Processing to export and localize

from previous lectures.....

Alternative splicing control:

ESE, ESS, ISE, ISS  cis-elements
SR/hnRNP ratio

Tissue-specific splicing factors (Nova, ESPR, nPTB)

Promoters/Enhancers can confer higher/lower speed to RNA Pol II and speed is proportional to higher exclusion of “wseak” exons.

Promoters/enhancers can “load” splicing factors to the RNA PolII elongation complex

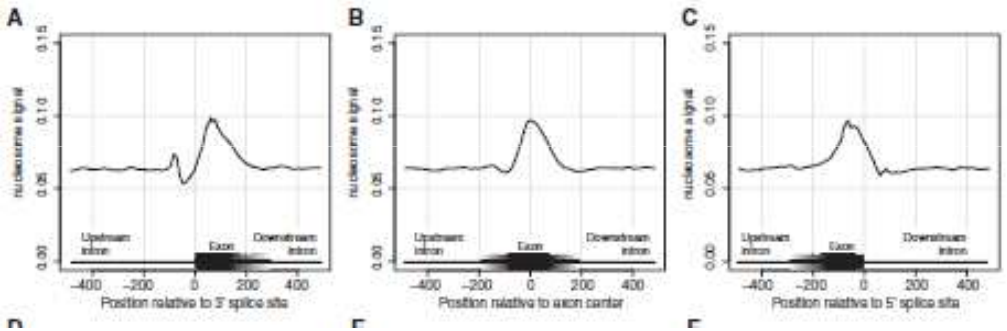
Effects of nucleosomes on RNA Pol II elongation rate
Nucleosomes are preferentially positioned on exons as compared to introns

Molecular Cell 36, 245–254, October 23, 2009

# Biased Chromatin Signatures around Polyadenylation Sites and Exons

Noah Spies,[1,2,5] Cydney B. Nielsen,[1,5,6] Richard A. Padgett,[3] and Christopher B. Burge[1,4,*]
[1]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[2]Whitehead Institute for Biomedical Research, Cambridge, MA 02142, USA
[3]Cleveland Clinic Foundation, Cleveland, OH 44195, USA
[4]Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02142, USA
[5]These authors contributed equally to this work
[6]Present address: Michael Smith Genome Sciences Centre, Vancouver, BC V5Z 4S6, Canada
*Correspondence: cburge@mit.edu

Core RNA-processing reactions in eukaryotic cells occur cotranscriptionally in a chromatin context, but the relationship between chromatin structure and pre-mRNA processing is poorly understood. We observed strong nucleosome depletion around human polyadenylation sites (PAS) and nucleosome enrichment just downstream of PAS. In genes with multiple alternative PAS, higher downstream nucleosome affinity was associated with higher PAS usage, independently of known PAS motifs that function at the RNA level. Conversely, exons were associated with distinct peaks in nucleosome density. Exons flanked by long introns or weak splice sites exhibited stronger nucleosome enrichment, and incorporation of nucleosome density data improved splicing simulation accuracy. Certain histone modifications, including H3K36me3 and H3K27me2, were specifically enriched on exons, suggesting active marking of exon locations at the chromatin level. Together, these findings provide evidence for extensive functional connections between chromatin structure and RNA processing.




nucleosomes are positioned at exons
(data from MNase-Seq experiments)

ChIP-Seq experiments allow measurement of histone modification frequency on exons / introns

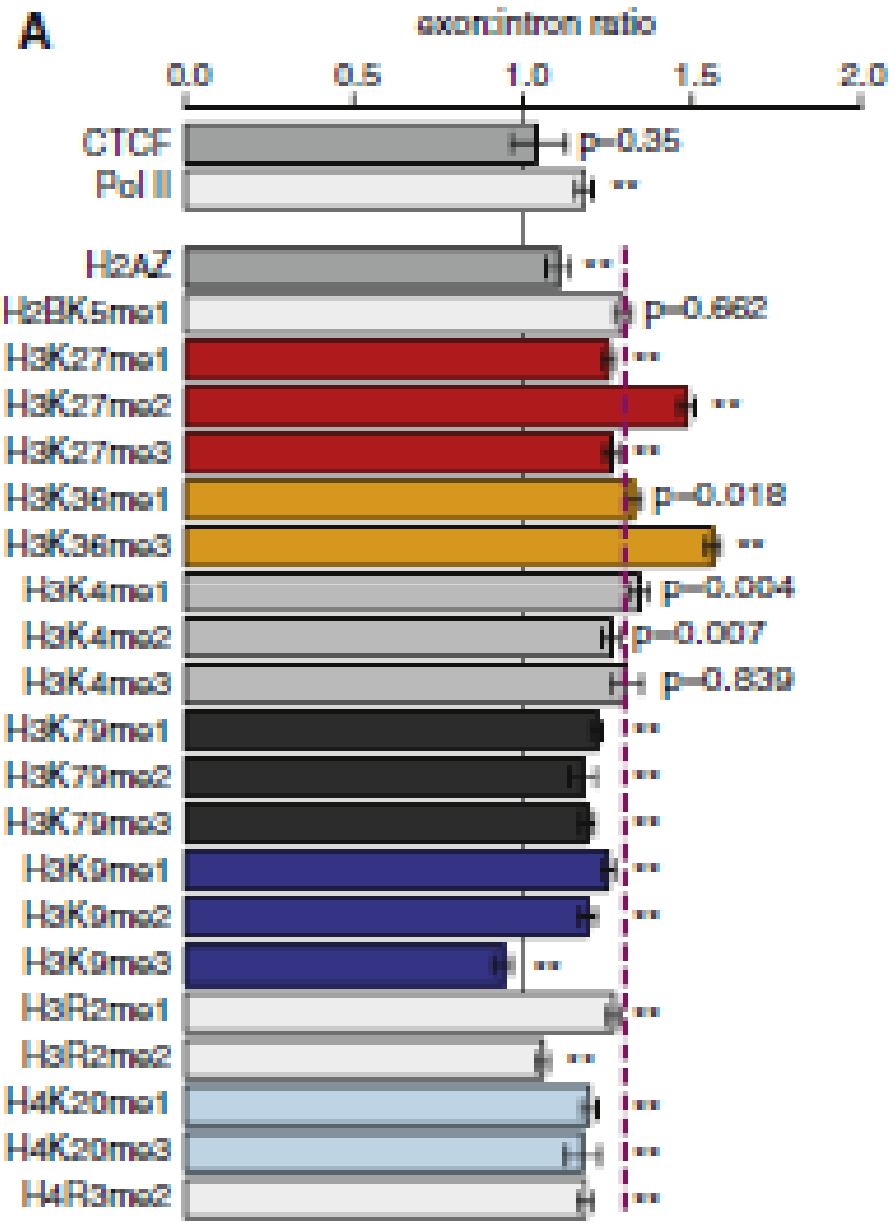



Figure 2. Exon-Biased Distribution of Specific Histone H3 Methylation Marks. (A) ChIP enrichment for exons, relative to flanking intronic regions, compared to 1.0 (CTCF and Pol II) or histone overall average of 1.3 (purple dashed line). Error bars are 95% confidence intervals (resampling). **p < 0.01 after correction for multiple testing (resample test, Bonferroni corrected).

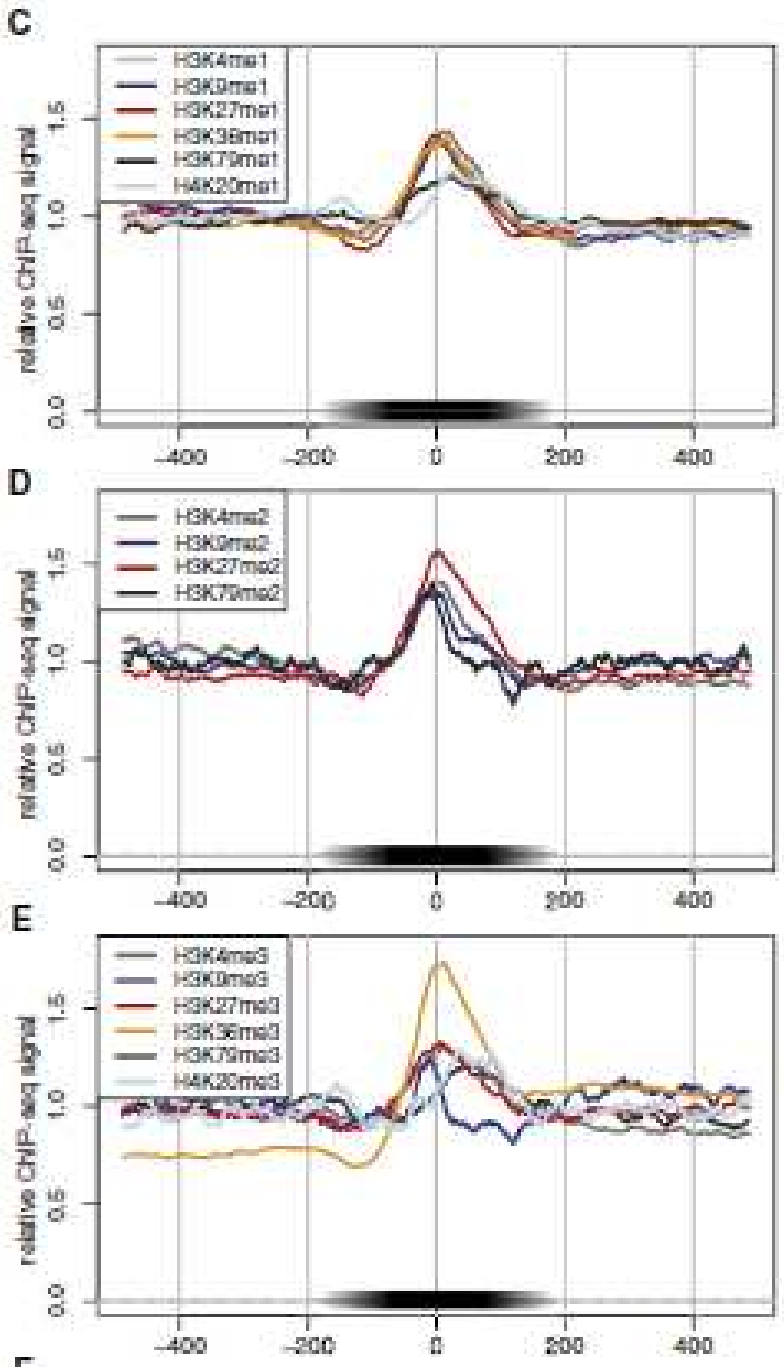



Figure 2. Exon-Biased Distribution of Specific Histone H3 Methylation Marks. (B–F) (B) Histone marks are similarly enriched in highly and lowly expressed genes. Profiles centered on exons for (C) monomethyl histone marks, (D) dimethyl histone marks, and (E) . (C)–(F) are normalized to average library ChIP signal across the displayed region.

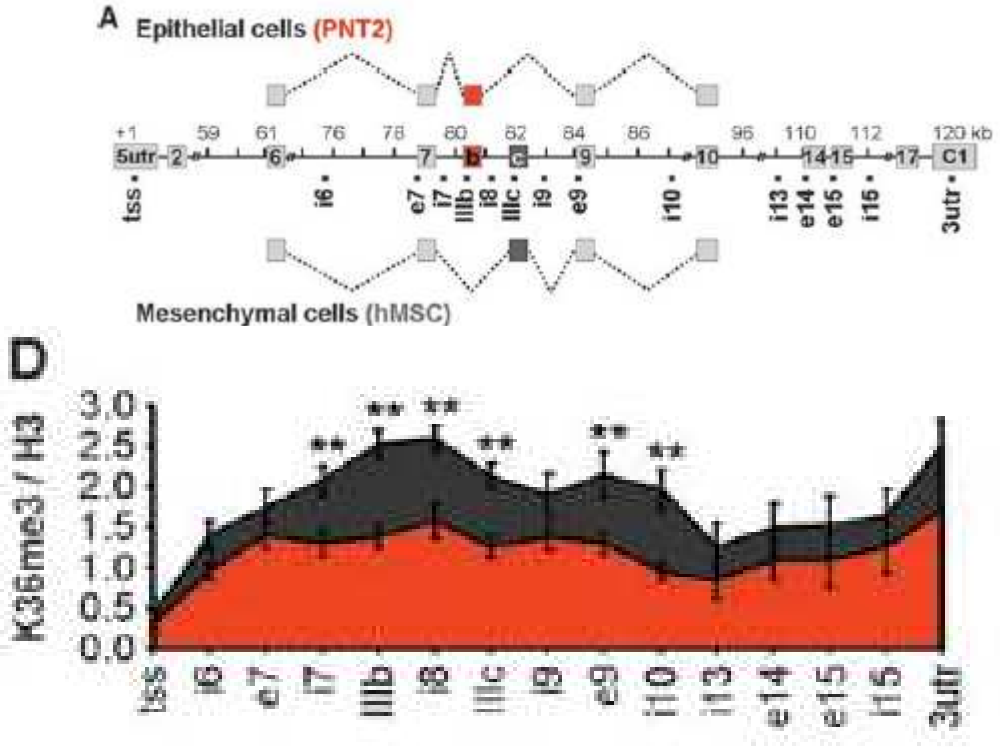
# Regulation of Alternative Splicing by Histone Modifications

Reini F. Luco,[1] Qun Pan,[2] Kaoru Tominaga,[3] Benjamin J. Blencowe,[2]
Olivia M. Pereira-Smith,[3] Tom Misteli[1]*

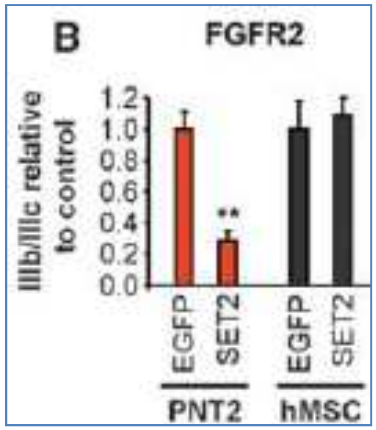
Alternative splicing of pre-mRNA is a prominent mechanism to generate protein diversity, yet its regulation is poorly understood. We demonstrated a direct role for histone modifications in alternative splicing. We found distinctive histone modification signatures that correlate with the splicing outcome in a set of human genes, and modulation of histone modifications causes splice site switching. Histone marks affect splicing outcome by influencing the recruitment of splicing regulators via a chromatin-binding protein. These results outline an adaptor system for the reading of histone marks by the pre-mRNA splicing machinery.

FGFR2 pre-mRNA tissue-specific exon IIIb / IIIc alternative splicing was studied in PNT2 (prostate normal epithelium) and in hMSC (human mesenchimal stem) cells.

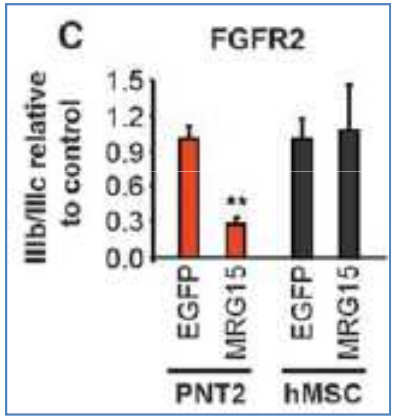
The level of H3K36 trimethylation was assessed by ChIP-qPCR along the gene in these cells. Cell-specific over-representation in hMSC was observed around exons/introns interested by alternative splicing:



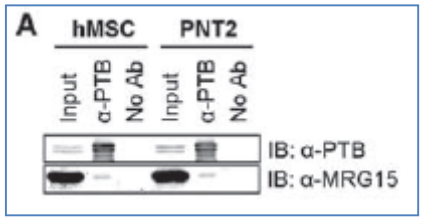
The HMT specific to H3K36 is SET2.
When SET2 is overexpressed in epithelial cells, IIIb/IIIc raio falls by 75%:



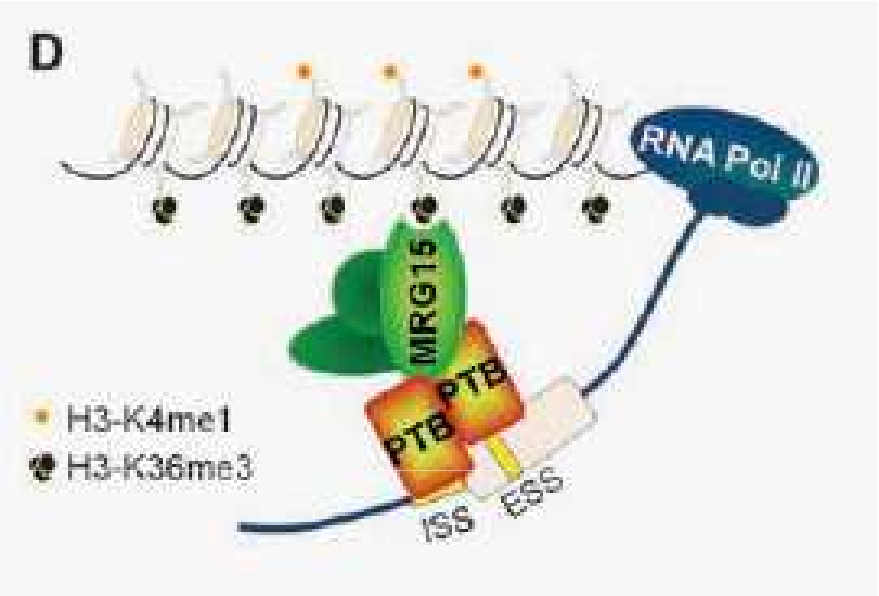

H3K36(me3) is recognized by the bromodomain protein MERG15.
When MERG15 is overexpressed in epithelial cells, IIIb/IIIc raio falls by 75%:




MERG15 co-immunoprecipitates with the RNA binding protein PTB



Suggested model:




Such a model makes the first link ever between epigenetics and alternative splicing

Importantly, this explain a possible mechanism of epigenetic AS “memory” at the cell level, that can also explain phenomena like Dscam

In addition this model is compatible with nucleosome positioning data.

The Authors present a more elaborated model in a Review article published this year
Luco et al., 2011, Cell, 144:16-26.  (available in previous lecture materials).

The Authors also present evidences from the literature that similar mechanisms may concern other chromatin binding proteins and other RNA binding proteins.

Finally, an integrative model is presented, that takes into account all the different mechanisms and models proposed for alternative splicing control.

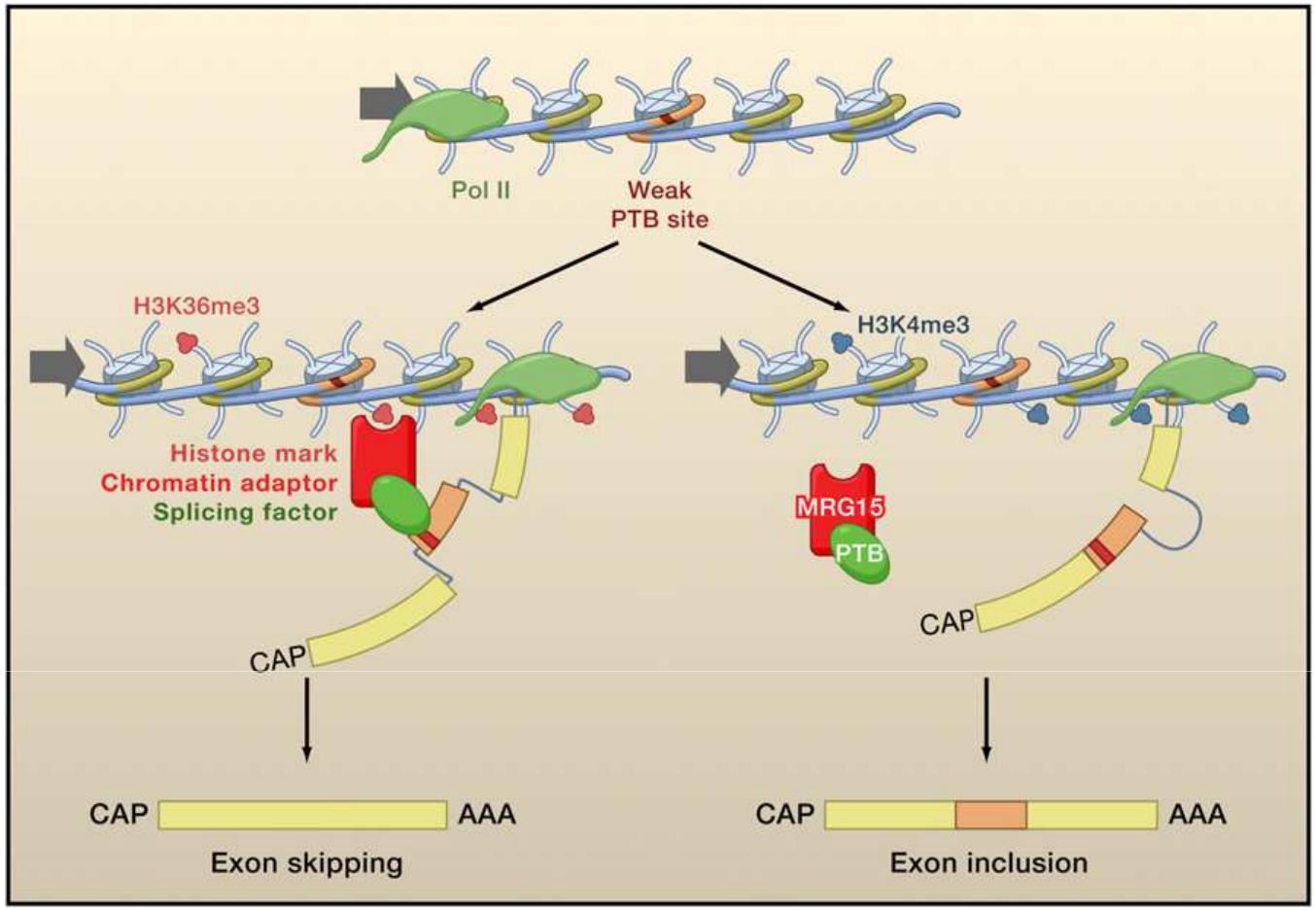



Figure 3. The Chromatin-Adaptor Model of Alternative Splicing. Histone modifications along the gene determine the binding of an adaptor protein that reads specific histone marks and in turn recruits splicing factors. In the case of exons whose alternative splicing is dependent on poly-pyrimidine tractbinding protein (PTB) splicing factor, high levels of trimethylated histone 3 lysine 36 (H3K36me3, red) attract the chromatin-binding factor MRG15 that acts as an adaptor protein and by protein-protein interaction helps to recruit PTB to its weaker binding site inducing exon skipping. If the PTBdependent gene is hypermethylated in H3K4me3 (blue), MRG15 does not accumulate along the gene, and PTB is not recruited to its target premRNA, thus favoring exon inclusion.

