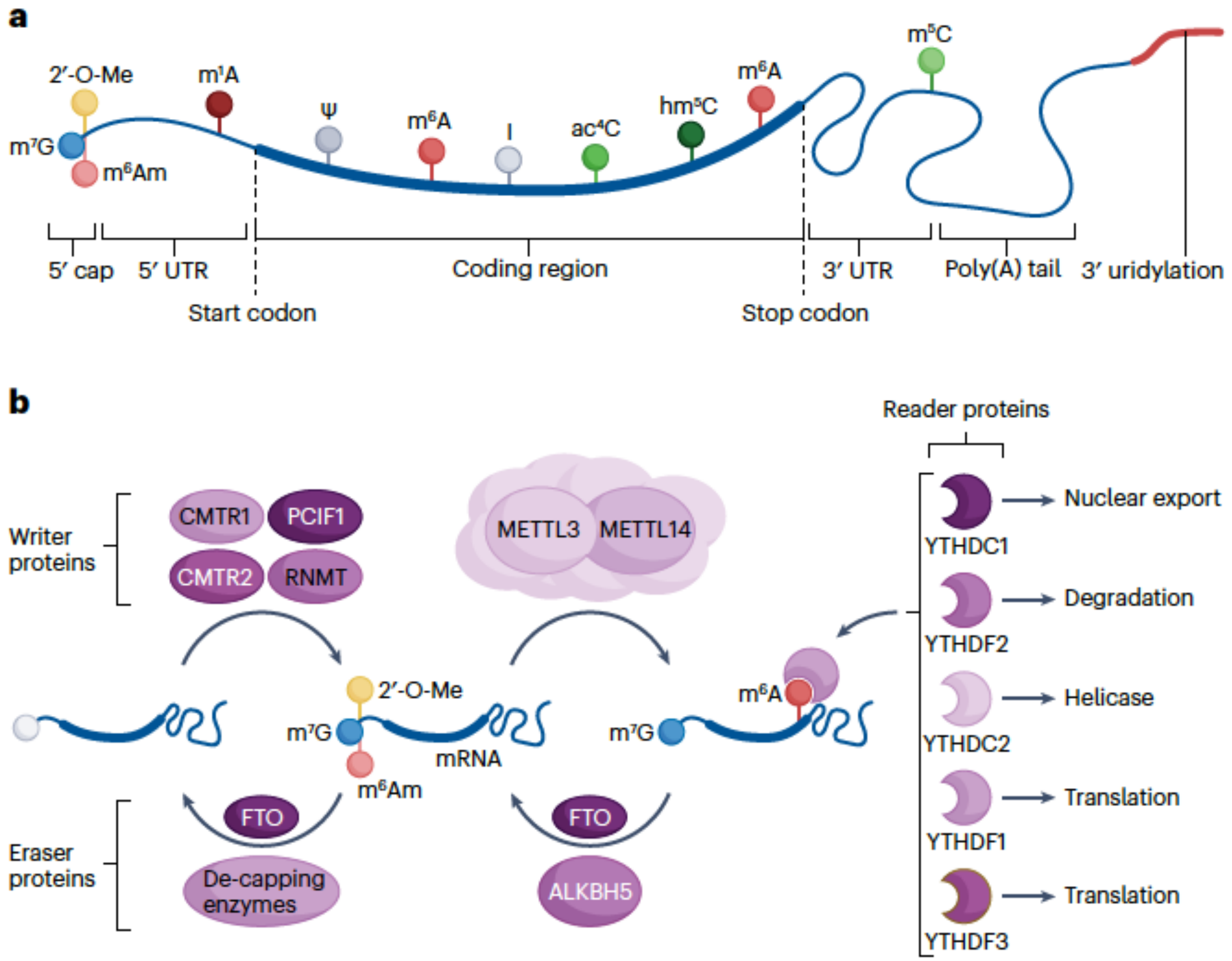
a Deadenylation-dependent mRNA decay

b Deadenylation-independent mRNA decay

c Endonuclease-mediated mRNA decay

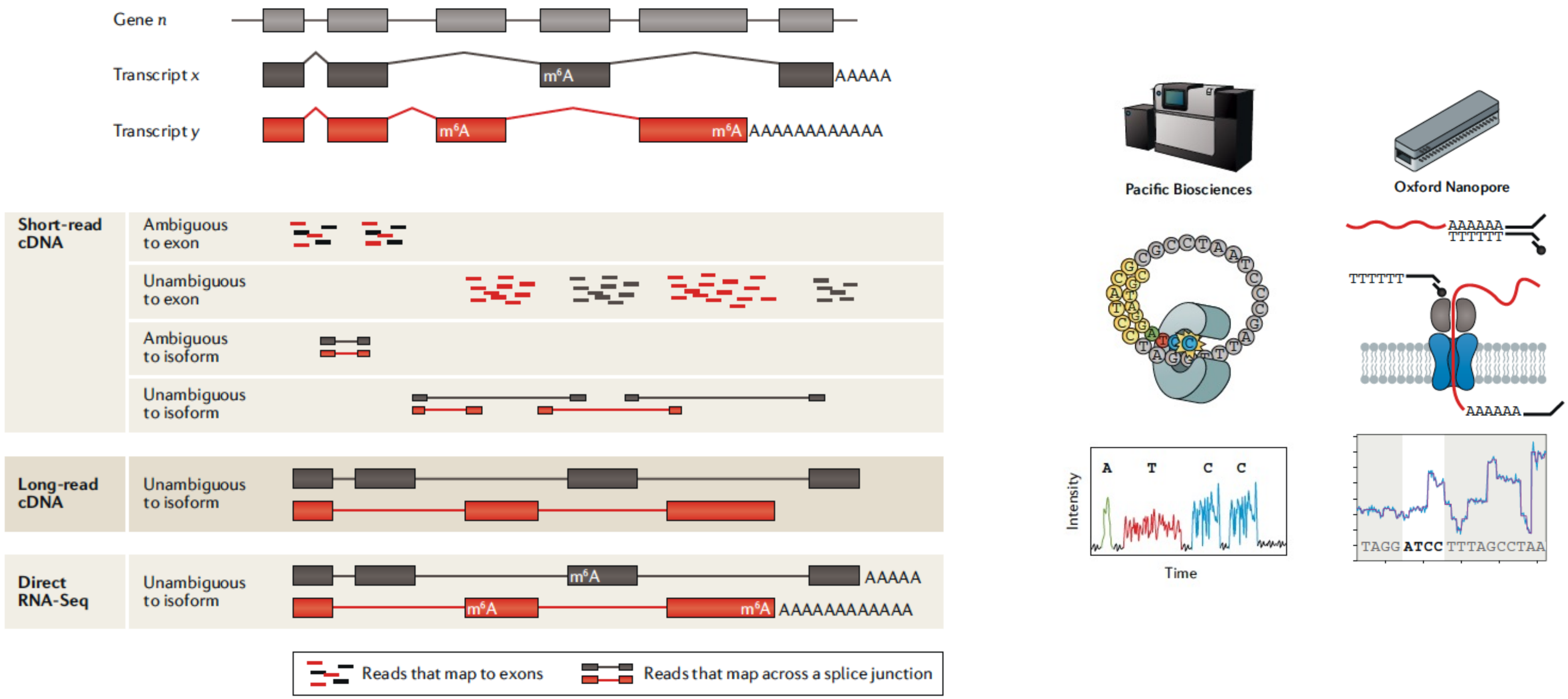
Garneau, Wilusz, Wilusz. *Nat Rev Mol Cell Biol.* 2007.