This is the first demonstration of a mechanistic link between chromatin and alternative splicing. Other protein-protein interaction between chromatin-competent proteins and RNA binding proteins is present in the literature, however no direct demonstration of a mechanism was given to date. Nonetheless, interactions suggest a possible functional role that should be worked out in the future.

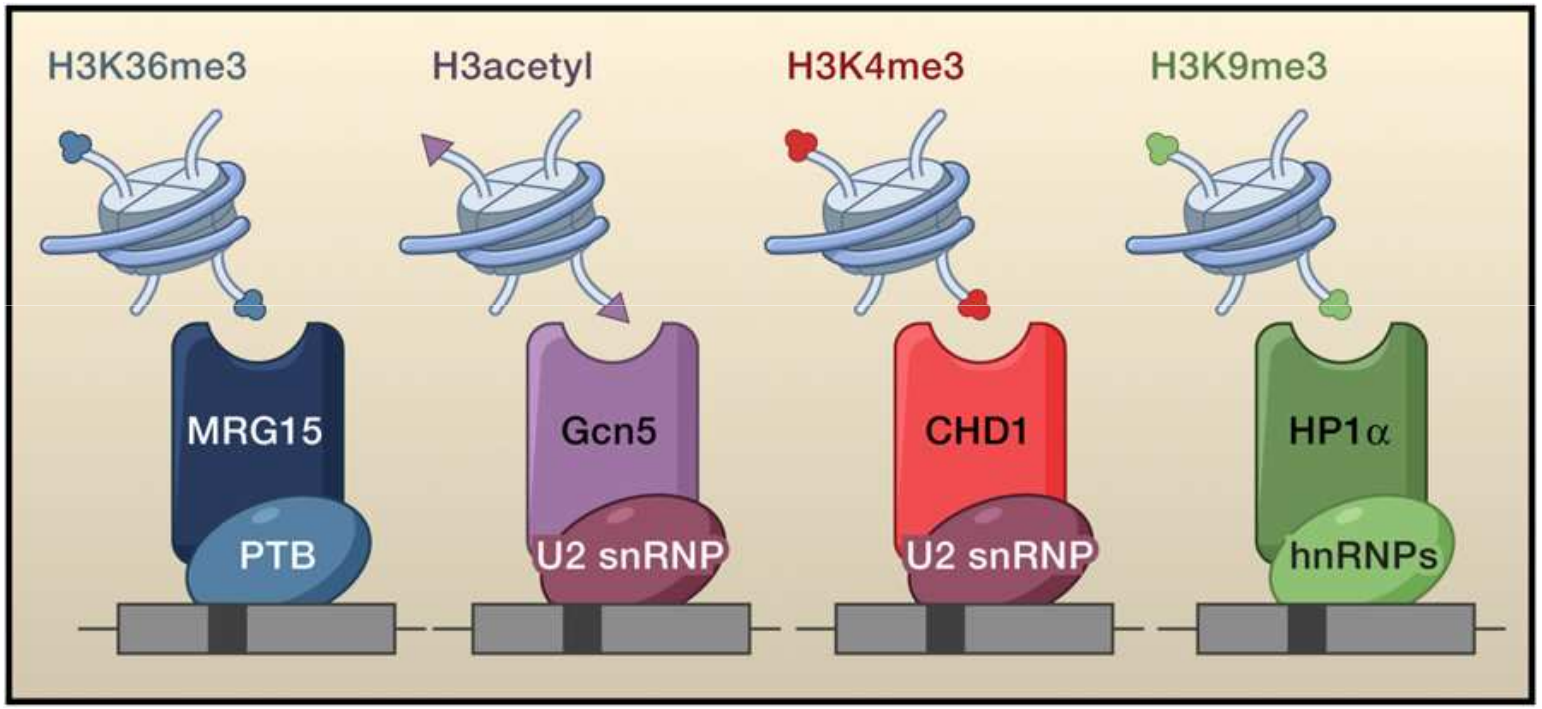



Figure 4. Chromatin-Adaptor Complexes
Several histone modification-binding chromatin proteins interact with splicing factors (Luco et al., 2010; Sims et al., 2007; Gunderson and Johnson, 2009; Piacentini et al., 2009; Loomis et al., 2009).

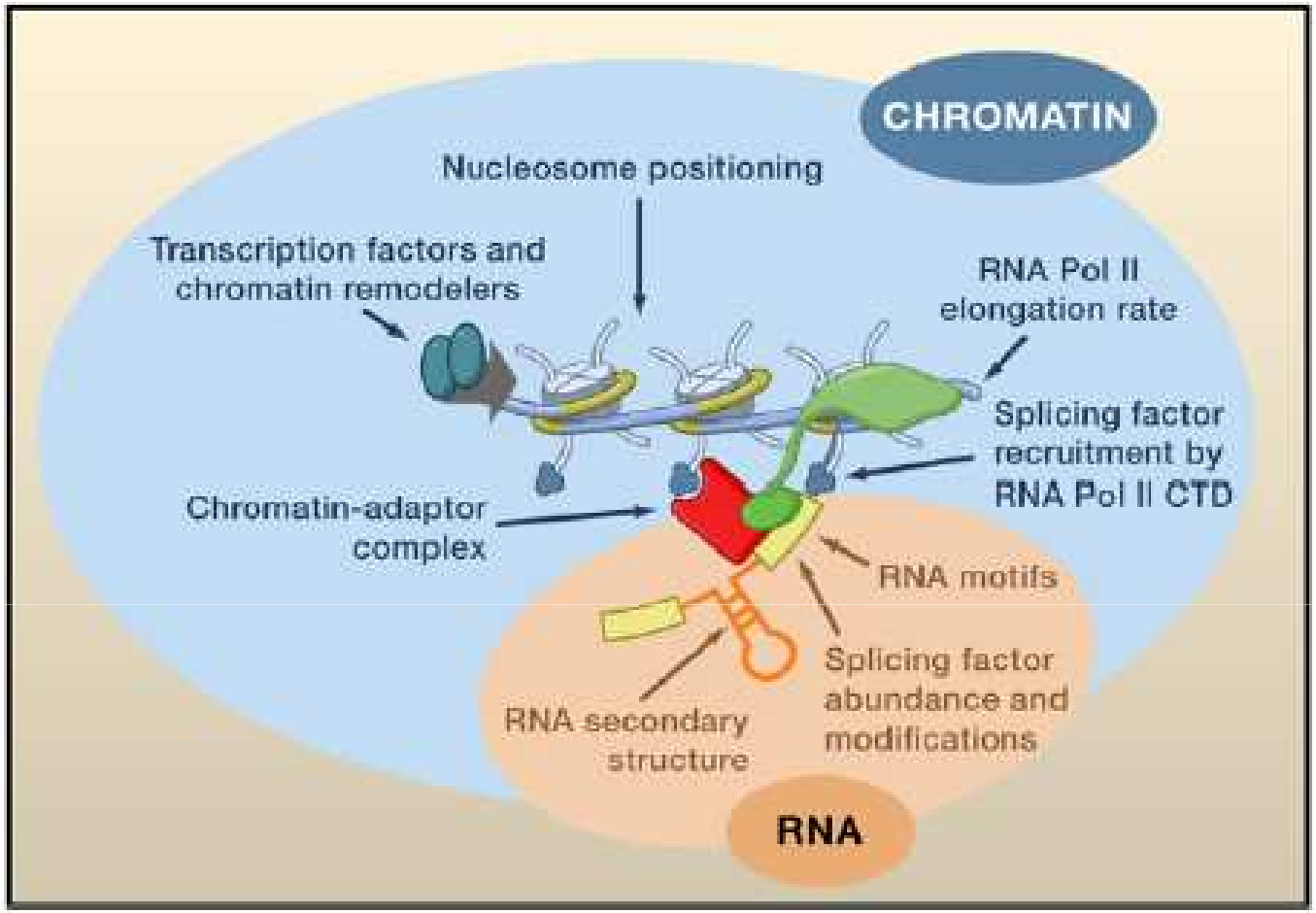



Figure 5. An Integrated Model for the Regulation of Alternative Splicing
Alternative splicing patterns are determined by a combination of parameters including cis-acting RNA regulatory elements and RNA secondary structures (highlighted in orange) together with transcriptional and chromatin properties (highlighted in blue) that modulate the recruitment of splicing factors to the premRNA.

A model of epigenetic “memory” of alternative splicing in the cells

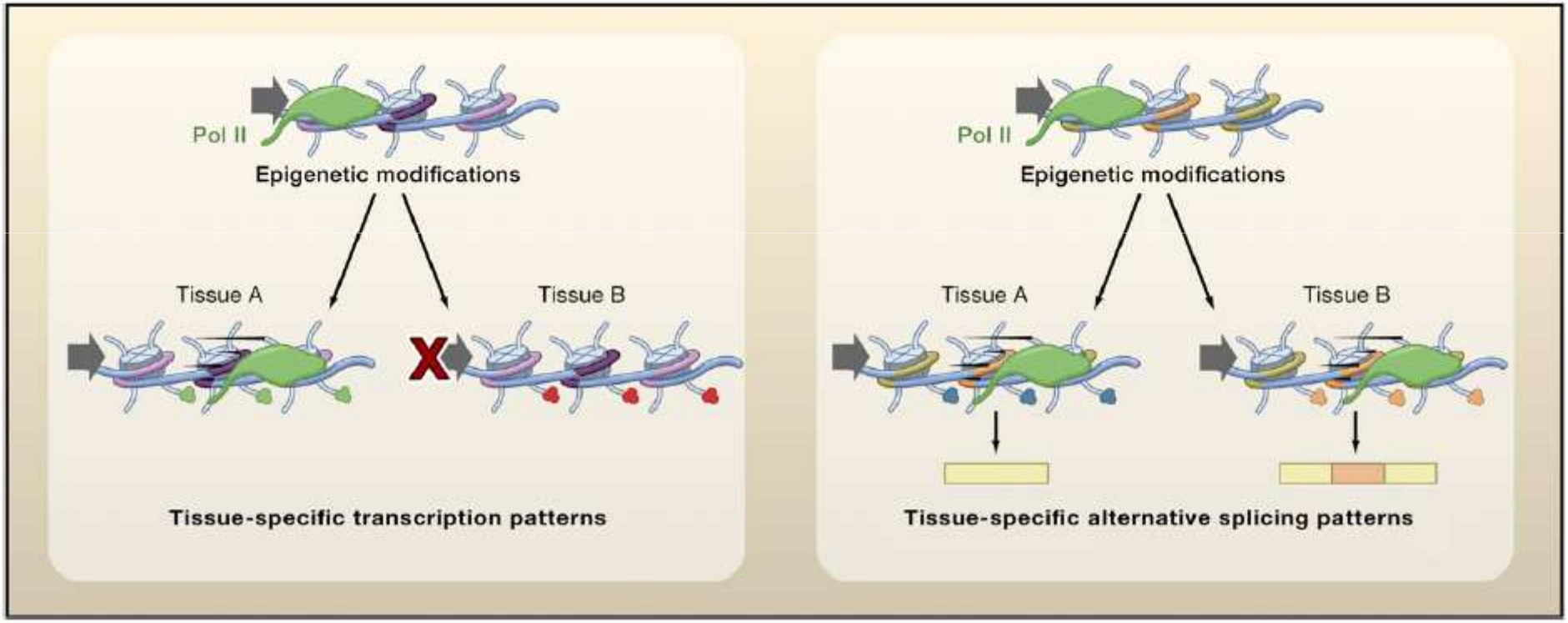



**Figure 6. The Epigenetics of Alternative Splicing**
The combination of histone modifications along a gene establishes and maintains tissue-specific transcription patterns (left panel), as well as heritable tissue-specific alternative splicing patterns (right panel).

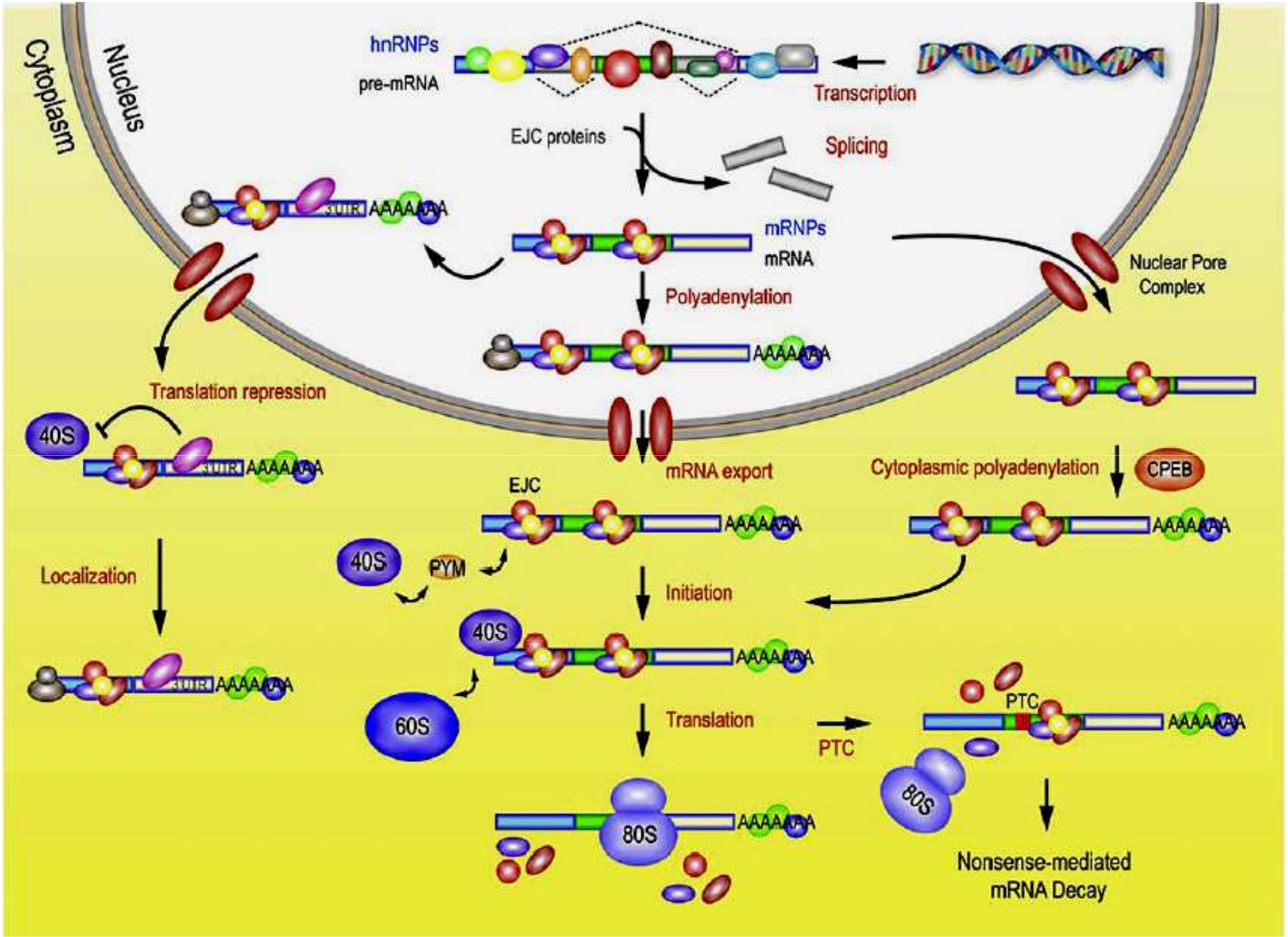
Next...

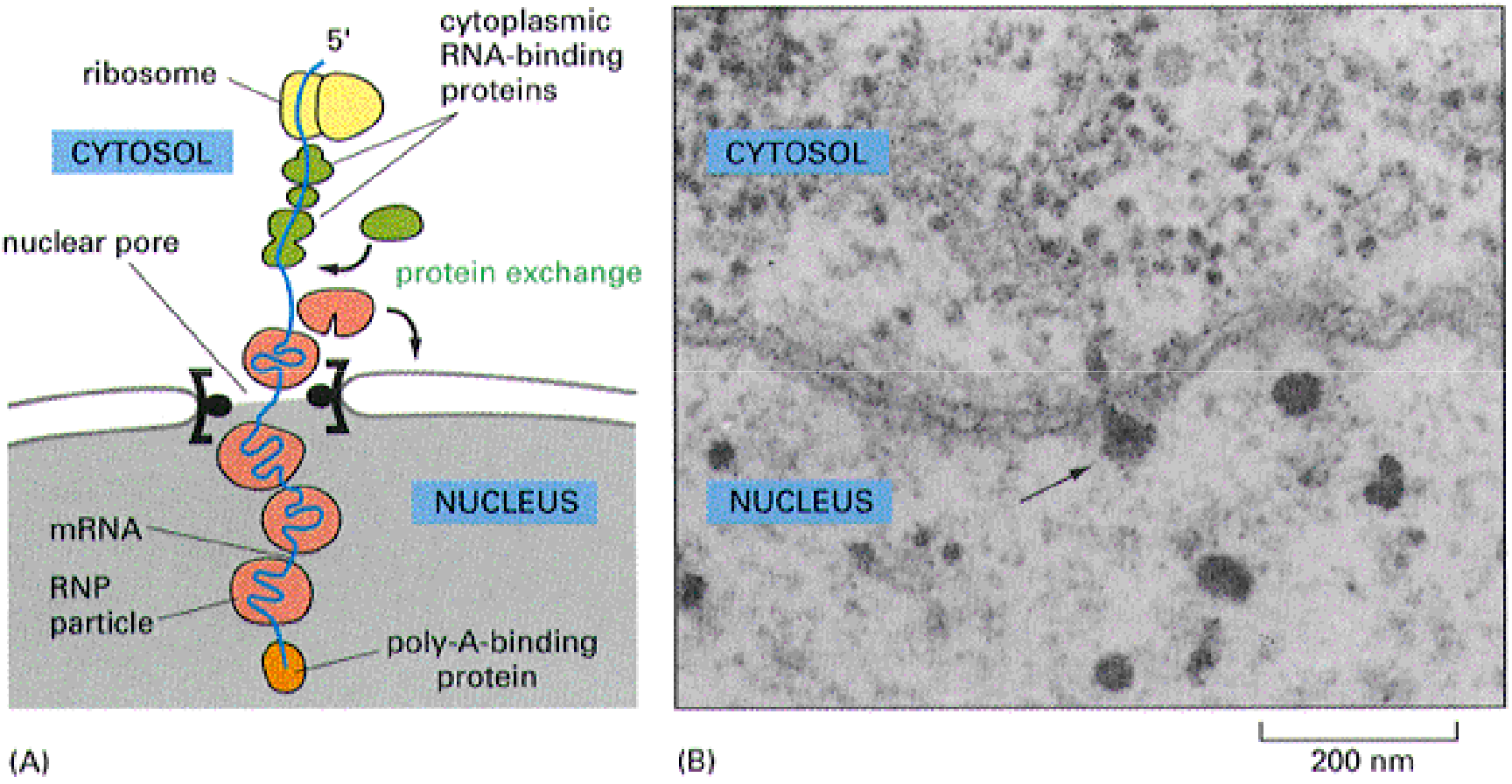
Nuclear→cytoplasmatic export

Cellular Localization

Translation

Degradation


Cytoplasm
Nucleus
hnRNPs
pre-mRNA
Transcription
EJC proteins
Splicing
mRNPs
mRNA
Polyadenylation
Nuclear Pore Complex
Translation repression
40S
mRNA export
EJC
Cytoplasmic polyadenylation
CPEB
Localization
40S
PYM
Initiation
40S
60S
Translation
PTC
PTC
80S
80S
Nonsense-mediated mRNA Decay



Chromosoma (2008) 117:319–331
DOI 10.1007/s00412-008-0158-4

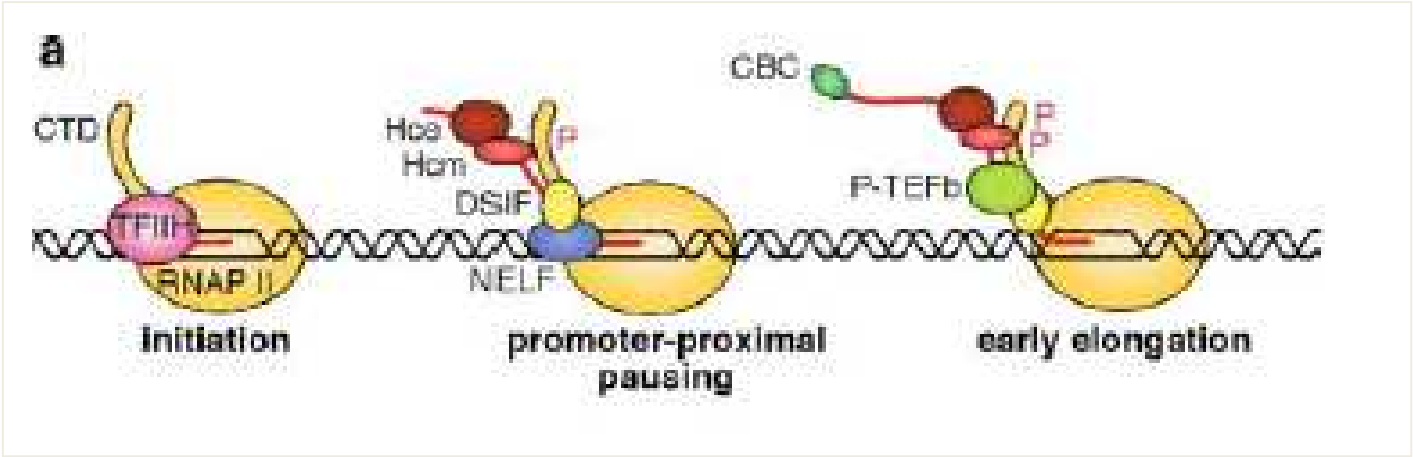
REVIEW

# Biogenesis of mRNPs: integrating different processes in the eukaryotic nucleus

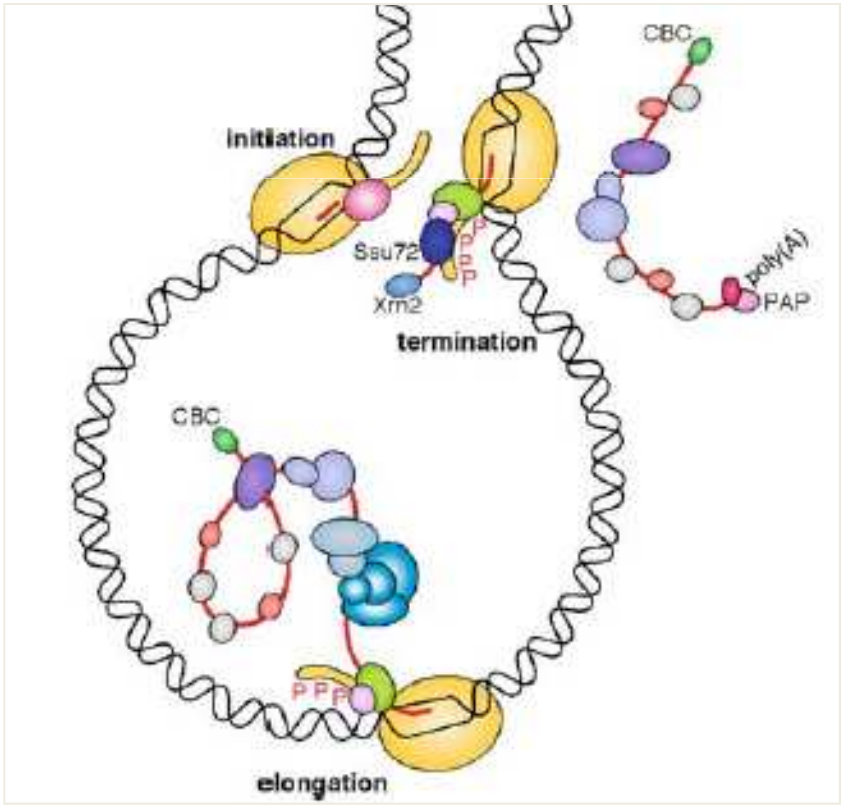
Rosa Luna · Hélène Gaillard ·
Cristina González-Aguilera · Andrés Aguilera

Review

Abstract. Transcription is a central function occurring in the nucleus of eukaryotic cells in coordination with other nuclear processes. During transcription, the nascent pre-mRNA associates with mRNA-binding proteins and undergoes a series of processing steps, resulting in export competent mRNA ribonucleoprotein complexes (mRNPs) that are transported into the cytoplasm. Experimental evidence increasingly indicates that the different processing steps (5′-end capping, splicing, 3′-end cleavage) and mRNP export are connected to each other as well as to transcription, both functionally and physically. Here, we review the overall process of mRNP biogenesis with particular emphasis on the functional coupling of transcription with mRNP biogenesis and export and its relationship to nuclear organization.




Several lines of evidence suggest that processing proteins are recruited to the Transcription Elongating Complex (i.e. active RNA Polymerase II) either at the promoter (before transcription begins) or during elongation.

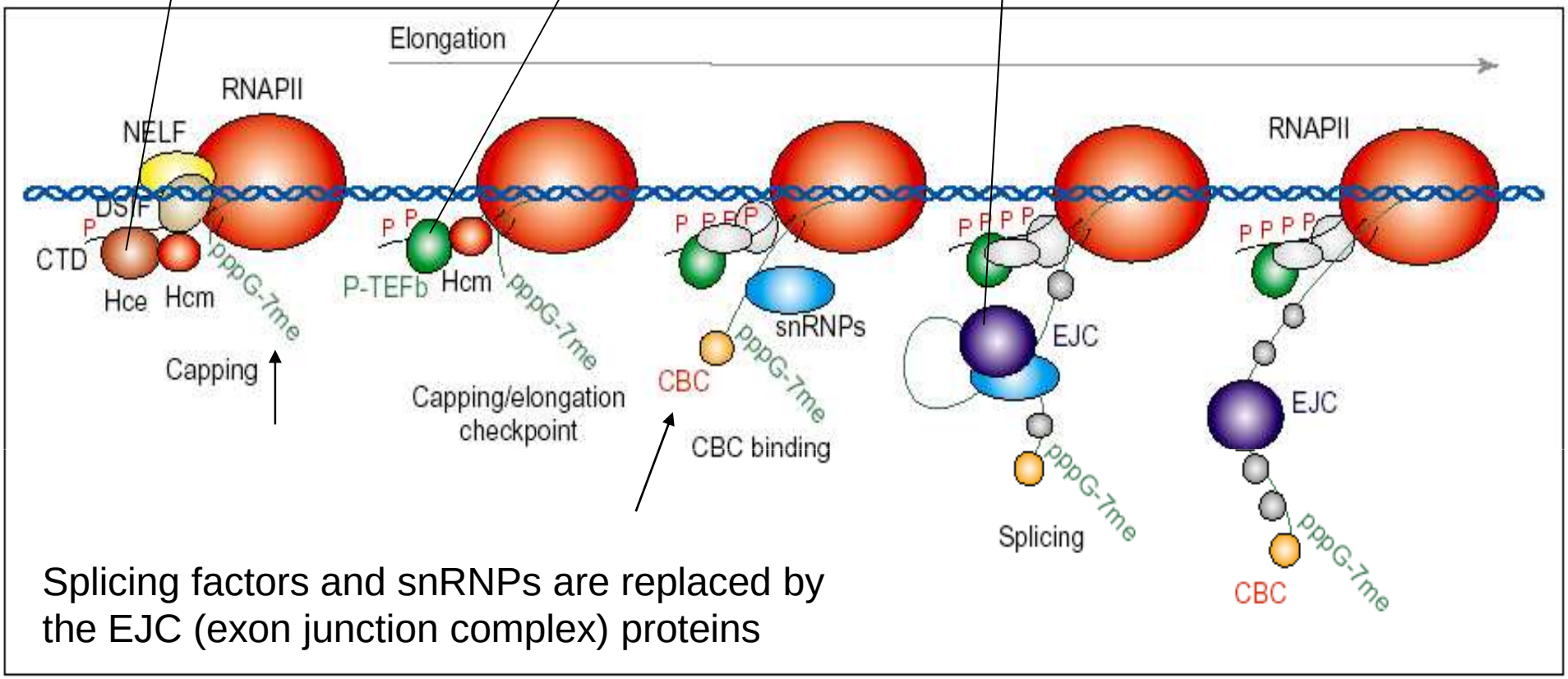


guanine-N7-methyltransferase

Elongation factor

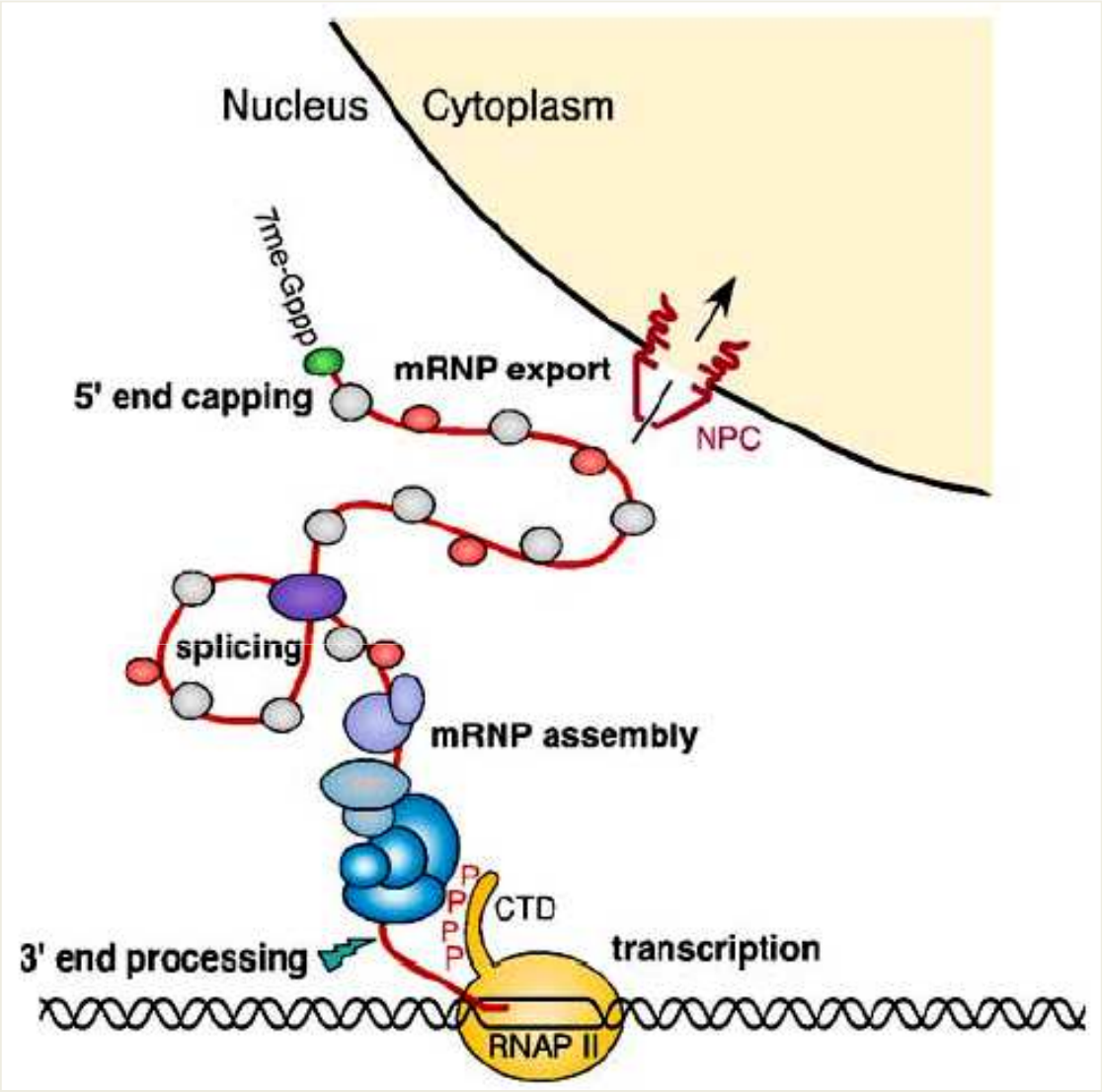
EJC deposited 20-24 nt upstream the exon-exon junction

**Figure 2**



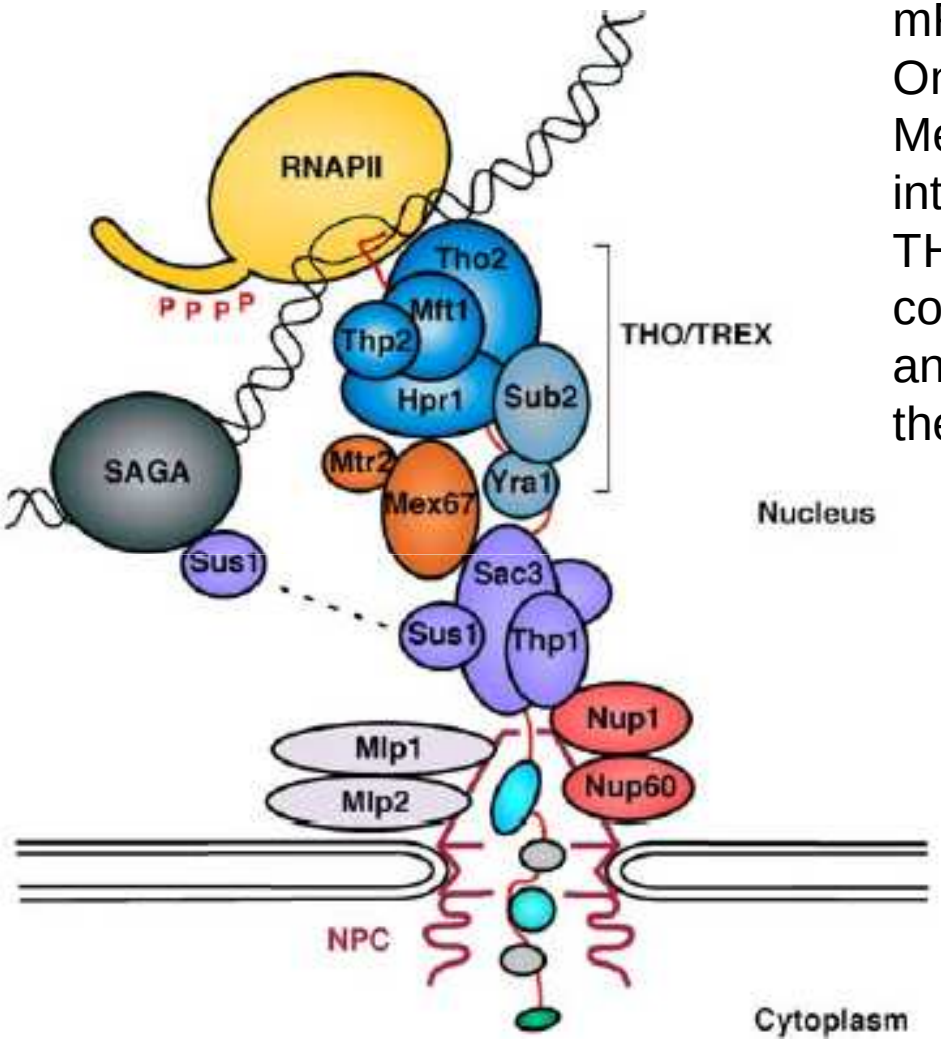

Schematic of the first steps of cotranscriptional mRNP assembly including 5′-end capping, CBC (cap-binding protein complex) loading and the splicing-dependent assembly of the exon junction complex (EJC), as it may occur in humans. After transcription initiation and phosphorylation of the CTD–Ser2 of RNAPII by TFIIH–Cdk7, transcription is paused by the joint action of the DSIF and NELF transcription elongation factors, allowing guanylyltransferase (Hce) and methyltransferase (Hcm) to cap the pre-mRNA 5′-end in two steps as soon as it emerges from the RNAPII. The capping machinery is displaced after completing its task, allowing the elongation factor P-TEFb to access the RNAPII CTD, which leads to massive phosphorylation of Ser5 and to productive elongation being resumed. This serves as a checkpoint to prevent extension of unproperly capped RNAs. Once the first intron is transcribed, the core splicing machinery of snRNPs is recruited to the intron and the EJC is assembled onto the resulting exon–exon junction.

From Aguilera 2005, Curr Op Cell Biol, 17:242.




Proteins necessary for mRNP transport through the nuclear pores are also charged to RNA during transcription

Fig. 1 Schematic view of the nuclear side of eukaryotic gene expression, from transcription to nuclear export. NPC Nuclear pore complex, CTD C-terminal domain of Rpb1, RNAPII RNA polymerase II.




mRNA export.
One of the most important proteins is Mex67/Mtr2, that is a briding protein, i.e. it interacts with adaptors such as the THO/TREX transport complex, that is composed of several RNA-binding proteins, and with other adaptors that connect it with the nucleoporins.

Interestingly, several interaction between TREX and /or Mex67(Mtr2 were found with SAGA (the major histone acetyltransferase complex in Yeast).
This allows a model where active genes (i.e. those that contain SAGA) are recruited close to the nuclear pores to allow transcription, processing and transport at the same time.
It should be noted, however, that several genes are trascribed also in the middle of the nucleus, rather than at the periphery.

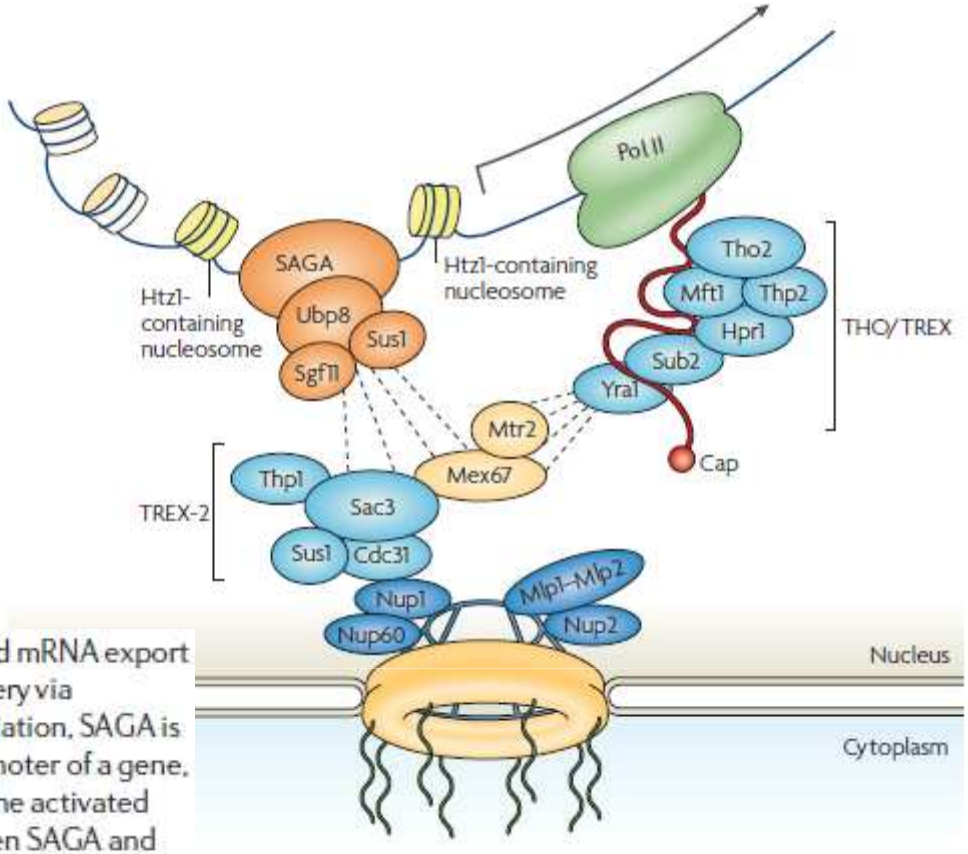



Figure 4 | **Gene gating and mRNA export.** Model of transcription-coupled mRNA export in yeast, which involves gating of the activated gene to the nuclear periphery via interaction with the nuclear pore complex (NPC). During transcription initiation, SAGA is recruited as part of a large transcription pre-initiation complex to the promoter of a gene, which at this point is located in the interior of the nucleus. Subsequently, the activated gene becomes tethered to the nuclear periphery via an interaction between SAGA and the nuclear pore-bound TREX-2 complex (which consists of the subunits Sac3, Thp1, Sus1 and Cdc31). Transcription of the tethered gene generates a nascent transcript that becomes assembled, with the aid of the THO/TREX complex, into an export-competent messenger ribonucleoprotein (mRNP). Thus, the mRNP is brought into the vicinity of the NPC-associated Mex67–Mtr2 mRNA export machinery. Mex67 can also directly contact the promoter region independently of the mRNA (not shown). Sus1 is a functional component of both TREX-2 and SAGA. Within SAGA, Sus1 is part of a heterotrimeric subcomplex together with the histone H2B deubiquitylating enzyme Ubp8 and Sgf11, a protein of unknown function. The nucleoporin Nup1 is part of the TREX-2 docking site at the nuclear basket. Nup2 was shown to associate with active promoters and might tether genes through the histone variant Htz1 in yeast. The Mlp1 and Mlp2 proteins contribute to RNA surveillance by retention of unspliced mRNA. Dotted lines represent predicted interactions between the complexes. TREX, transcription-coupled export complex. Proteins in the same complex have the same colour.

This will connect the processing story with the data illustrating the differential localization of hetrochromatin and euchromatin in the nucleus.