mRNA modification adds another layer of gene regulation



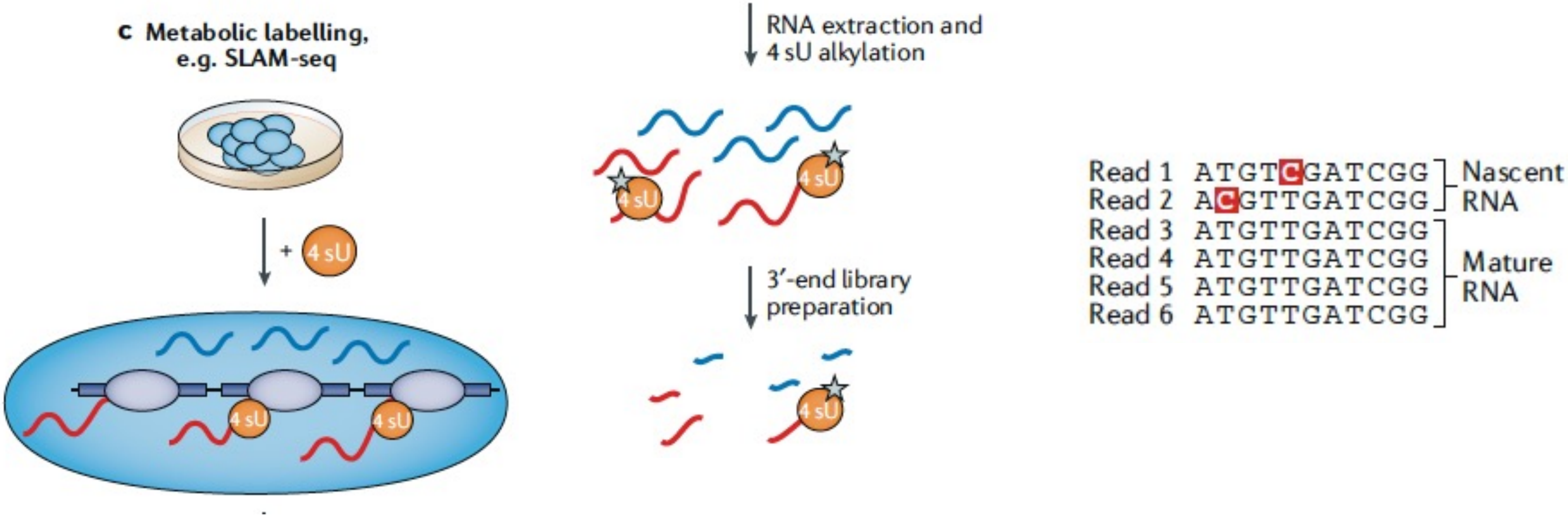
Delaunay, Helm, Frye. *Nat Rev Genet.* 2024.