You must “splice” these concepts to what you’ve learnt in the 22 October Cell Biology lecture

# The nuclear envelope and transcriptional control

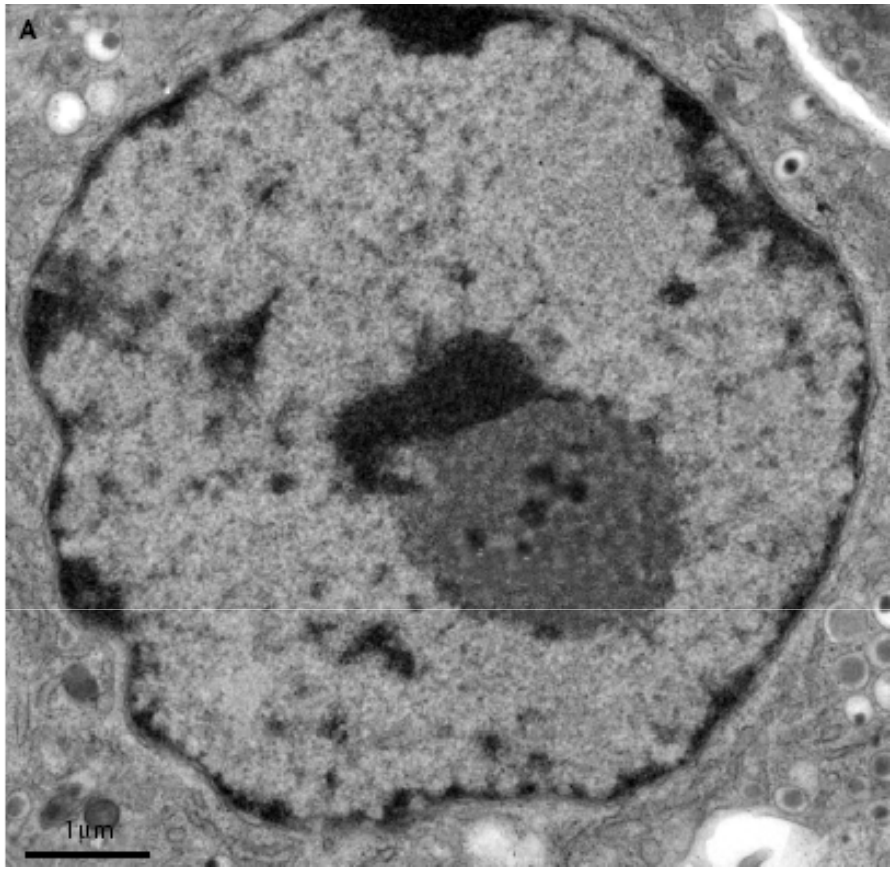
Review 1

*Asifa Akhtar* and Susan M. Gasser‡*

Abstract | Cells have evolved sophisticated multi-protein complexes that can regulate gene activity at various steps of the transcription process. Recent advances highlight the role of nuclear positioning in the control of gene expression and have put nuclear envelope components at centre stage. On the inner face of the nuclear envelope, active genes localize to nuclear-pore structures whereas silent chromatin localizes to non-pore sites. Nuclear-pore components seem to not only recruit the RNA-processing and RNA-export machinery, but contribute a level of regulation that might enhance gene expression in a heritable manner.

2007

A




B

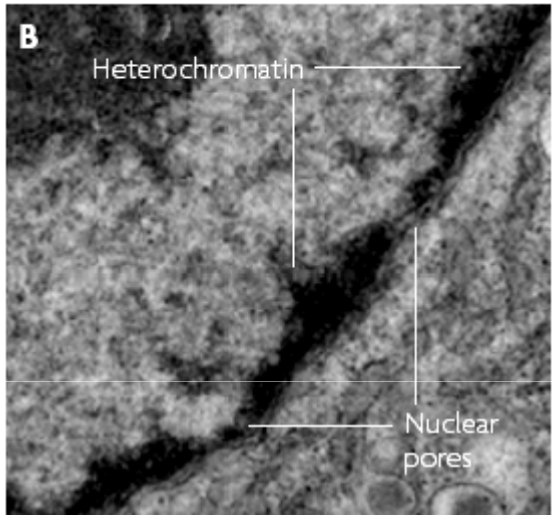



Figure 1 | **Heterochromatin in mammalian and yeast cells is distinct from nuclear pores.** **A** | An electron micrograph of the mammalian liver nucleus (with an enlarged section shown in part **B**), showing dense-staining heterochromatin located around the nucleolus and against the nuclear envelope. Nuclear pores open onto lighterstaining open chromatin.

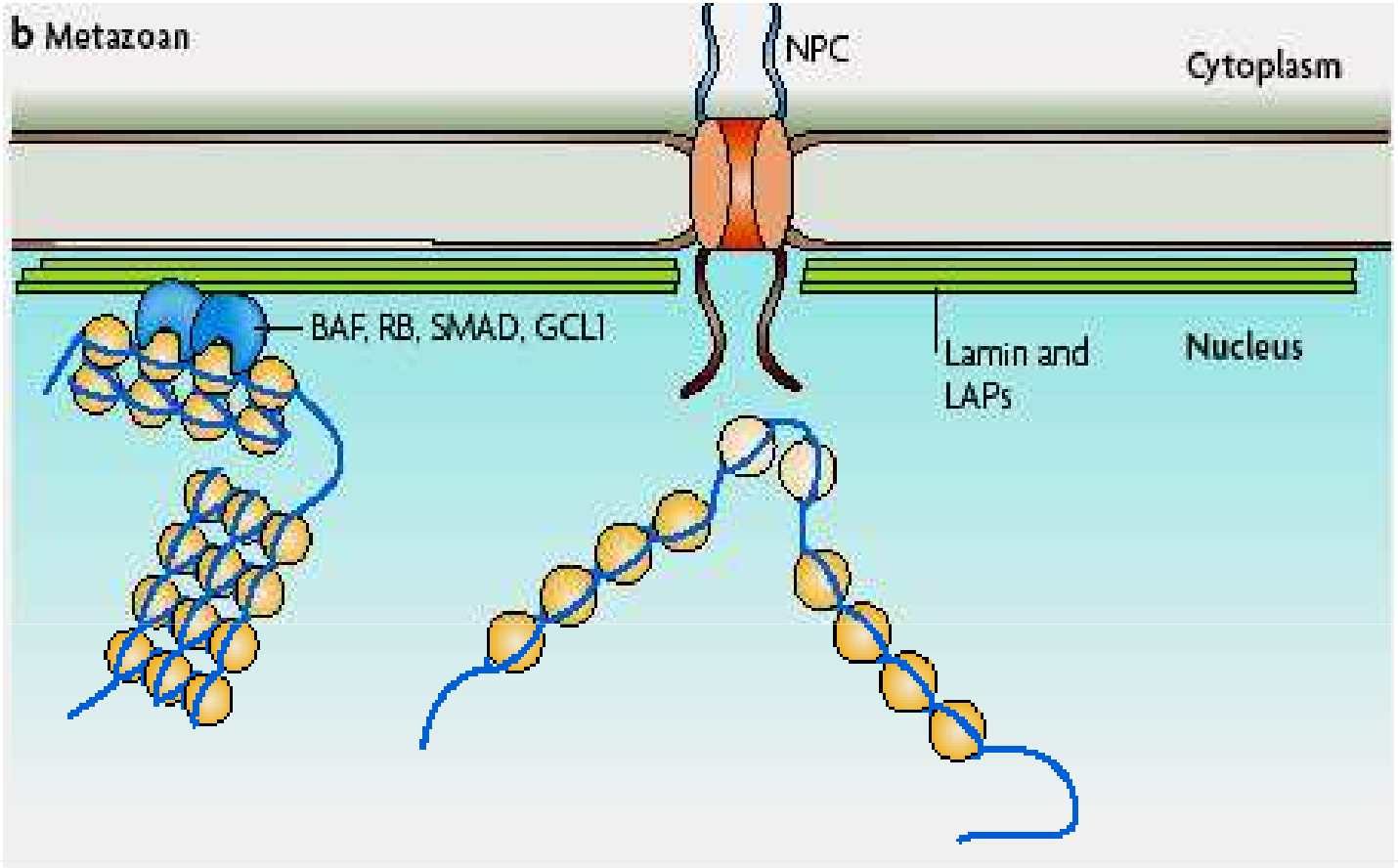



Figure 2 | **The nuclear periphery in metazoans and yeast.** In eukaryotic cells, the nuclear compartment is separated from the cytoplasm by the inner and outer nuclear membranes. This membrane bilayer is perforated by nuclear pores, which are constituted by a large multiprotein complex (the nuclear pore complex (NPC)) that is composed of about 30 proteins. This nuclear membrane, together with the pores, is commonly referred as the 'nuclear envelope' (NE). **b** | In metazoan nuclei, the nuclear envelope is underlaid by a continuous meshwork of lamins and lamin-associated proteins (LAPs), which preferentially associate with inactive chromatin regions. Increasing evidence implicates interactions of chromatin with various nuclear-envelope components in gene repression as well as gene activation. BAF, barrier to autointegration factor; GCL1, germ-cell-less homologue; RB, retinoblastoma 1.

An extraordinary number of different proteins were found involved in RNA export in yeast.

Most of them were recruited to the elongating RNA Polymerase co-trascriptionally

This astonishing number of RBPs poses the question whether all of them are present together on the same RNA, or whether these proteins show a certain degree of specificity for different classes of RNA.

A study in yeast using RNA immunoprecipitation followed by microarray identification suggested a degree of specificity.

From Aguilera 2005, Curr Op Cell Biol, 17:242.

**Table 1**

**Abundance and localization of yeast proteins involved in mRNA biogenesis and metabolism and their respective metazoan orthologs.**

| Yeast protein | Metazoan protein | Yeast/metazoan protein complex | Primary subcellular localization in yeast[a] | Molecules per cell in yeast[b] | Co-transcriptional recruitment in yeast[c] |
|---|---|---|---|---|---|
| Rpb2 | Rpb2 | RNAPII | Nucleus | 18 700 | + |
| Rpb3 | Rpb3 | RNAPII | Nucleus | 10 000 | + |
| Med2 | Med2 | RNAPII | Nucleus | 10 800 | |
| Dst1 | TFIIS | | Nucleus | 6680 | + |
| Spt4 | Spt4 | Spt4,5/DSIF | Nucleus | 4490 | + |
| Spt5 | Spt5 | Spt4,5/DSIF | Nucleus | 2340 | + |
| Spt6 | Spt6 | | Nucleus | 8890 | + |
| Tho2 | Tho2 | THO | Nucleus | 521 | + |
| Hpr1 | Hpr1 | THO | Nucleus | 1340 | + |
| Mft1 | | THO | Nucleus | 5910 | + |
| Thp2 | | THO | Nucleus | 2840 | + |
| Sub2 | UAP56/HEL | TREX | Nucleus | 51 700 | + |
| Yra1 | Aly/REF | TREX | Nucleus | – | + |
| Npl3 | | | Nucleus | 78 700 | + |
| Nab2 | | | Nucleus | 9670 | + |
| Rrp6 | Rrp6 | Nuclear exosome | Nucleus and nucleolus | 2160 | (+) |
| Gbp2 | | | Nucleus | 2540 | + |
| Hrb1 | | | Nucleus | 1990 | + |
| Rna14 | CStF-77 | CFI | Nucleus | 5350 | + |
| Pcf11 | Pcf11 | CFIA | Nucleus | 2800 | + |
| Rat1 | Xrn2 | | Nucleus | 623 | + |
| Dbp5 | Dbp5 | | Nuclear periphery | 14 900 | (+)[d] |
| Mex67 | TAP | Mex67-Mtr2/TAP-p15 | Nuclear periphery | 2830 | |
| Mtr2 | p15 | Mex67-Mtr2/TAP-p15 | Nuclear periphery | | |
| Thp1 | | Thp1-Sac3 | Nuclear periphery | 1140 | |
| Sac3 | | Thp1-Sac3 | Nuclear periphery | 339 | |
| Sus1 | | SAGA and Thp1-Sac3 | Nuclear periphery | | |
| Nup1 | Nup1 | NPC | Nuclear periphery | 468 | – |
| Nup60 | Nup60 | NPC | Nuclear periphery | 4590 | – |
| Hmt1 | | | Nucleus and cytoplasm | 37 600 | + |
| Elp1 | Elp1 | Elongator | Cytoplasm | 10 500 | +[e] |
| Elp3 | Elp3 | Elongator | Cytoplasm | 6090 | +[e] |
| Pab1 | PABPC4/PABP1 | | Cytoplasm | 198 000 | + |
| Pan3 | PAN3 | | Cytoplasm | 1600 | |

Blanks indicate that information is not known. [a]Data taken from [70]. [b]Data taken from [32]. [c]Data obtained from different references as cited in the text. [d]It has been shown for *Chironomus* but there are no data regarding the yeast Dbp5. [e]Shown to associate with RNAPII-driven pre-mRNA.

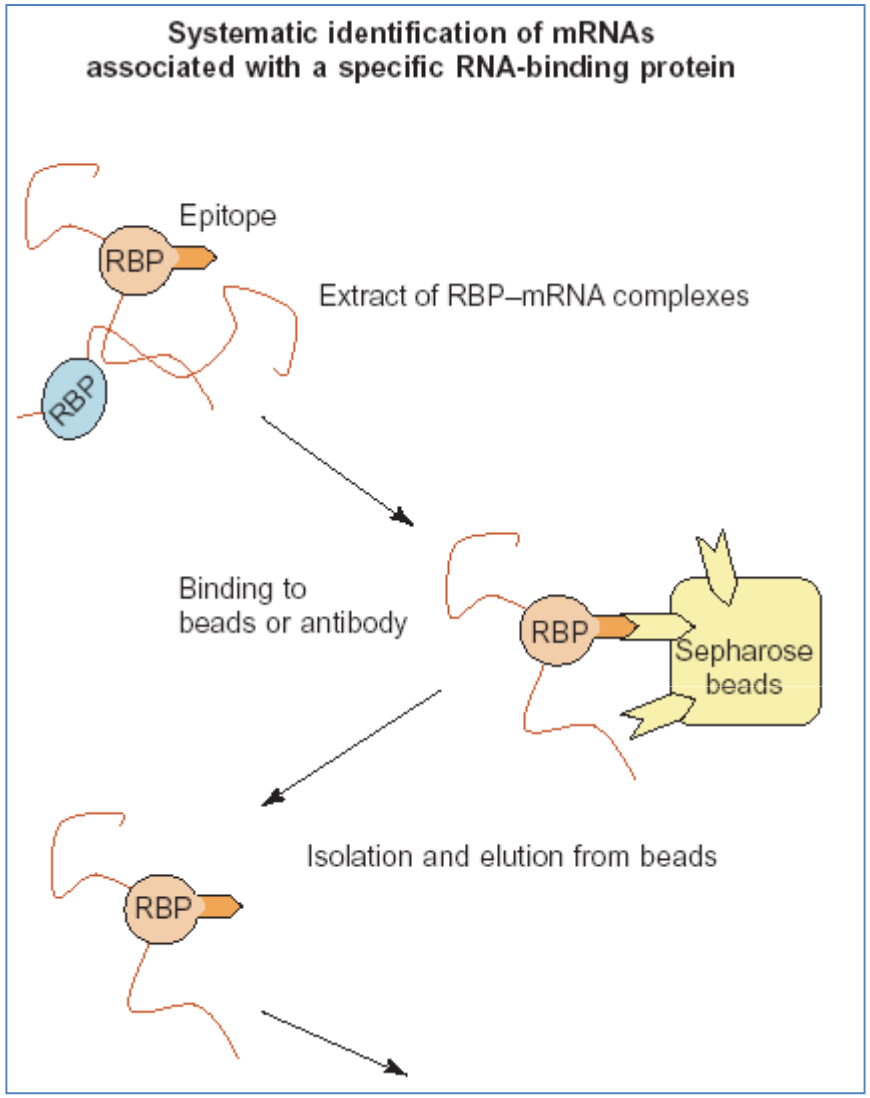
# Genome-wide analysis of RNA–protein interactions illustrates specificity of the mRNA export machinery

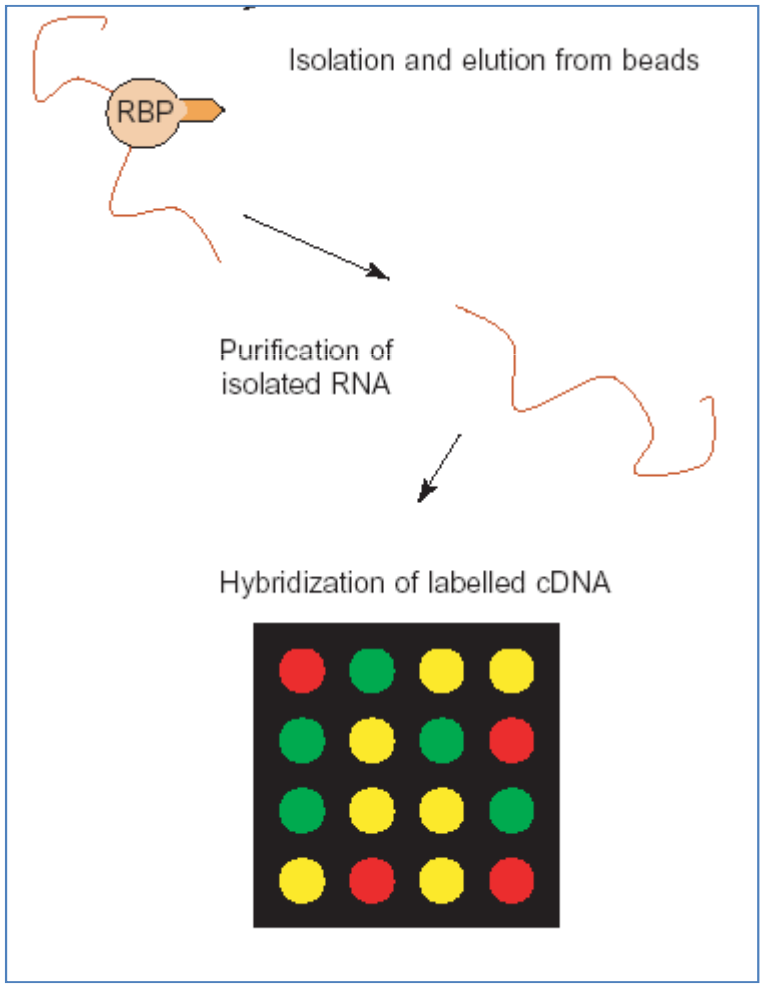
Haley Hieronymus & Pamela A. Silver

nature genetics • volume 33 • february 2003

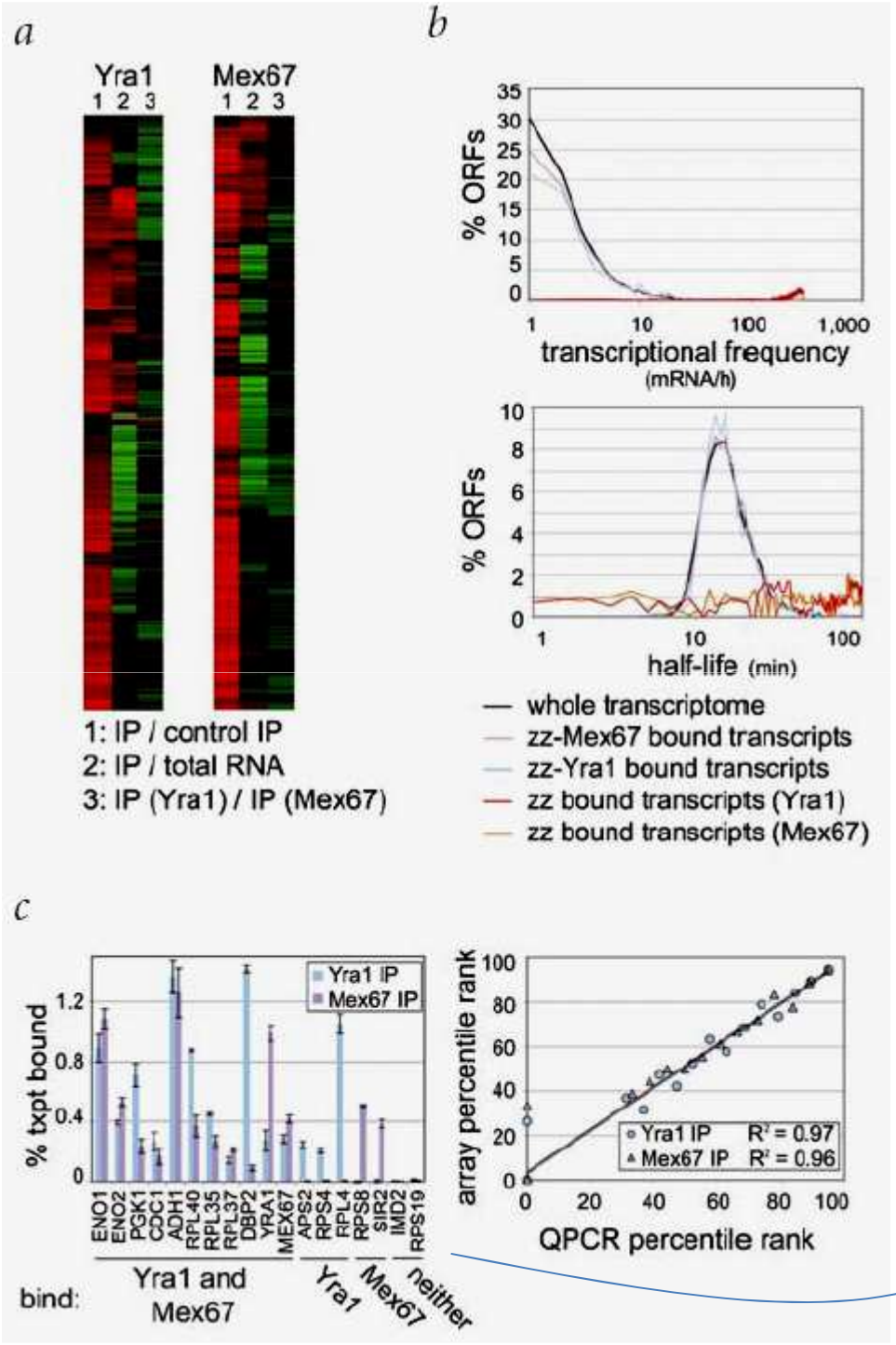
33:155-161.