Long read sequencing identifies RNA isoforms and modifications



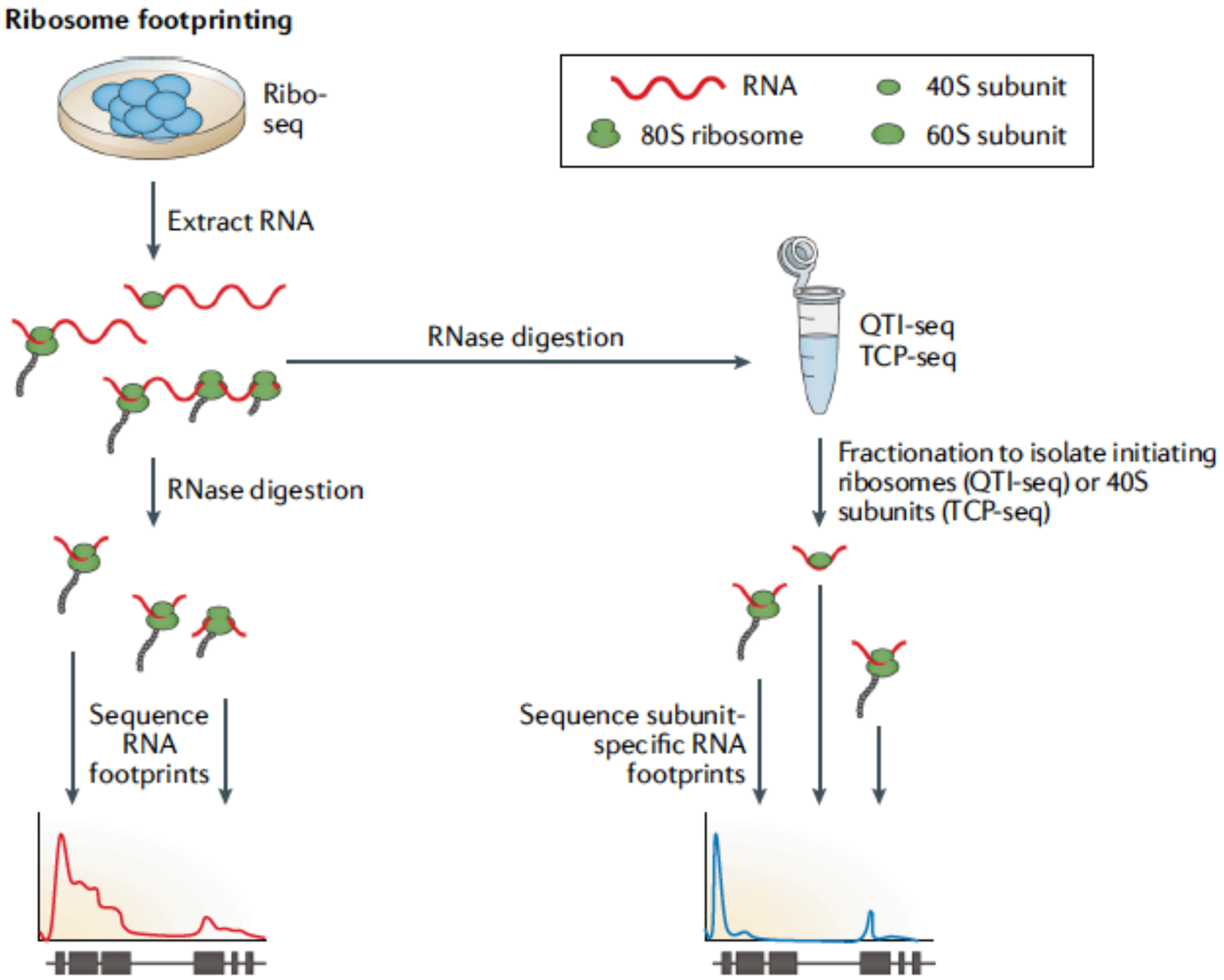
Stark, Grzelak, Hadfield. *Nat Rev Genet.* 2019.

SLAM-seq



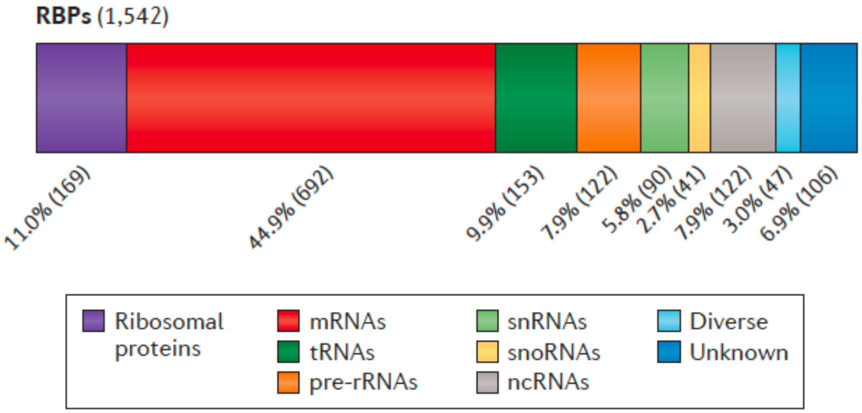
Stark, Grzelak, Hadfield. *Nat Rev Genet.* 2019.

Polysome profiling vs. Ribo-seq

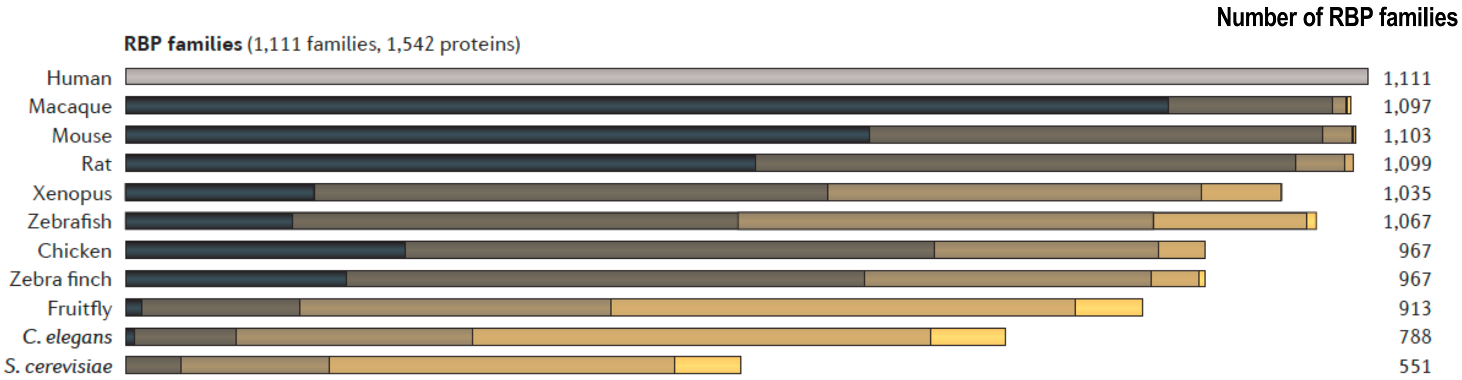


Stark, Grzelak, Hadfield. *Nat Rev Genet.* 2019.