Nuclear export of mRNA is mediated by a complex machinery of RNA-binding proteins that recognizes and routes mRNAs through a messenger ribonucleoprotein (mRNP) network. The full spectrum of mRNA cargoes for any dedicated mRNA export factor is unknown. We identified the mRNAs that bind two conserved yeast mRNA export factors, Yra1 (refs. 1–5) and Mex67 (refs. 6,7), on a genome-wide scale and determined their level of binding. Yra1 and Mex67 bind approximately 1,000 and 1,150 mRNAs, respectively, corresponding to almost 20% of the yeast genome and roughly 36% of all transcriptional events each. The binding level of Yra1 targets is related to their transcriptional frequency, but that of Mex67 targets is not. Yra1-bound transcripts are enriched in mRNAs that are regulated by a number of transcription factors. Yra1- and Mex67-bound populations also show enrichment of mRNAs encoding distinct functional classes of proteins, some of which are regulated by these transcription factors. We determined that one such transcription factor, Abf1 (refs. 8–10), associates with Yra1. These results indicate a previously unidentified specificity of mRNA export factors, which coordinates the export of transcriptionally co-regulated, functional classes of transcripts, perhaps through interactions with the transcriptional machinery.



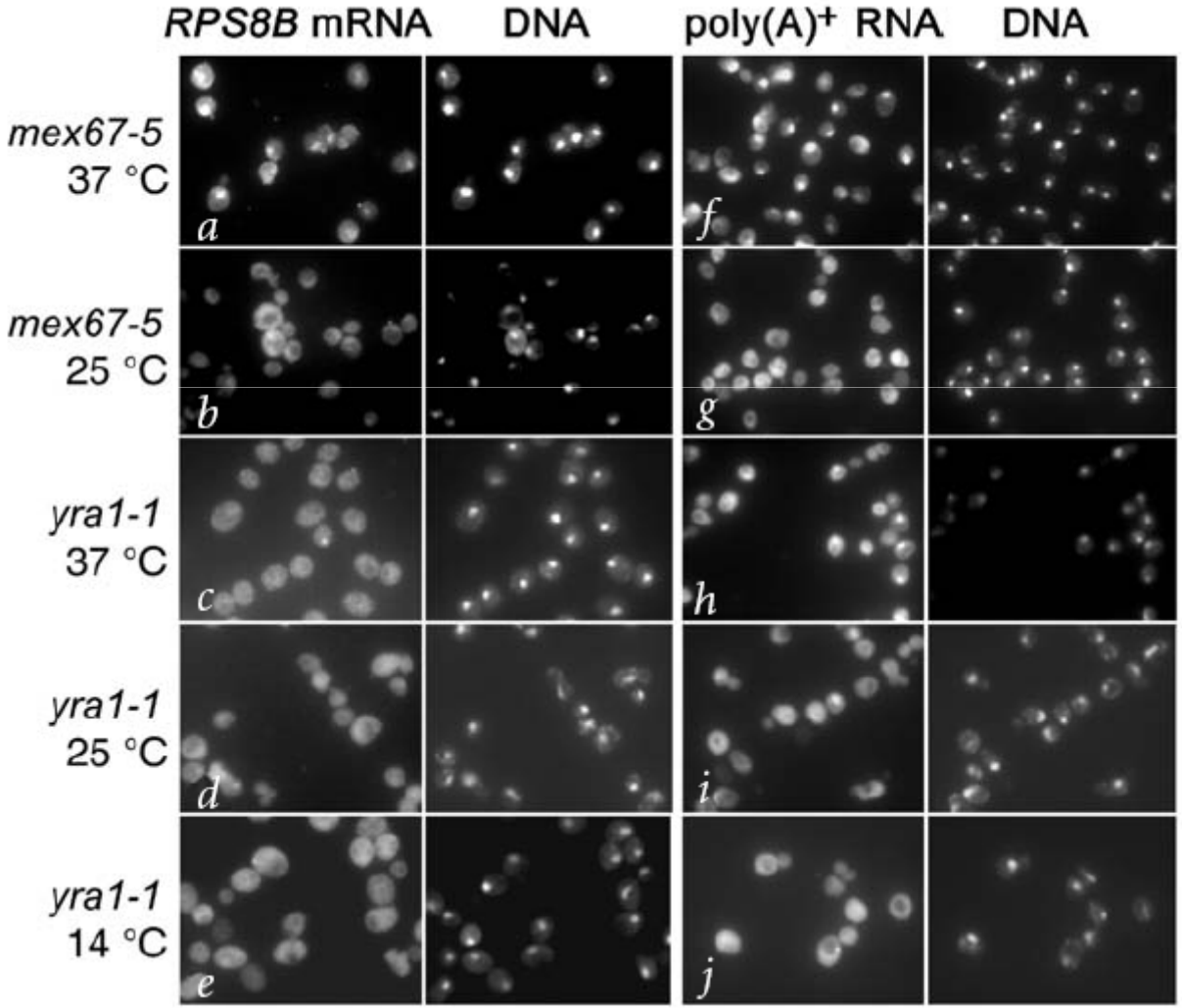




From: Mata et al.(2005) Trends Biochem Sci 30:506-514.




mRNA that are differentially found in complex with Mex67 or Yra1.

Example of validation of results. One of these mRNA is RPS8B, which is immunoprecipitated with Mex67, but not with Yra1.
Temperature-sensitive mutants of Mex67 and Yra1 demostrate that Mex67, but not Yra1, is necessary for the export of this mRNA.



To identify all the factors involved in mRNA nuclear to cytoplasmatic transport, researchers have used genome-wide RNAi screens

In this technique, double-stranded interfering RNAs (siRNA) directed against all known protein-encoding transcripts are individually transfected in cells.

To increase througput, reverse transfection is used: RNAi is mixed wih gelatin and delivered to the wells of 384-wells culture plates. Cells are plated on the gelatin in these plates together with lipofection reagents. After several hours, cell phenotype is measured with appropriate techniques.

For this particular project, cells were hybridized to fluorescent oligo-dT probes, which will reveal the localization of all poly(A+)-containing RNAs .

# Definition of global and transcript-specific mRNA export pathways in metazoans

Natalie G. Farny,[1] Jessica A. Hurt,[1] and Pamela A. Silver[2]

Department of Systems Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

**Eukaryotic gene expression requires export of messenger RNAs (mRNAs) from their site of transcription in the nucleus to the cytoplasm where they are translated. While mRNA export has been studied in yeast, the complexity of gene structure and cellular function in metazoan cells has likely led to increased diversification of these organisms' export pathways. Here we report the results of a genome-wide RNAi screen in which we identify 72 factors required for polyadenylated [poly-($A^+$)] mRNA export from the nucleus in *Drosophila* cells. Using structural and functional conservation analysis of yeast and *Drosophila* mRNA export factors, we expose the evolutionary divergence of eukaryotic mRNA export pathways. Additionally, we demonstrate the differential export requirements of two endogenous heat-inducible transcripts—intronless heat-shock protein 70 (HSP70) and intron-containing HSP83—and identify novel export factors that participate in HSP83 mRNA splicing. We characterize several novel factors and demonstrate their participation in interactions with known components of the *Drosophila* export machinery. One of these factors, *Drosophila melanogaster* PCI domain-containing protein 2 (dmPCID2), associates with polysomes and may bridge the transition between exported messenger ribonucleoprotein particles (mRNPs) and polysomes. Our results define the global network of factors involved in *Drosophila* mRNA export, reveal specificity in the export requirements of different transcripts, and expose new avenues for future work in mRNA export.**

Genes and Development, 2008

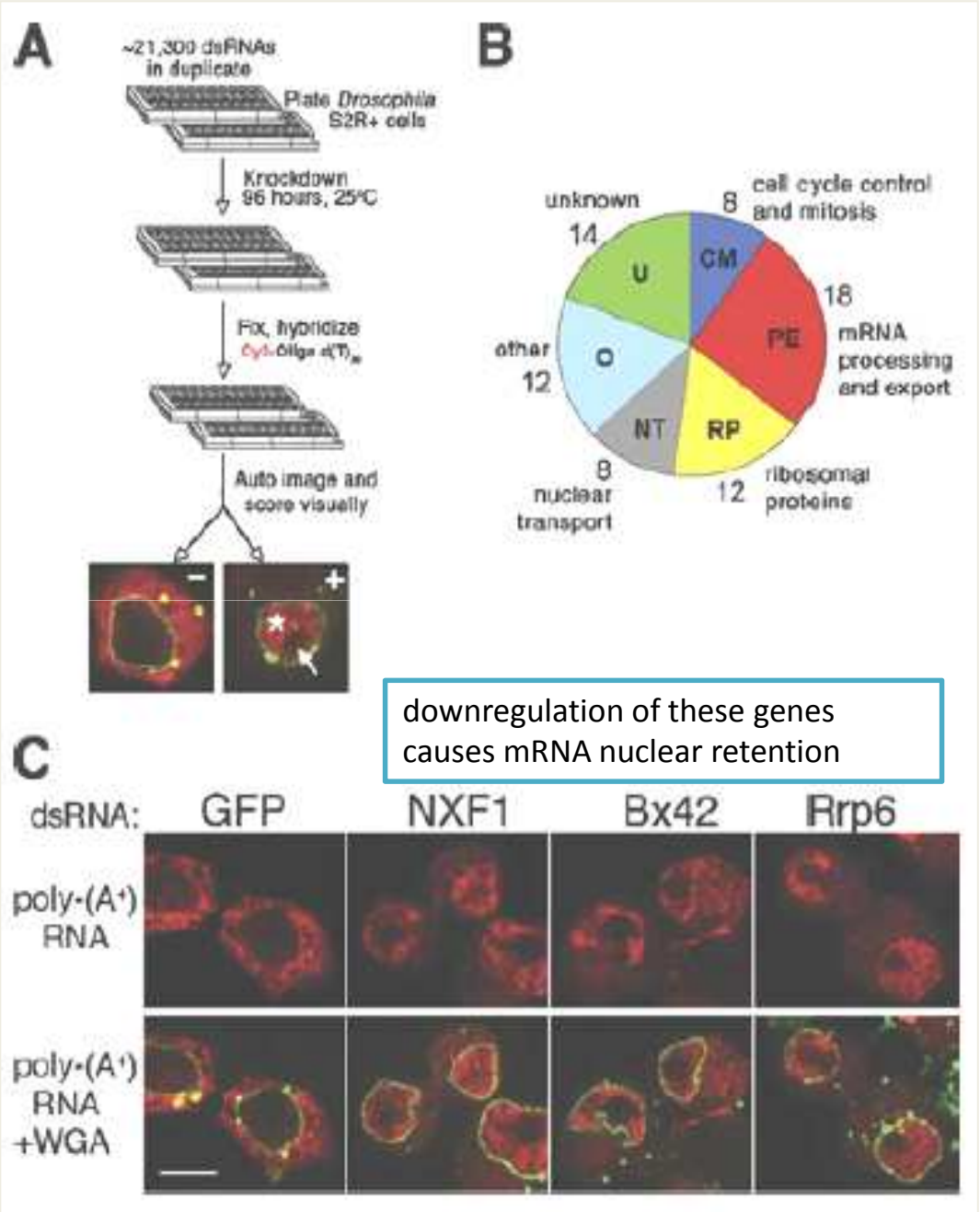



**Figure 1.** Identification and classification of genes affecting mRNA export in *Drosophila* S2R+ cells. (*A*) A library of 21,300 dsRNAs was arrayed in black, clear-bottomed 384-well plates. S2R+ cells were plated on these dsRNAs, incubated for 4 d, then fixed and assayed for localization of poly-(A+) RNA using a Cy3-labeled Oligo-d(*T*)30 probe. Images were collected by automated microscopy and visually inspected for nuclear accumulation of poly-(A+) RNA. Images were scored negative (*left* image) or positive (*right* image) for nuclear poly-(A+) RNA accumulation. In positive cells, mRNA accumulates in the nucleus (asterisk) and is excluded from the nucleolus (arrow). (*B*) Gene ontology categories of the 72 mRNA export factors identified in this screen.
(*C*) The cells shown were treated with dsRNAs for 4 d, targeting the genes indicated. After fixation and hybridization, poly-(A+) RNA localization (red, *top* panels) was determined by overlaying images with the nuclear membrane marker, wheat germ agglutinin (WGA) (green, *bottom* images). Bar, 4 µm.

# mRNA localization

A number of mRNAs are not translated immediately when they enter the cytoplasmatic compartment, but are redistributed to specific subcellular localizations before they can be translated.

This phenomenon in some cases depends on direct active transport of specific mRNA, but in other cases relies on other mechanisms, such as trapping, or local protection against diffused degradation, or local degradation of a translational inhibitor.

REVIEW

Review

# Subcellular mRNA Localization in Animal Cells and Why It Matters

Christine E. Holt[1] and Simon L. Bullock[2]*

Subcellular localization of messenger RNAs (mRNAs) can give precise control over where protein products are synthesized and operate. However, just 10 years ago many in the broader cell biology community would have considered this a specialized mechanism restricted to a very small fraction of transcripts. Since then, it has become clear that subcellular targeting of mRNAs is prevalent, and there is mounting evidence for central roles for this process in many cellular events. Here, we review current knowledge of the mechanisms and functions of mRNA localization in animal cells.

Alcuni mRNA vengono localizzati in posizioni specifiche della cellula.

Quanti?

circa il 10% degli mRNA di origine materna negli oociti di Drosophila

circa 400 mRNA vengono localizzati nei dendriti/assoni dei neuroni, nei mammiferi

Recent estimates brought to up to 70% of early drosophila mRNAs showing some degree of preferential oocyte localization

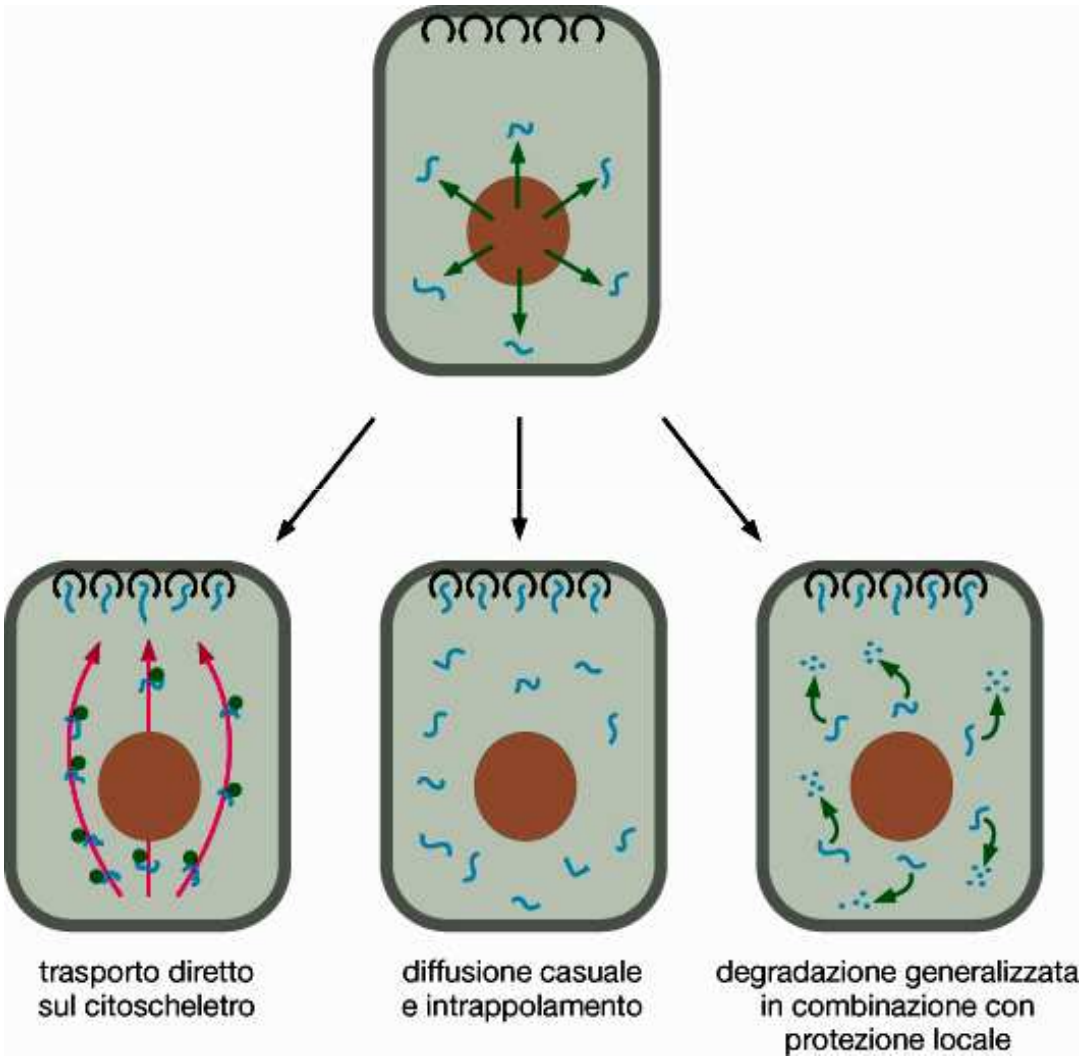
Some genes have localization signals in both mRNA and protein (sometimes redundant)

Why to localize mRNA instead of proteins?

1) to localize proteins correctly but preventing their presence elsewhere
2) regulate gene expression differentially in different cell localizations
3) immediateness of specific local responses

# Possibili meccanismi di localizzazione dell'RNA

trasporto vettoriale dal nucleo



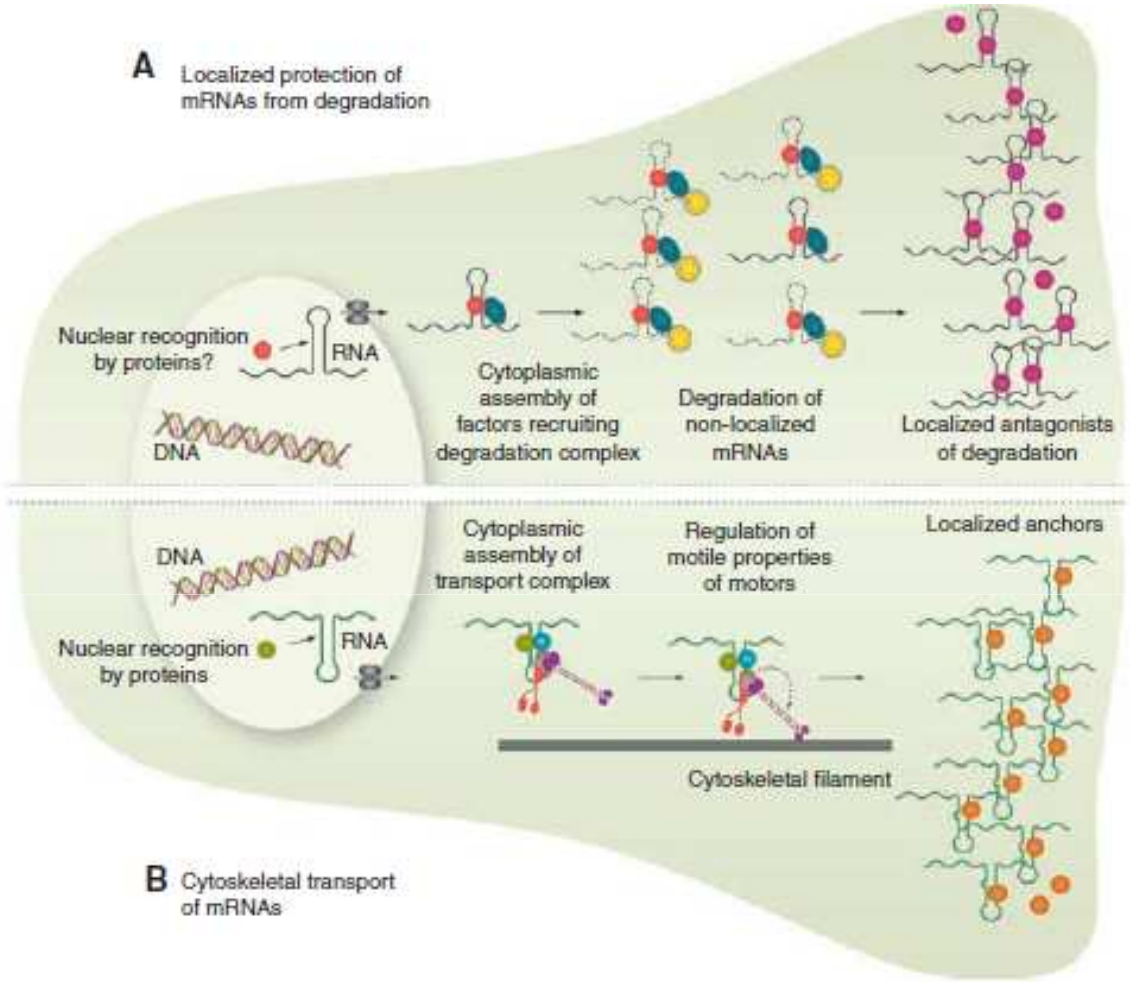



**Fig. 1.** mRNA localization is a multi-step process. Shown is an illustration of two stylized cells, depicting mechanisms that can contribute to mRNA localization. (**A**) Protection of mRNAs from degradation. Red, nuclear RNA recognition factor; dark blue, cytoplasmic RNA recognition factor; yellow, ribonuclease; purple, agonist of degradation. (**B**) Motor-based transport. Green, nuclear RNA recognition factor; light blue and light gray, cytoplasmic RNA recognition factor; red and purple, molecular motors; orange, anchorage factor. In reality, different combinations of these mechanisms may be used to localize a single mRNA species in the same cell.