A Large Number of RBPs in the Human Genome



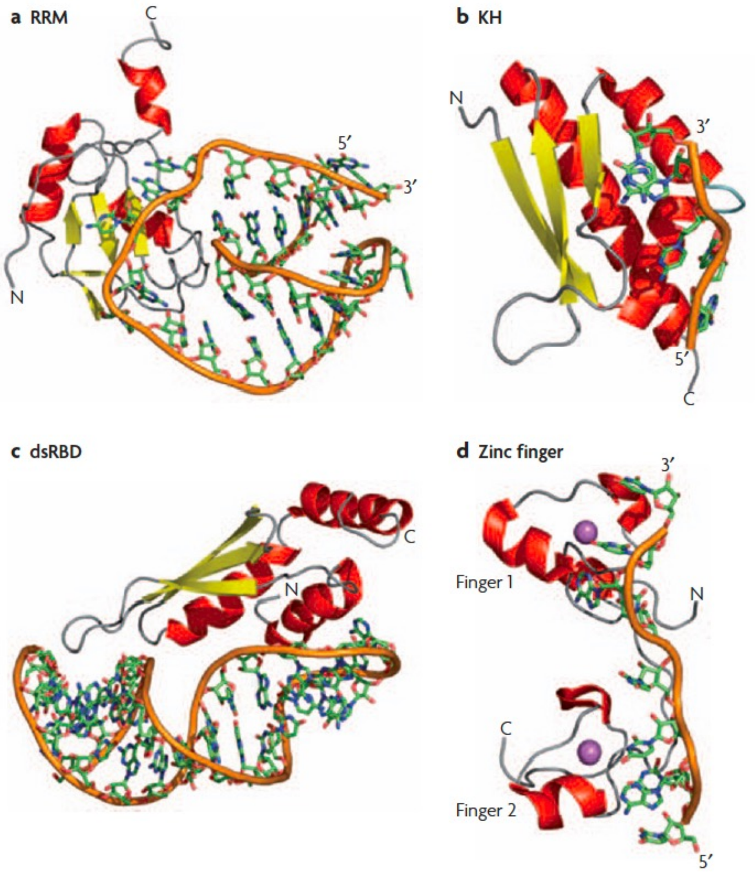
- >1,500 RBPs in the human genome.
- RBPs target all types of RNAs.
- RBPs can be grouped into families based on sequence homology.



Gerstberger et al. *Nat Rev Genet*. 2014.

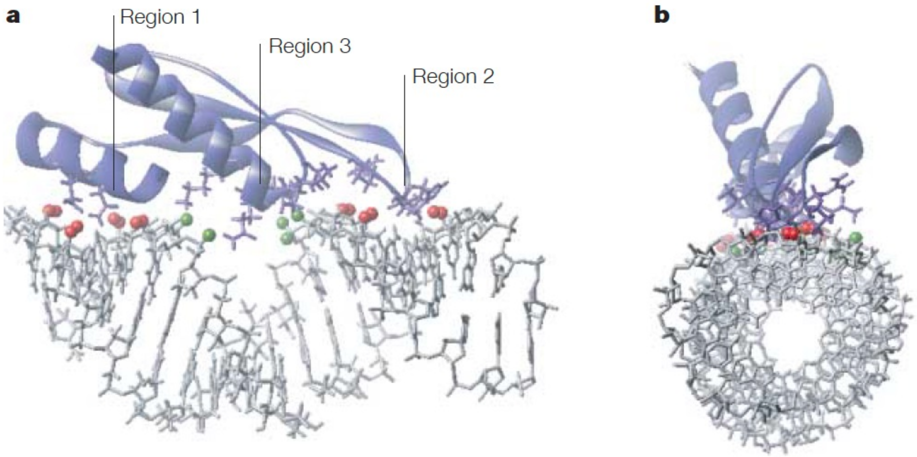
RNA-binding domains (RBDs) are responsible for RNA interactions

Domain	Topology	RNA-recognition surface
RRM	αβ	Surface of β-sheet
KH (type I and type II)	αβ	Hydrophobic cleft formed by variable loop between β2, β3 and GXXG loop. Type II: same as type I, except variable loop is between α2 and β2
dsRBD	αβ	Helix α1, N-terminal portion of helix α2, and loop between β1 and β2
ZnF-CCHH	αβ	Primarily residues in α-helices
ZnF-CCCH	Little regular secondary structure	Aromatic side chains form hydrophobic binding pockets for bases that make direct hydrogen bonds to protein backbone



Lunde et al. *Nat Rev Mol Cell Biol.* 2007.

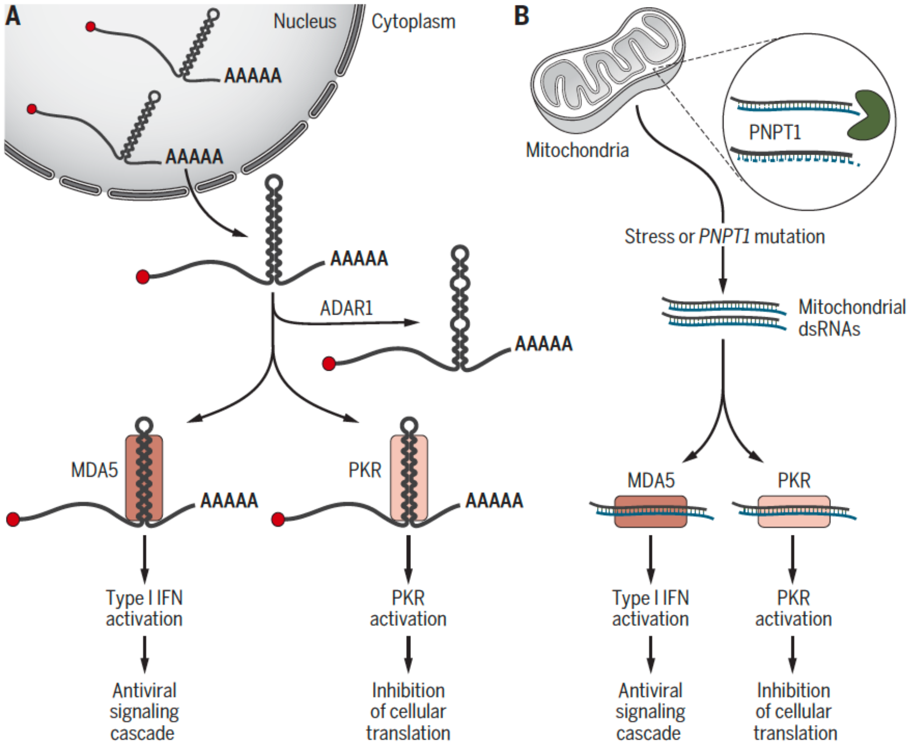
Double-stranded RNA-binding domain (dsRBD) interacts with RNA structures



Tian et al. *Nat Rev Mol Cell Biol.* 2004.

- dsRBDs bind the RNA duplex structure without sequence specificity.
- RNA interactions are through phosphate groups, rather than bases.
- dsRBPs are involved in dsRNA biology, including processing of dsRNAs, surveillance of dsRNAs, and miRNA biogenesis and functions.

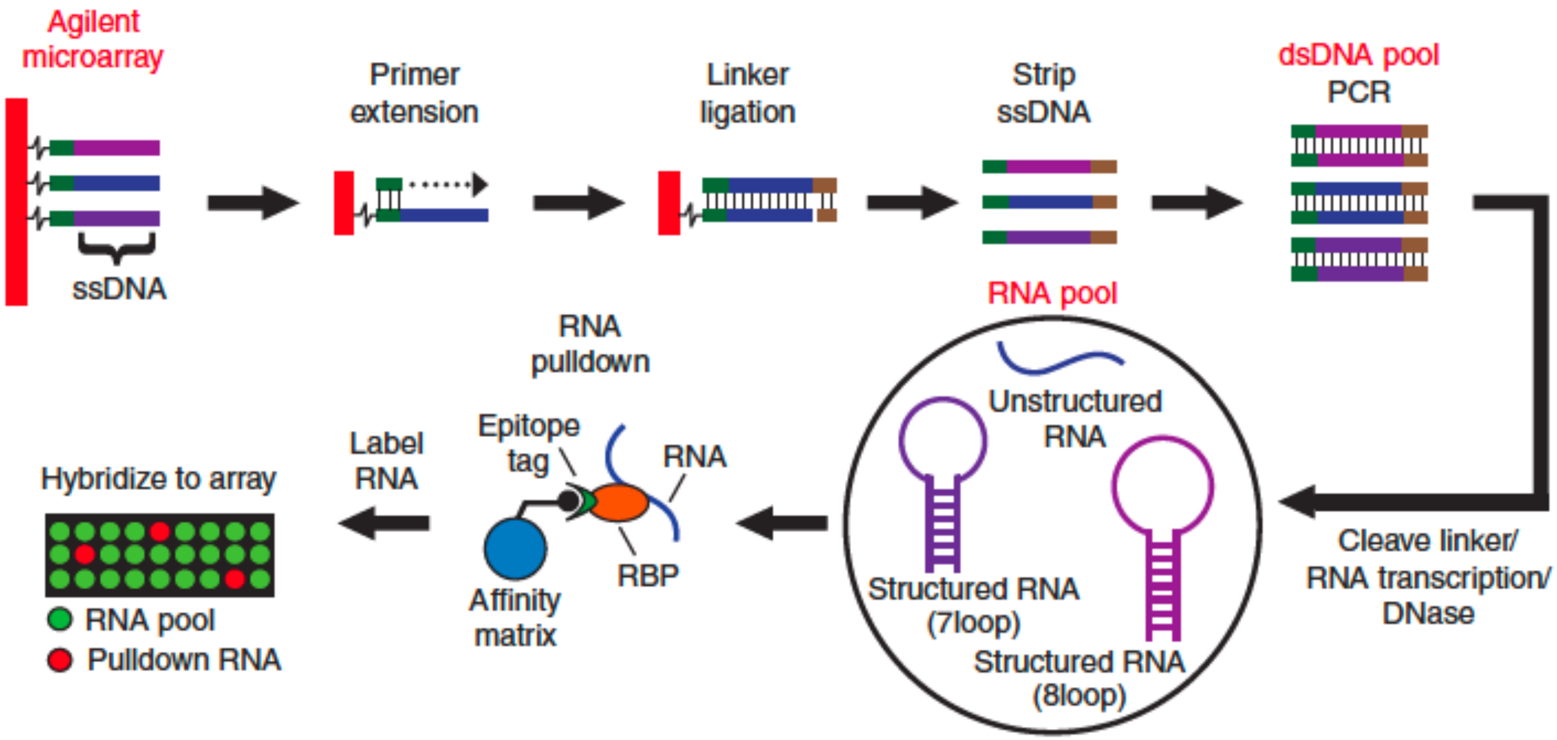
Surveillance of dsRNAs



Wolin & Maquat. *Science*. 2019.

High throughput RIP assays to study RBP:RNA interactions

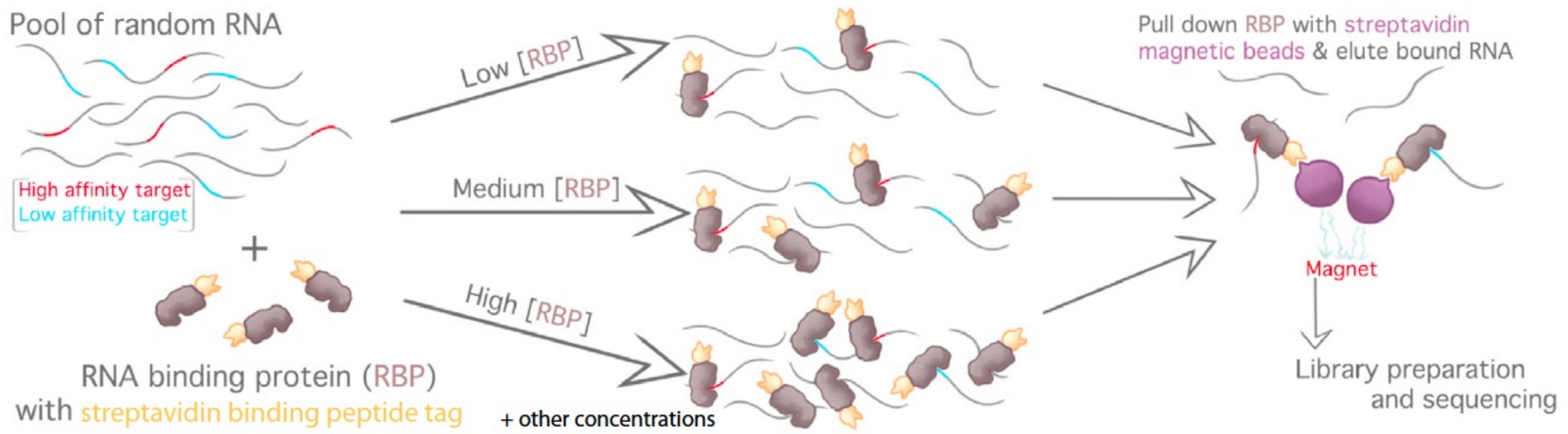
RNAcompete



Ray…Hughes. *Nat Biotech.* 2009.