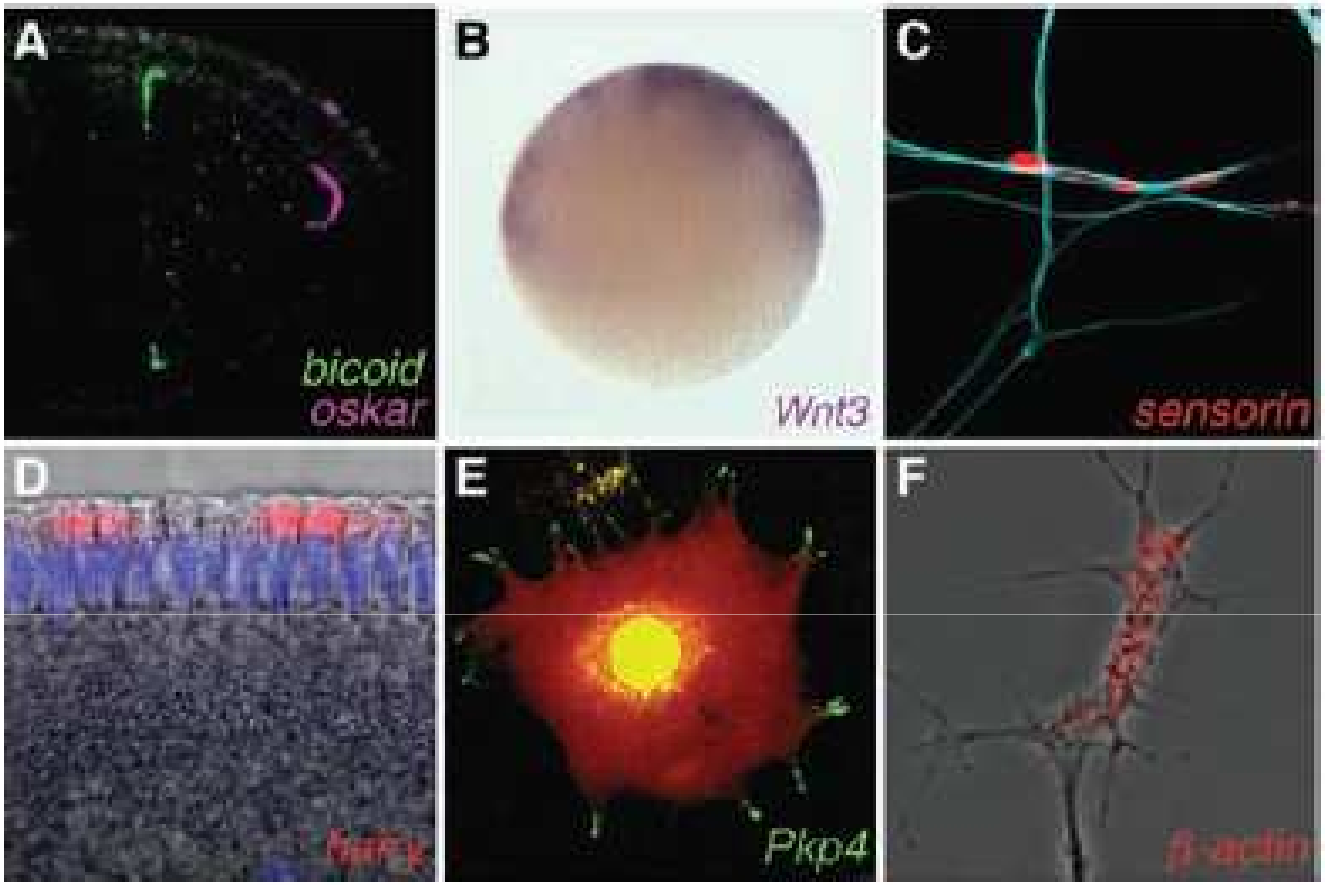
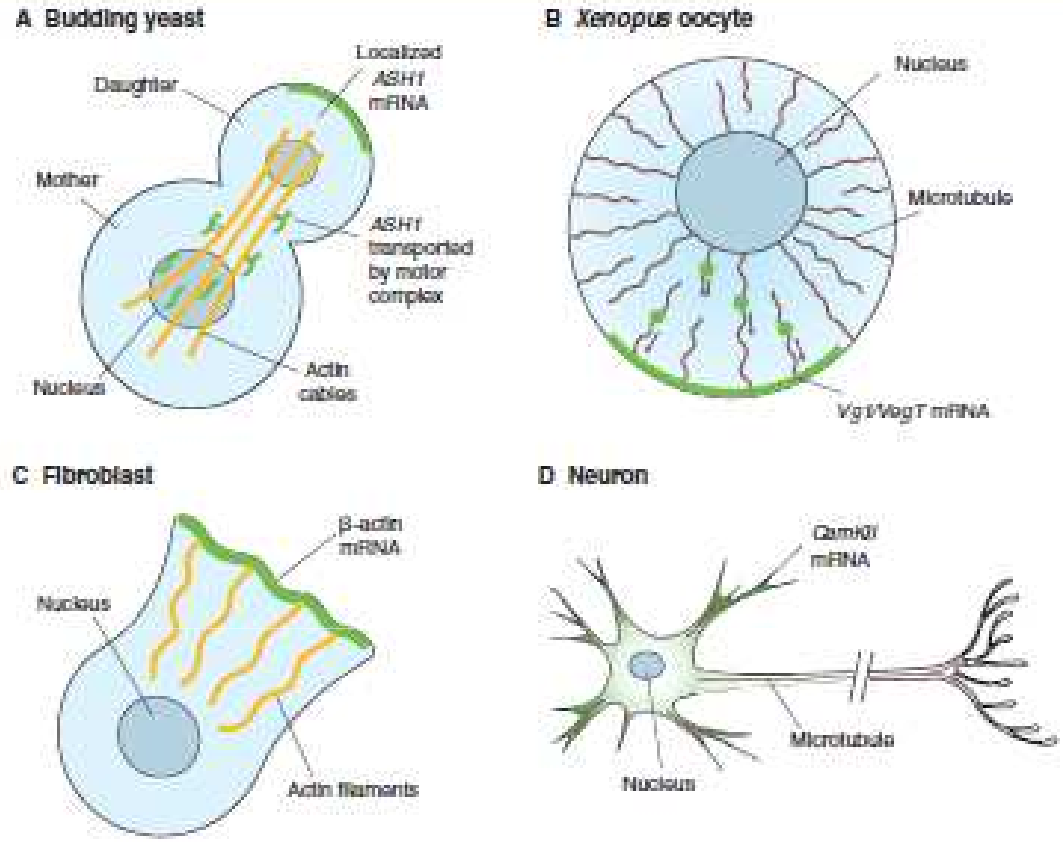



Fig. 2. Examples of asymmetrically localized mRNAs. (A) Differential localization of mRNA determinants within the Drosophila oocyte. (B) Animal localization of a transcript encoding a signalingmolecule required for axis development in the egg of a cnidarian, Clytia. (C) mRNA enrichment in synapses of an Aplysia sensory neuron in response to contact with a target motor neuron (blue). (D) Apical localization of an mRNA in the Drosophila embryo, which facilitates entry of its transcription factor product into the nuclei (purple). (E) mRNA localization in pseudopodial protrusions of a cultured mammalian fibroblast (red signal indicates the cell volume). (F) mRNA enrichment within a Xenopus axonal growth cone. mRNAs were visualized by means of in situ hybridization except in (E), in which the MS2–green fluorescent protein (GFP) system was used.

**Box 1. Key examples of localized mRNAs**




In budding yeast, localization of *ASH1* mRNA to the bud tip by myosin-mediated transport on actin cables (see A in figure) targets Ash1p to the daughter cell, where it is required to repress mating type switching (reviewed by Gonsalvez et al., 2005). Thus, mating type switching occurs only in the mother cell, thereby ensuring that both yeast mating types are present in the population. Germ layer specification in *Xenopus* embryos relies on localized mRNAs, including *Vg1* and *VegT* (reviewed by King et al., 2005) (see B in figure), which are transported to the vegetal pole of the oocyte by kinesin motors and anchored to the cortex by an actin-dependent mechanism. After fertilization, *Vg1* and *VegT* RNAs are inherited by the embryonic vegetal cells, where Vg1 protein, a TGFβ homolog, participates in mesoderm specification, while VegT, a transcription factor, regulates endoderm specification and mesoderm induction. mRNA localization plays an important role in the polarization of somatic cells, such as fibroblasts and neurons (see C,D in figure). Localization of β-actin mRNA to the leading edge of migrating fibroblasts provides a high local concentration of actin monomers that drives assembly of the actin filaments needed for forward movement. Similarly, β-actin mRNA localization to growth cones in developing axons promotes the motility required for axon guidance (reviewed by Condeelis and Singer, 2005). Dendritic localization of RNAs like calcium/calmodulin-dependent protein kinase IIα (CaMKIIα) mRNA in hippocampal neurons facilitates a rapid response to synaptic activity in the form of local protein translation and contributes to learning and memory-related synaptic plasticity (reviewed by Martin et al., 2000).

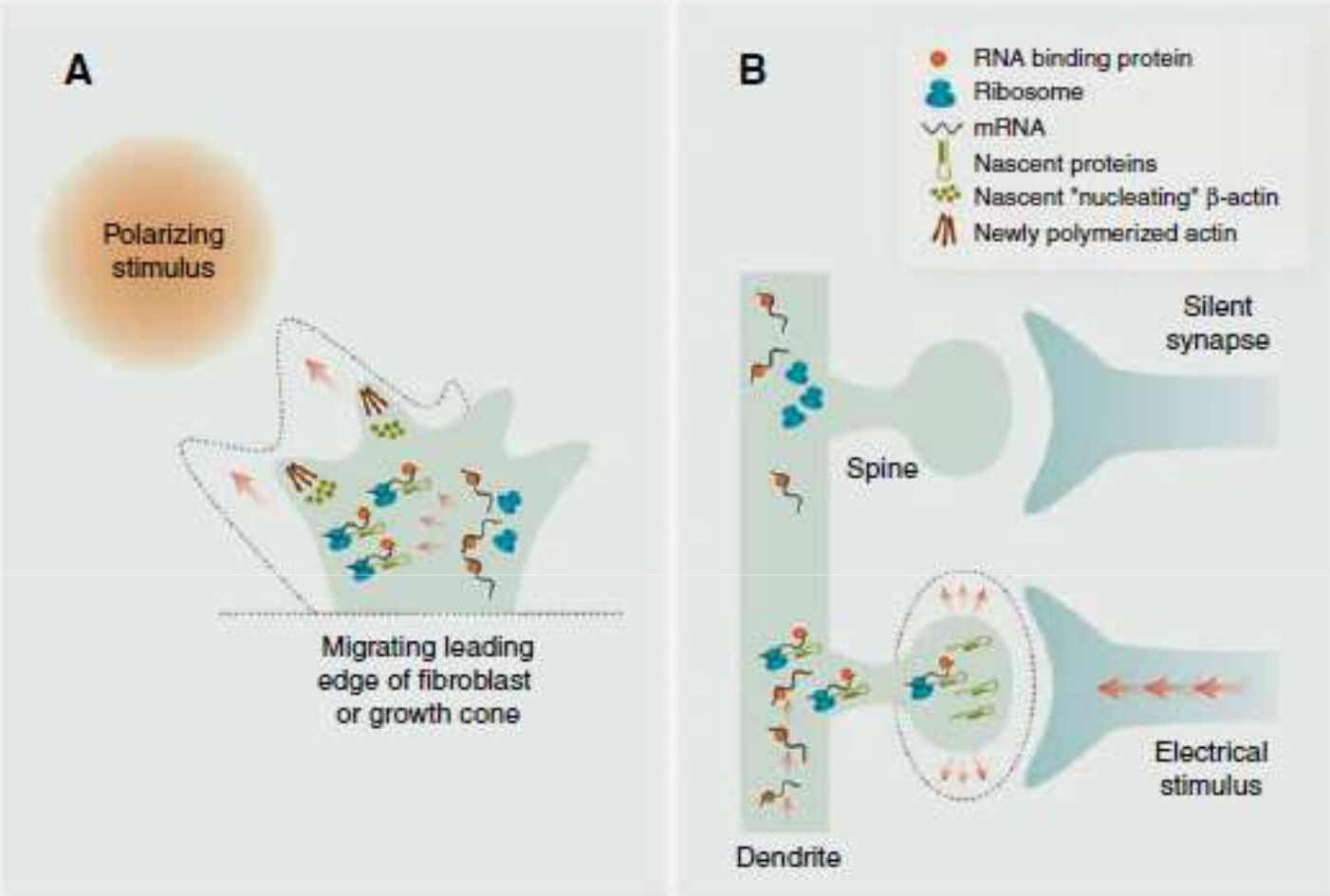
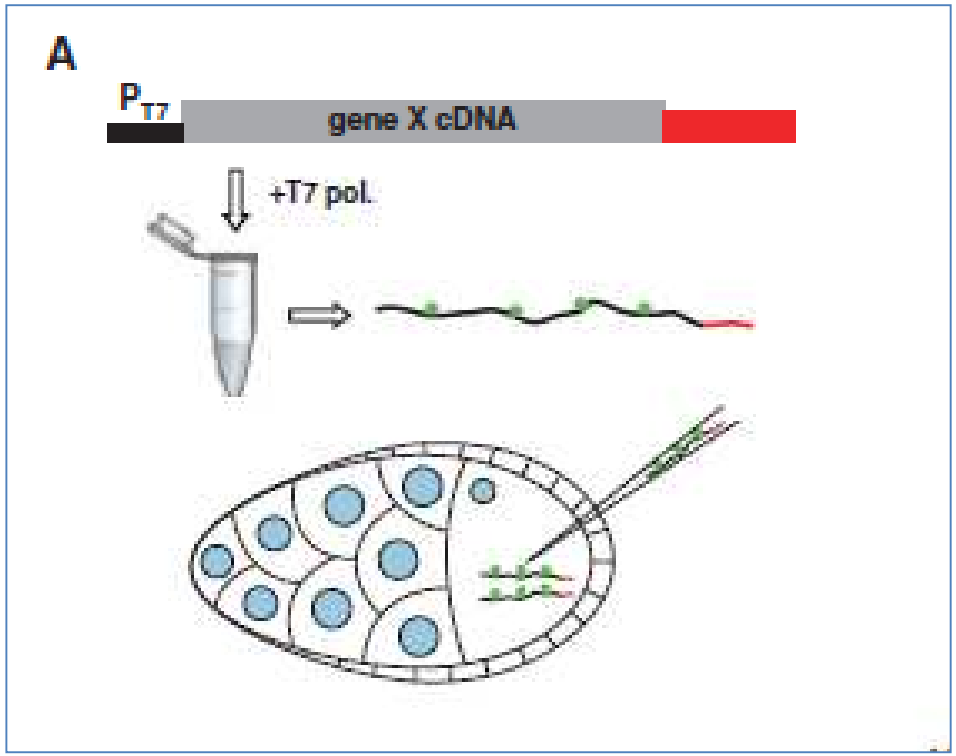



**Fig. 3.** Extrinsic stimuli elicit changes in subcellular mRNA localization and translation. (**A**) A polarizing stimulus elicits asymmetric localization and translation of mRNAs encoding β-actin and actin regulators on the near-stimulus side of the leading edge of migrating cells, such as fibroblasts and axonal growth cones, thus contributing to polarized cell movement and directional steering. The dashed outline denotes the post-stimulus trajectory. (**B**) Electrical input from presynaptic contacts selectively induces localized trafficking and translation of specific mRNAs in dendrites that mediate changes in spine morphology (dashed outline) and plasticity. Several aspects of these models are speculative.

To study specfic mRNA localization, several techniques have been used.

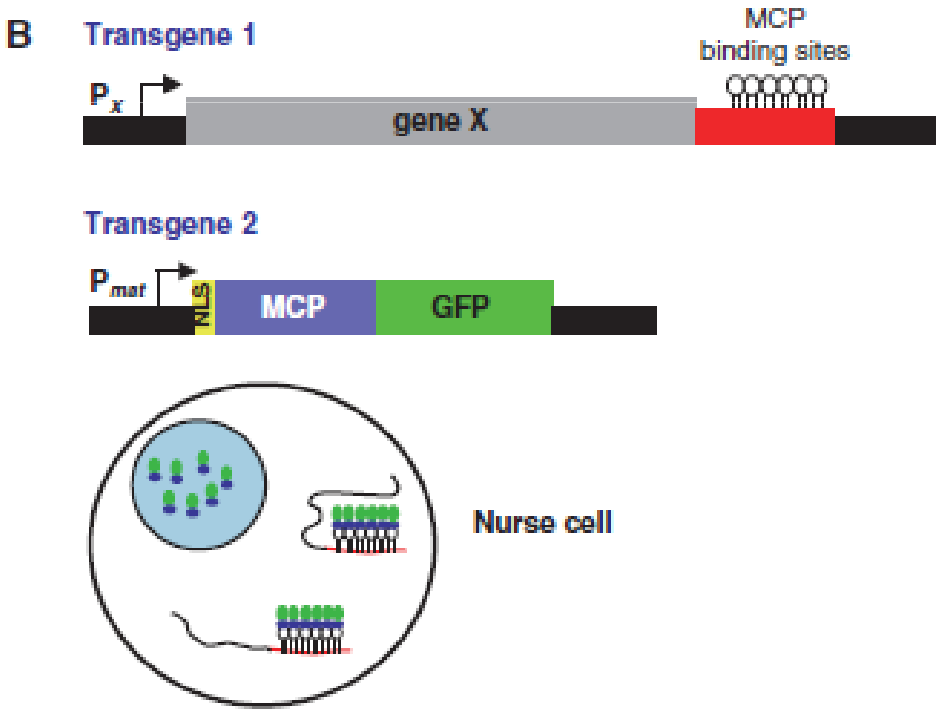
One method is to transcribe the RNA  “in vitro” labellig with fluorochromes.

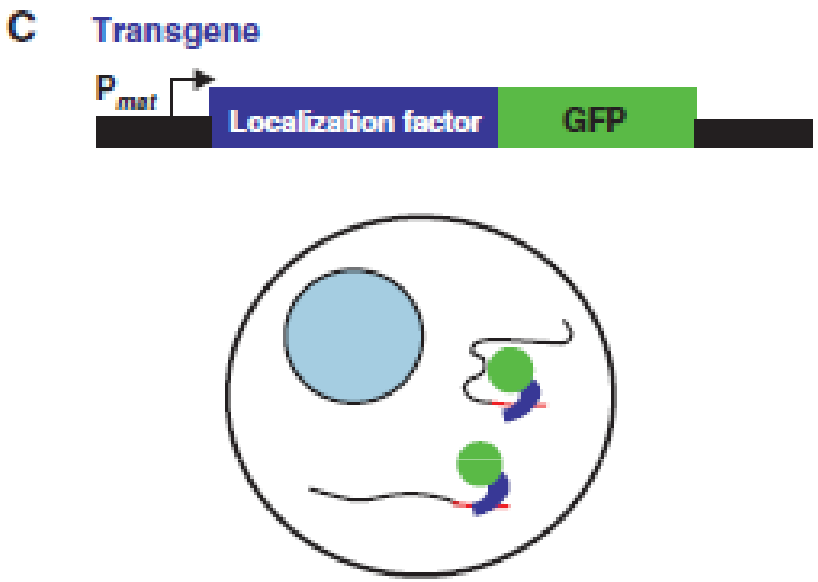
The RNA is then injected in cells or embryos.