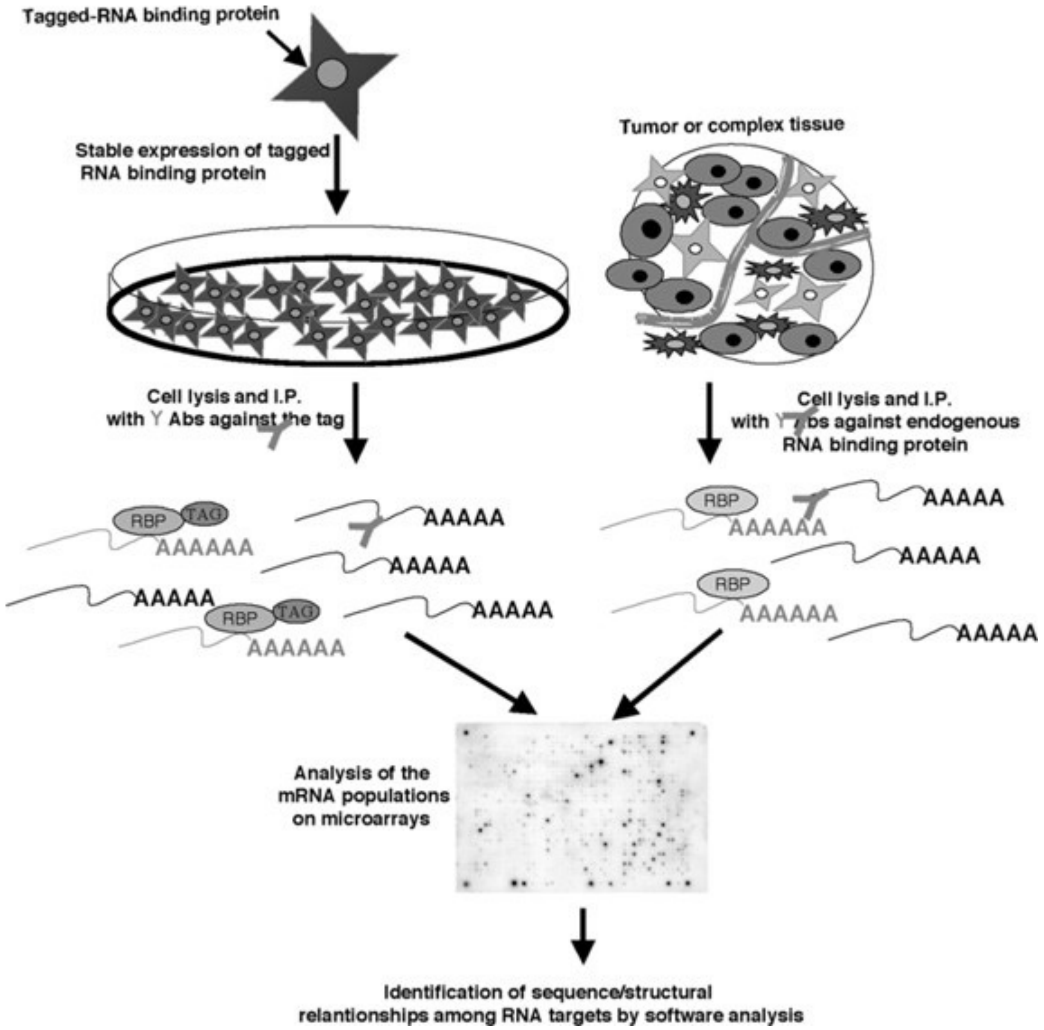
High throughput RIP assays to study RBP:RNA interactions

RNA Bind-n-Seq



Lambert…Burge. *Mol Cell*. 2014.

RIP assay to study RBP:RNA interactions

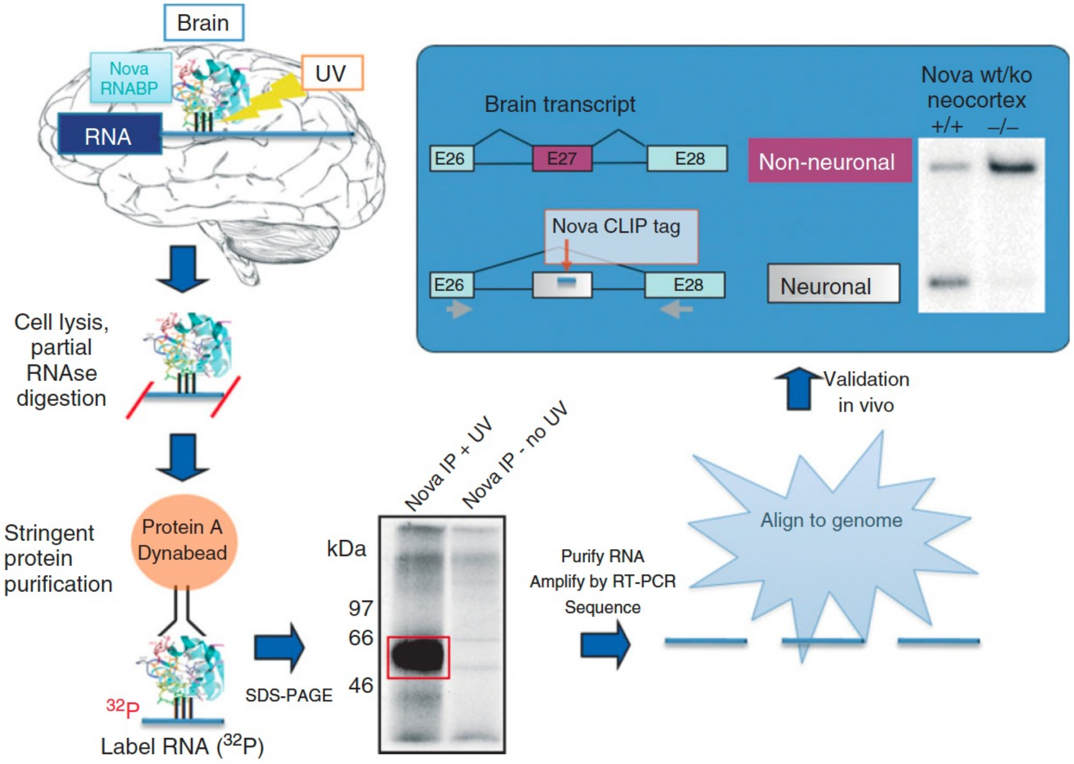


RIP: RNA immunoprecipitation

- Use of an antibody against a specific RBP.
- Immunoprecipitated RNA is subject to RT-qPCR, microarray, or sequencing.
- Crosslinking by formaldehyde may help stabilize RNA:RBP interactions.
- Some new interactions may take place during sample processing (Mili & Steitz, RNA, 2004).

Penalva...Keene. *Methods Mol Biol.* 2004.

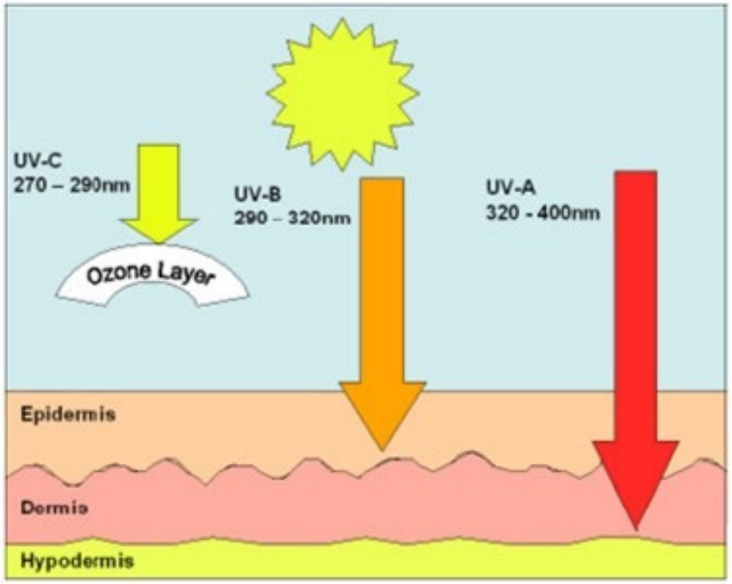
CLIP assay to study RBP:RNA interactions



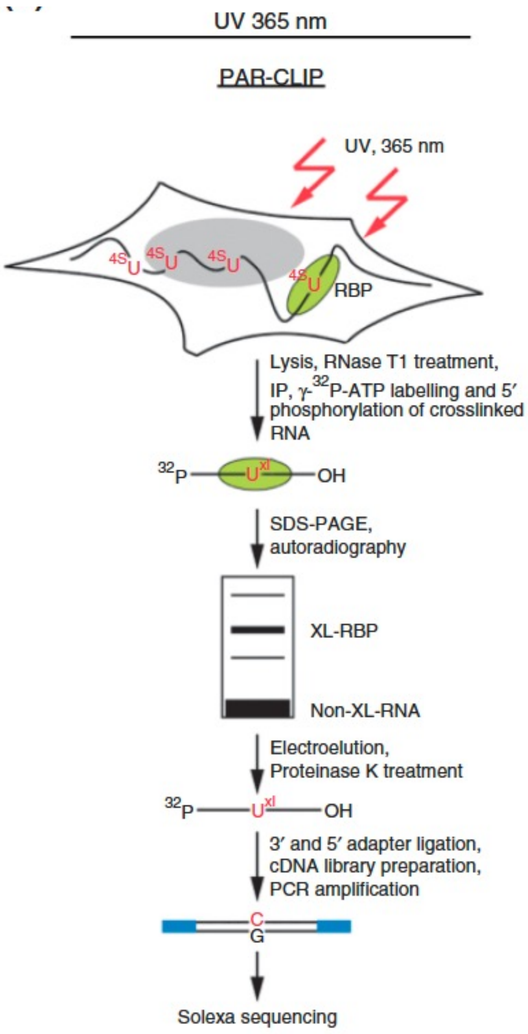
Darnell. *WIREs RNA*. 2010.

CLIP: Crosslinking immunoprecipitation

- **HITS-CLIP (aka CLIP-seq) : CLIP coupled with high throughput sequencing.**
- **Irradiation by UV 254 nm.**
- **iCLIP, eCLIP, etc.**
- **Mutations/deletions introduced during RNA reverse transcription indicate RBP binding site.**