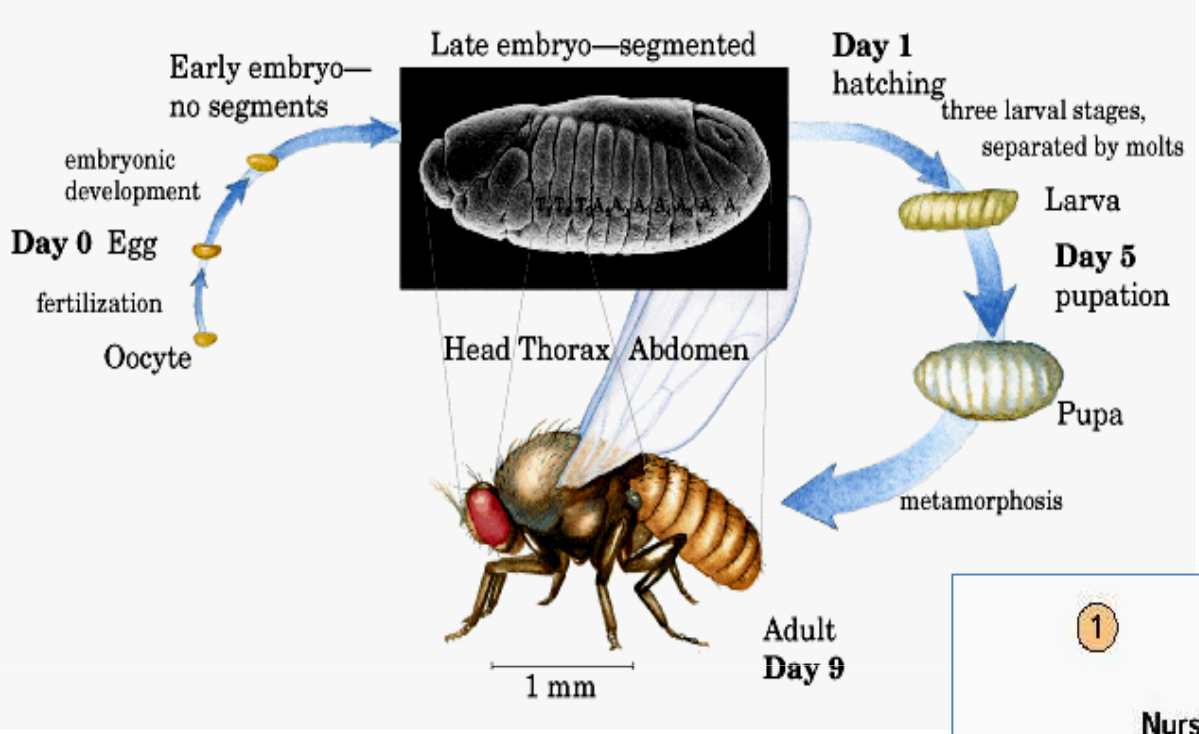

This approach works only if the factor that localizes the RNA does not require to be charged on the RNA during transcription.

Another method is to transfect the cells with a minigene fusing a repeated hairpin –forming sequence to our mRNA of interest, as for example the MCP operator.
The cells are transfected at the same time with a vectro encoding the MCP-binding domain fused to GFP.
GFP fluorescence will reveal the localization of our gene mRNA.

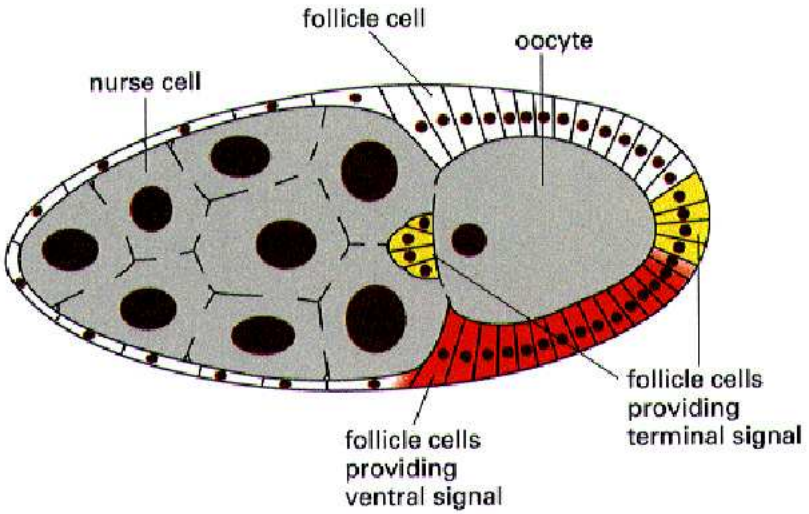


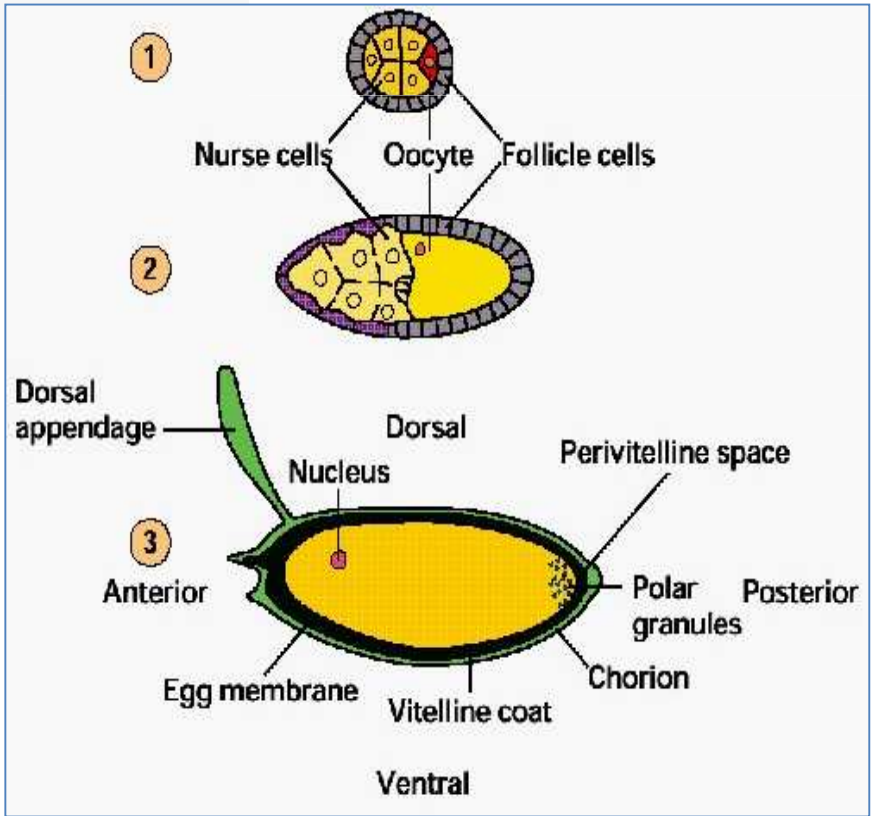


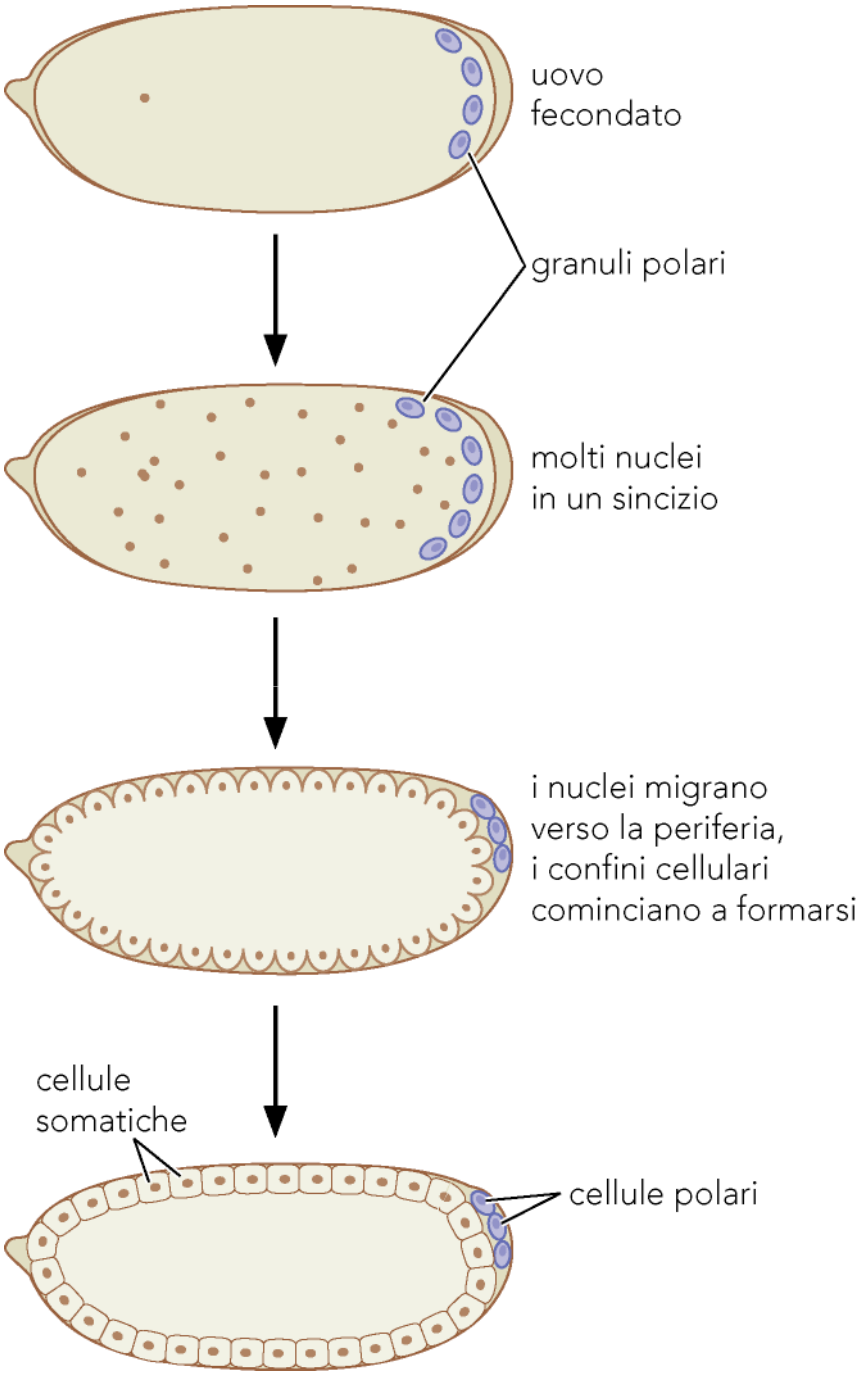

Finally, if the “localizator” is known, one can use a GFP fusion of this factor.

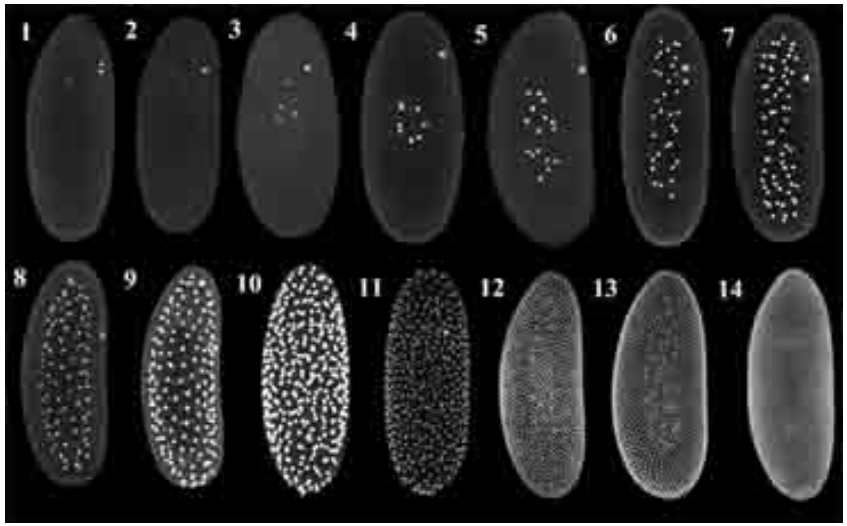



Most of our knowledge of mRNA localization comes from Drosophila early development.






Short after fertilization, nuclear divisions produce a syncithial blastocist, followed by migration of nuclei at the periphery and only later cellularization. Nuclei are then exposed to gradients of proteins and mRNAs that are present in the oocyte.

Maternal mRNA and proteins are produced by nurse cells and follicle cells. Nurse cells shrink and pass their RNAs to the oocyte, where different mechanisms of mRNA take place, resulting in the differential localization of more than 70% of the oocyte mRNAs.

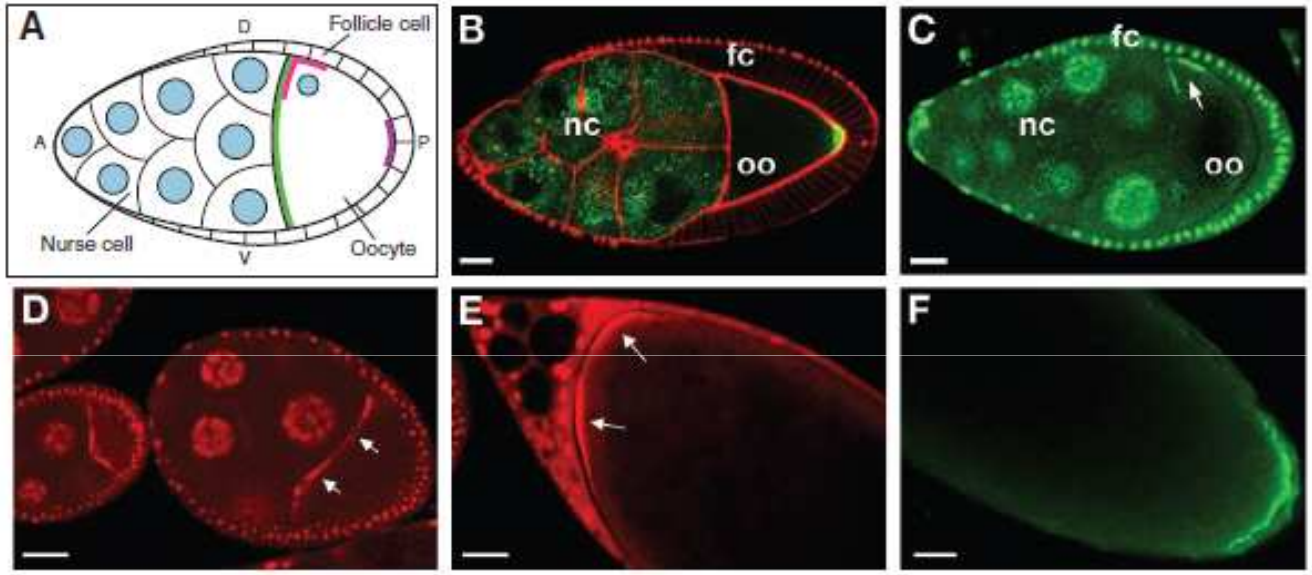
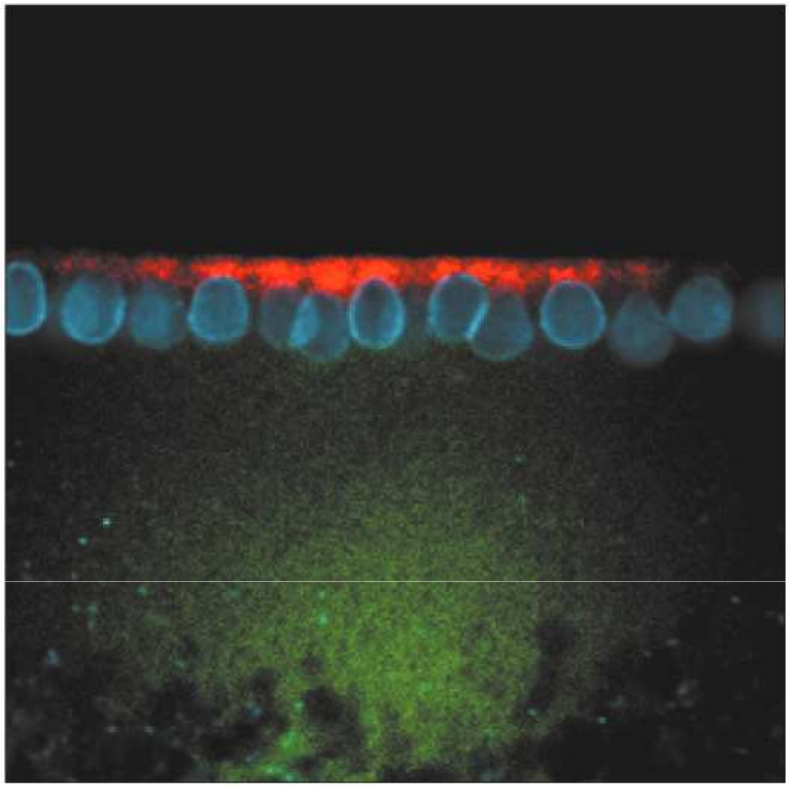



**Fig. 1. Localized distributions of *grk*, *bcd*, *osk* and *nos* mRNAs.** (**A**) Schematic showing *grk* (pink), *bcd* (green) and *osk* (purple) mRNA localization in mid-oogenesis (stage 9). *nos* mRNA is not yet localized at this stage. The anteroposterior (AP) and dorsoventral (DV) axes are indicated. (**B**) GFP-Stau (green), as proxy for *osk* mRNA, at the posterior pole of the oocyte (oo) during mid-oogenesis. GFP-Stau is also detected in the nurse cell (nc) cytoplasm. The actin cytoskeleton is highlighted in red with phalloidin. fc, follicle cells. Orientation is the same as in A. (**C-F**) Visualization of endogenous mRNAs using the MS2 system: (C) *grk* and (D) *bcd* during mid-oogenesis; (E) *bcd* and (F) *nos* in late oocytes. Owing to the promoter used, the MCP-GFP and MCP-RFP fusion proteins are expressed in both the nurse cells and follicle cells, whereas the MS2-tagged mRNAs are produced only in the nurse cells. MCP-GFP/RFP that is not bound to mRNA enters both the nurse cell and follicle cell nuclei. Scale bars: 20 µm. Image in B was modified, with permission, from Huynh et al. (Huynh et al., 2004); image in C was modified, with permission, from Jaramillo et al. (Jaramillo et al., 2008); images in D and E are reproduced, with permission, from Weil et al. (Weil et al., 2006). Image in F is courtesy of K. Sinsimer (Princeton University, Princeton, NJ, USA). *bcd*, *bicoid*; *grk*, *gurken*; GFP, green fluorescent protein; MCP, MS2 coat protein; *nos*, *nanos*; *osk*, *oskar*; RFP, red fluorescent protein.



*Hairy* mRNA localization in syncithial Drosophila blastocyst, as determined by microinjection of fluorochrome-labeled mRNAs.

**Red** = wild-type *hairy* mRNA

**Green** = mRNA from a mRNA lacking 3'-UTR

In most cases, *cis*-acting elements in RNA are found in 3'-UTR

(cases of elements within the coding sequence are known, however)

This is interesting because at least 50% of alternative 3'-UTR are tissue-specific:.

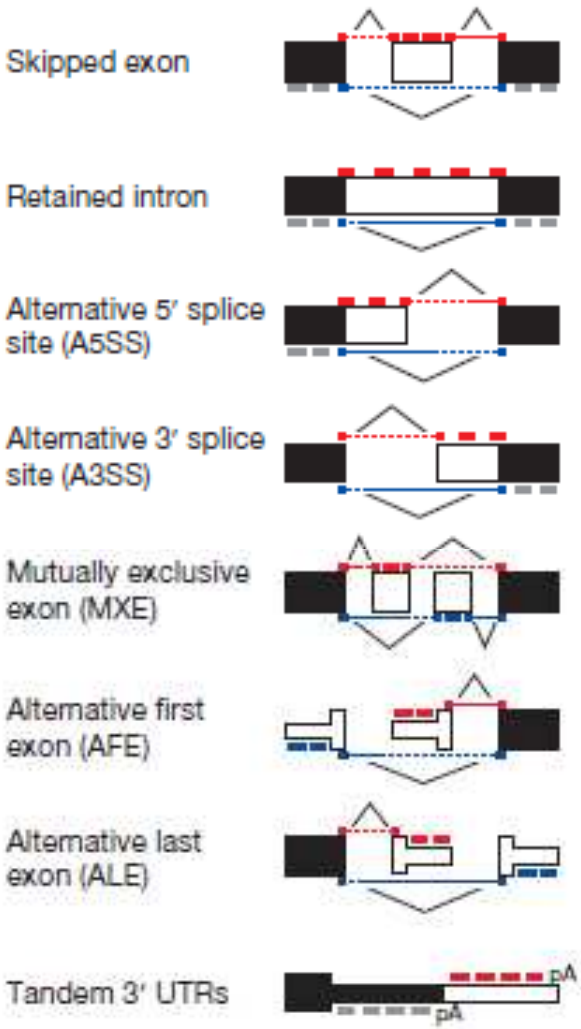
Large number of AS events involving the 3' UTR

- Alternative Poly(A) site
- Alternative last exon
- Introns within 3' UTR

This may imply that differential regulation of RNA fate (localization, stability) is an important option for many genes that is controlled during RNA maturation.

From Wang's RNA Seq study

| Alternative transcript events | Total events ($\times 10^3$) | Number detected ($\times 10^3$) | Both isoforms detected | Number tissue-regulated | % Tissue-regulated (observed) | % Tissue-regulated (estimated) |
|---|---|---|---|---|---|---|
| Skipped exon | 37 | 35 | 10,436 | 6,822 | 65 | 72 |
| Retained intron | 1 | 1 | 167 | 96 | 57 | 71 |
| Alternative 5′ splice site (A5SS) | 15 | 15 | 2,168 | 1,386 | 64 | 72 |
| Alternative 3′ splice site (A3SS) | 17 | 16 | 4,181 | 2,655 | 64 | 74 |
| Mutually exclusive exon (MXE) | 4 | 4 | 167 | 95 | 57 | 66 |
| Alternative first exon (AFE) | 14 | 13 | 10,281 | 5,311 | 52 | 63 |
| Alternative last exon (ALE) | 9 | 8 | 5,246 | 2,491 | 47 | 52 |
| Tandem 3′ UTRs | 7 | 7 | 5,136 | 3,801 | 74 | 80 |
| Total | 105 | 100 | 37,782 | 22,657 | 60 | 68 |



How is localization regulated?