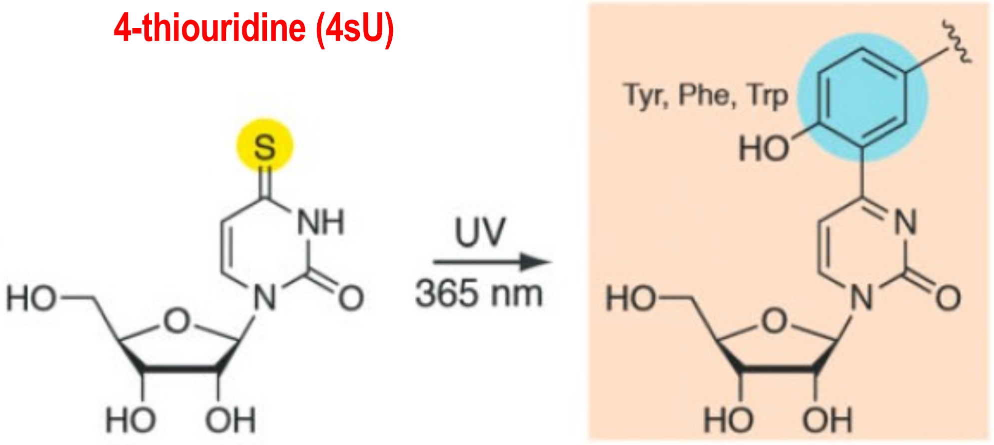
PAR-CLIP assay to study RBP:RNA interactions



Photoactivatable-ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP)

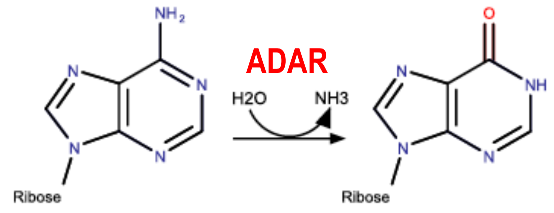
- **Use of photoactivatable ribonucleoside analogs.**
- **Irradiation by UV at 365 nm.**
- **T->C mutation (4sU) used for binding site identification.**

4-thiouridine (4sU)

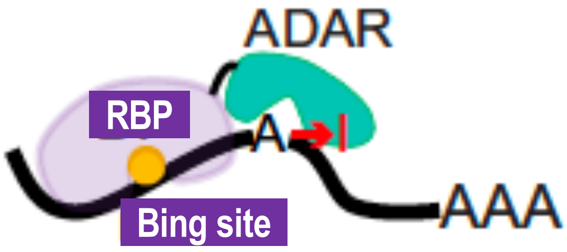


Ascano…Tuschl. *WIREs RNA* 2012.

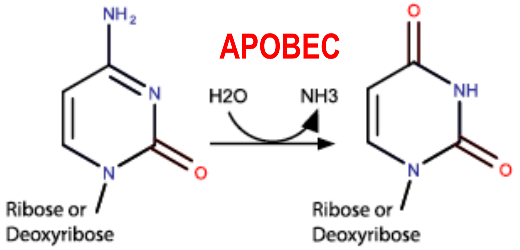
Base-editing tools to study RBP:RNA interactions



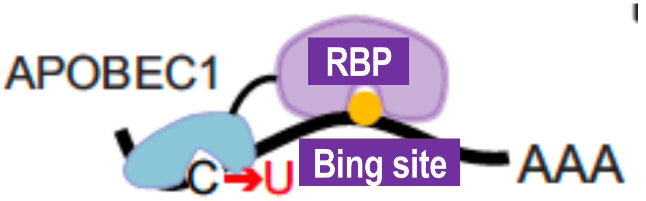
Adenosine Deamination to Inosine by ADAR



TRIBE: McMahon…Rosbash. (2016). Cell.

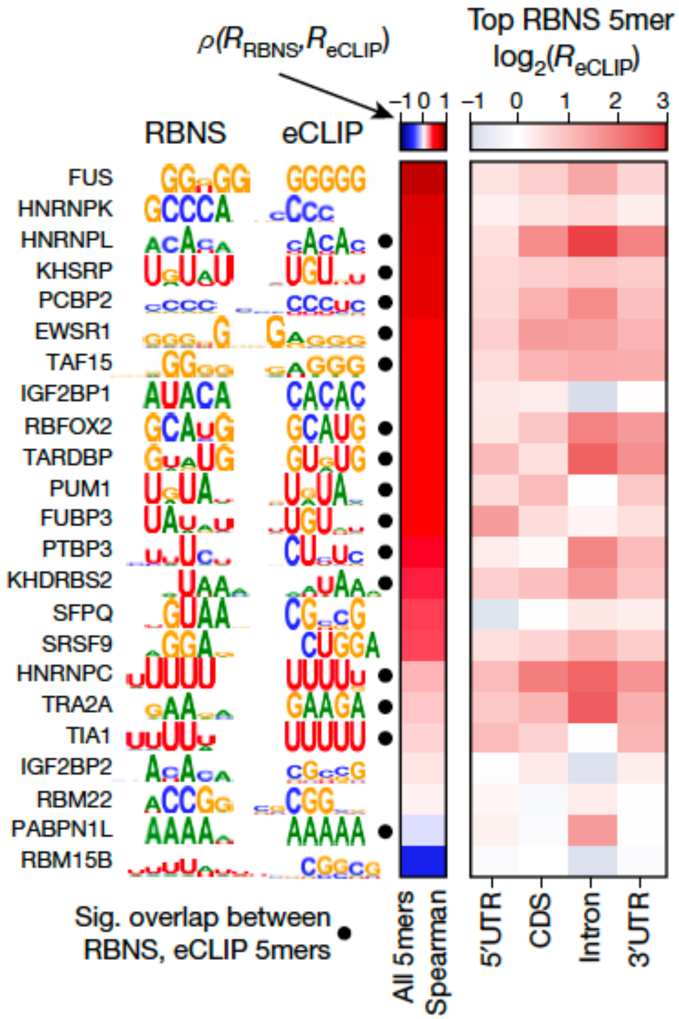


Cytosine Deamination to Uridine by APOBEC



DART-seq: Meyer (2019). Nat. Methods.
STAMP: Brannan … Yeo (2021). Nat. Methods.

RBPs have diverse sequence specificities and diverse roles



- **Sequence logo captures binding specificity.**
- **The height of a position indicates its information content or nucleotide bias.**
- **The height of each nucleotide at a given position indicates its frequency.**
- **RBP binding sequences are often short (4-6 nucleotides)**
- **Some RBPs have different binding specificities in vivo compared to in vitro**

The Encyclopedia of RNA Elements (ENCORE) project.

Van Nostrand et al. *Nature* 2020.

Questions?