RNA localization depends on *cis*-acting elements (usually –but not exclusively- within the 3’ UTR)

AND

*trans*-acting factors:

- RNA-binding proteins
- Vescicles
- Microtubules

OR

localized protection / degradation mechanisms

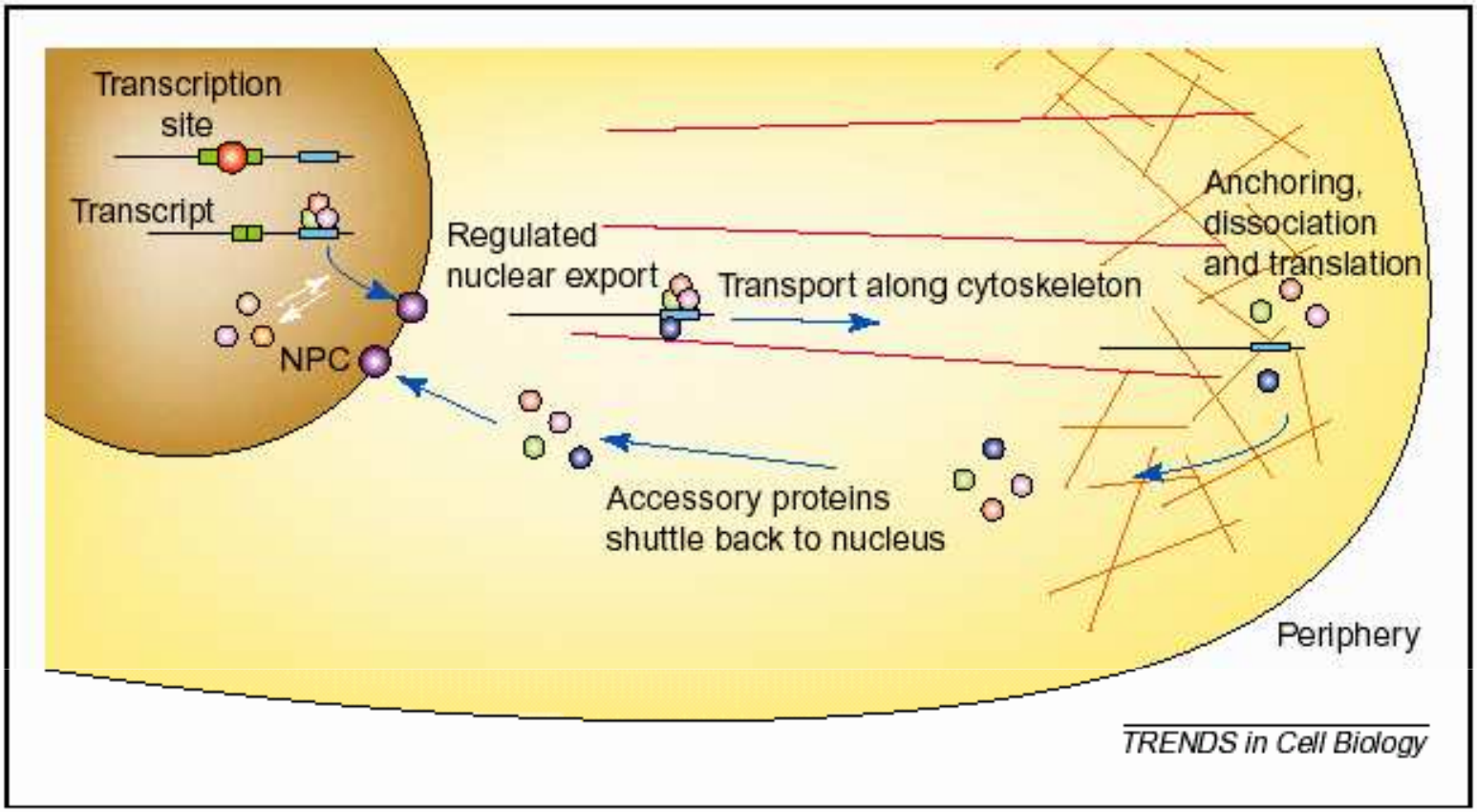



**Fig. 2.** Model of localization factors shuttling between the nucleus and the cytoplasm. Factors associate with the RNA at the transcription site to form a localization complex or 'locasome' - a particle that is designated to localize associated RNAs, by riding with the RNA through the nuclear pore complex (NPC) where additional factors join and assemble motors. The locasome then travels to the cell periphery directed by the polarity of the cytoskeletal filaments and anchors on the isotropic microfilament network, where it releases the shuttling factors and becomes competent for translation. The shuttling factors then return to the nucleus. The cell periphery might be the growth cone of a neuron, the lamellipod of a fibroblast, the vegetal pole of a *Xenopus* oocyte or the posterior pole of a *Drosophila* embryo.

from: Farina & Singer 2002, Trends in Cell Biol, 12: 466.

Protein factors bind to sequences present in 3’-UTR and mediate translational inhibition and other mechanisms of localization.

As illustrated in the next slide for example

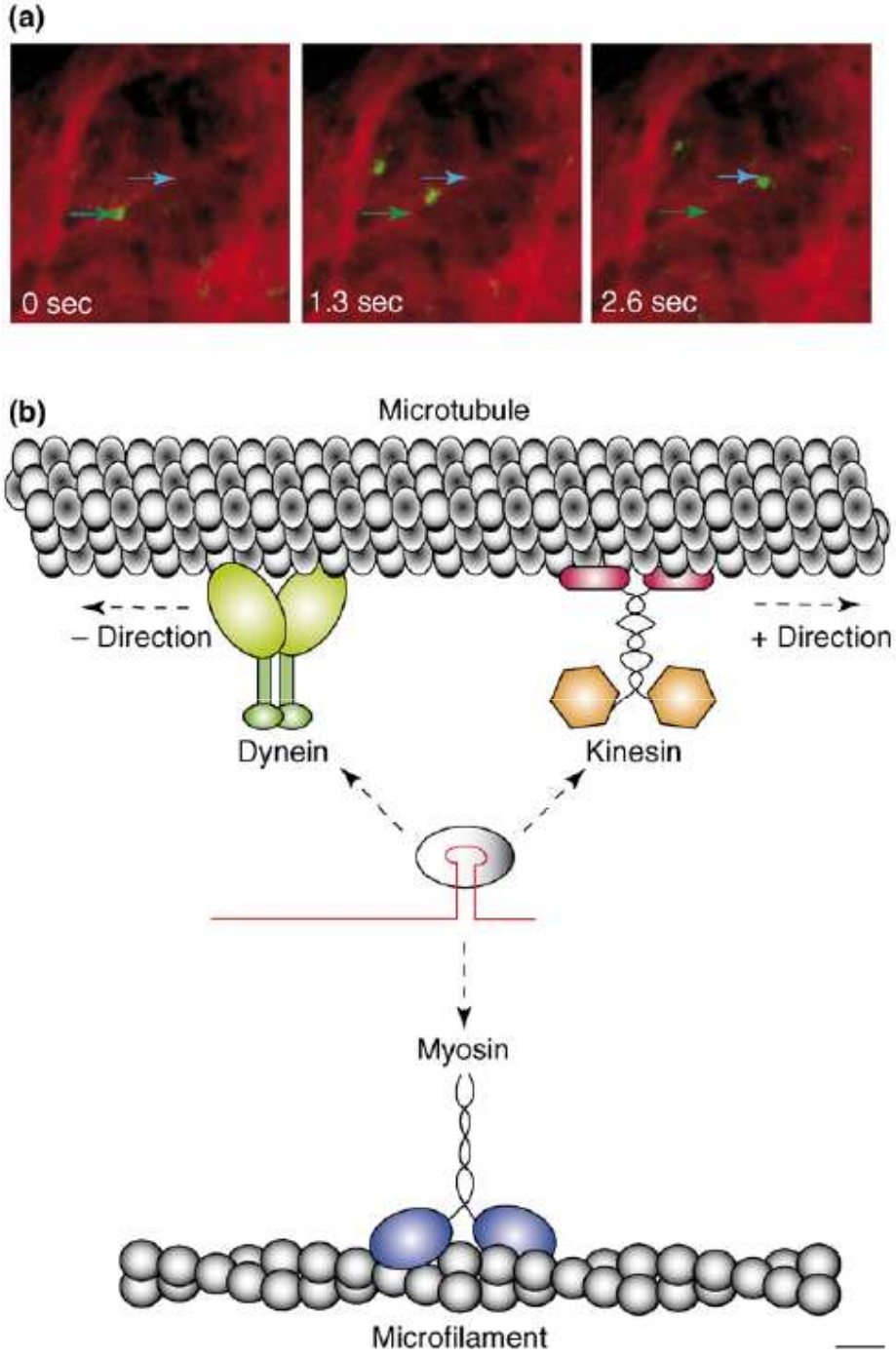



Figure 4. Interaction with cytoskeletal motors during mRNA localization. (a) A LEcontaining mRNA labeled in vivo with a GFP tag shows tracking along a microtubule in time-lapsed images. Red, tubulin; green arrow, starting position; blue arrow, ending position [70]. (b) The interactions of L-RNPs with microtubule
motor complexes or microfilament motor complexes have been implicated in localization. All three types of cytoskeleton-dependent motors (kinesin, and dynein for microtubules and myosin for microfilaments) have been suggested to have roles in either directed movement or local anchoring. Based on the identified examples of L-RNP–motor protein complexes, localizing mRNA is depicted here as interacting with motors indirectly through LE-binding proteins as a RNP complex (represented for simplicity as a gray oval bound to a localizing mRNA, but can in fact be extremely large with multiple RNA-binding proteins and mRNAs). In metazoans, the molecular details of these interactions are not known.

Several localization factors have been identified to date. Quite often, these are proteins known to play also different roles in RNA metabolism.

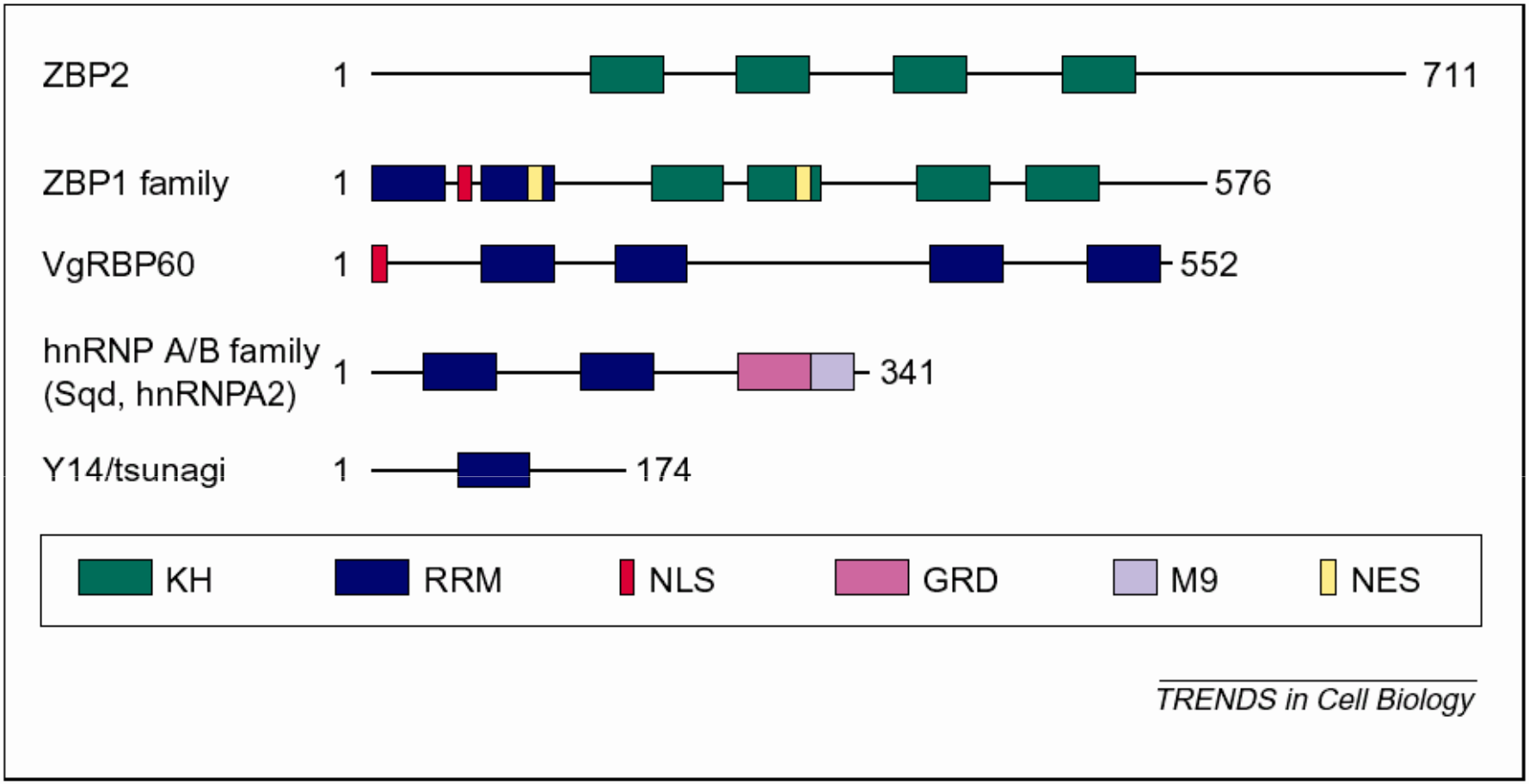



**Fig. 1.** Domain structures of RNA-binding proteins with a nuclear role in cytoplasmic RNA localization. The hnRNP K homology (KH) domain is coloured green, the RNA recognition motif (RRM) dark blue, the nuclear localization signal (NLS) red, the glycine-rich domain (GRD) purple, the M9 nucleocytoplasmic shuttling domain light blue, and the putative nuclear export signal yellow.

from: Farina & Singer 2002, Trends in Cell Biol, 12: 466.

**Table 1. Properties of *trans*-acting factors associated with localized mRNAs in the nucleus and the cytoplasm[a]**

| Protein | Organism | Cell type | mRNA | Homologues | Associated functions |
|---|---|---|---|---|---|
| **hnRNP A/B family** | | | | | |
| HnRNPA2 | Human | Oligodendrocyte | MBP | | Splicing |
| Sqd (hrp40) | *Drosophila* | Oocyte | *gurken* | Human hnRNPA1, *S. cerevisiae* Np13, Hrp1, *C. tentans* hrp36 | mRNA processing and transport, NMD |
| **ZBP1 family** | | | | | |
| ZBP1 | Chicken | Fibroblast, neuron | β-Actin | CRD-BP, IMP1, IMP2, IMP3, KOC, HCC, dIMP | mRNA stability, translational regulation, overexpressed in cancer |
| Vg1RBP | *Xenopus* | Oocyte | *Vg1, Veg T* | See above | See above |
| ZBP2 | Chicken | Fibroblast, neuron | β-Actin | KSRP, MARTA | Splicing |
| She2p | *S. cerevisiae* | Budding yeast | *ASH1* | NK | NK |
| Loc1p | *S. cerevisiae* | Budding yeast | *ASH1* | NK | NK |
| Tsunagi | *Drosophila* | Oocyte | *Oskar* | Human Y14 | EJC, NMD |

[a]Abbreviations: EJC, exon-exon junction complex; NK, not known; NMD, nonsense-mediated mRNA decay.

from: Farina & Singer 2002, Trends in Cell Biol, 12: 466.

Example 1

The *trans*-acting factor **hnRNPA2** is involved in the localization of myelin basic protein (MBP) mRNA to the myelin-forming processes of mammalian oligodendroctyes.

Like many of the proteins involved in RNA localization, hnRNPA2 has several other functions, including splicing, nuclear export, translational regulation and RNA stabilization.

As a *trans*-acting factor of MBPmRNA localization, hnRNPA2 binds to a 21-nucleotide (nt) *cis* element called the hnRNPA2 response element (A2RE), formerly known as the 'RNA trafficking sequence' (RTS), that is located in the 3' untranslated region (UTR) of the mRNA.

Although it is a predominantly nuclear protein, hnRNPA2 shuttles to the cytoplasm, by means of its M9 nucleocytoplasmic shuttling domain, where it localizes in cytoplasmic granules [16].

**Example 2**

**Sqd** localizes gurken mRNA

The *Drosophila* protein Sqd (hs hnRNPA1) is an RNA-binding protein of 42 kDa that is needed for the proper localization of *gurken* (*grk*) transcripts during oogenesis. Like hnRNPA2, Sqd is a member of a class of the hnRNPs that shuttle between the nucleus and the cytoplasm through an M9 shuttling motif.

Dorsoventral patterning in *Drosophila* requires localization of *grk* mRNAto the dorsoanterior corner of the oocyte. In *sqd* mutants, *grk* mRNA is mislocalized at the anterior of the oocyte, and does not accumulate anterodorsally.

Sqd protein is thought to associate with *grk* mRNA in the nucleus and to deliver it to cytoplasmic anchors at the dorsoanterior of the oocyte.

Sqd is known to bind directly to the 3′ UTR of *grk* mRNA.

Elements that mediate localization are called ZIP codes.

As very often with RNA, these elements are particulary difficult to identify, since very often they consist of short sequence motifs (4-8 nt) repeated and dispersed within fragment of hundred bp sizes.

In addition, very often more than one RNA binding protein is involved in the recognition of the LE (localizing element) and are required to mediate localization.

This is quite well illustrated by the Vg1 mRNA localizing element, illustrated in the next slide.

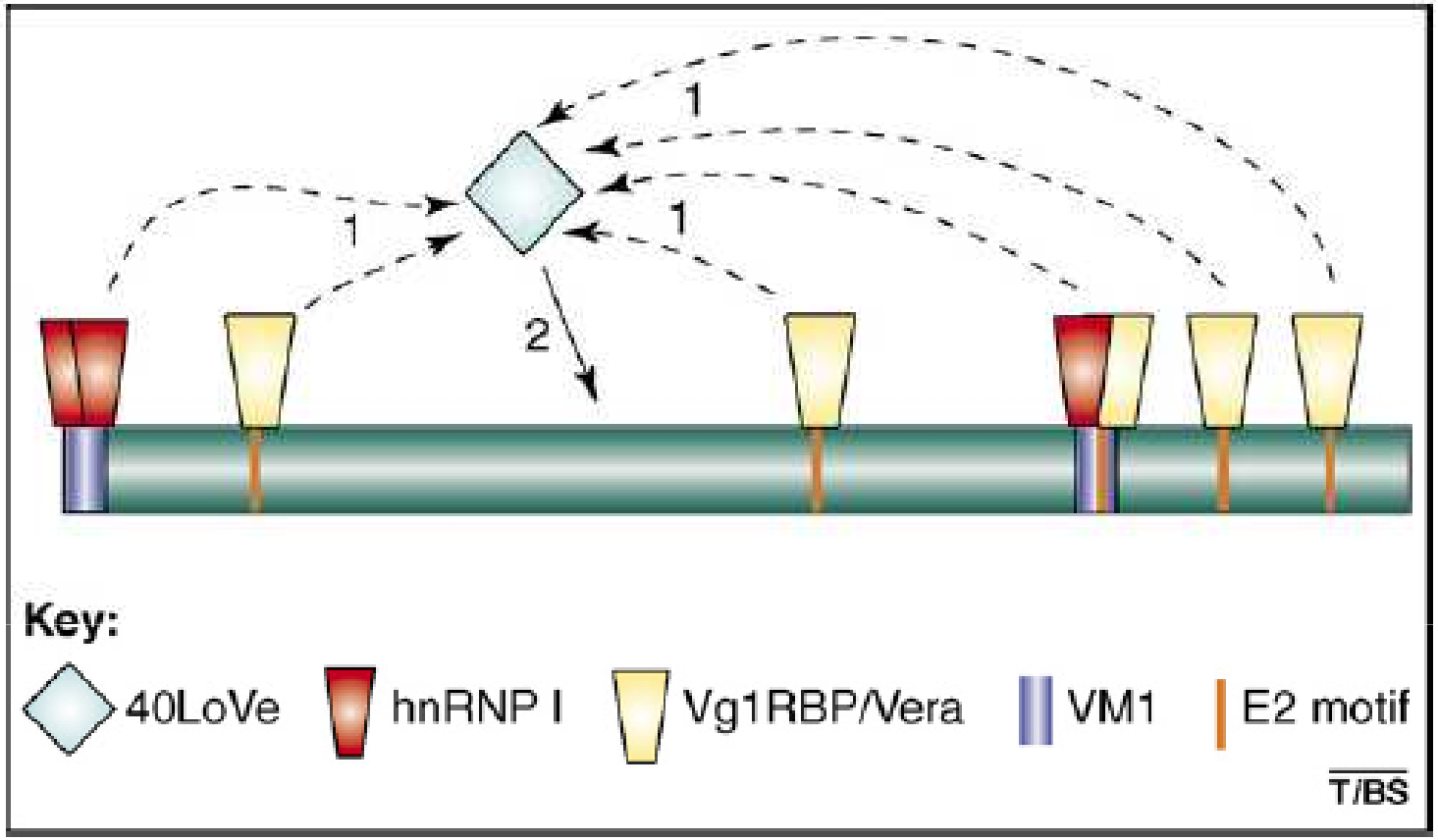



**Figure 1**. The LE of Vg1 mRNA. Vg1 is a TGF-β superfamily protein involved in mesoderm induction during *Xenopus* embryogenesis. The Vg1 mRNA localizes to the vegetal pole of oocytes through a 360-nucleotide LE in the 3′ untranslated region. The LE is depicted in green; the motifs that have been identified as required for localization are highlighted, as are the *trans*-acting factors that are known to bind these sequences. The binding of 40LoVe (pale blue diamond) to an unidentified site occurs after previous recognition of the Vg1 LE by the RNA-binding proteins hnRNP I and Vg1RBP/Vera (as indicated by the numbered arrows).

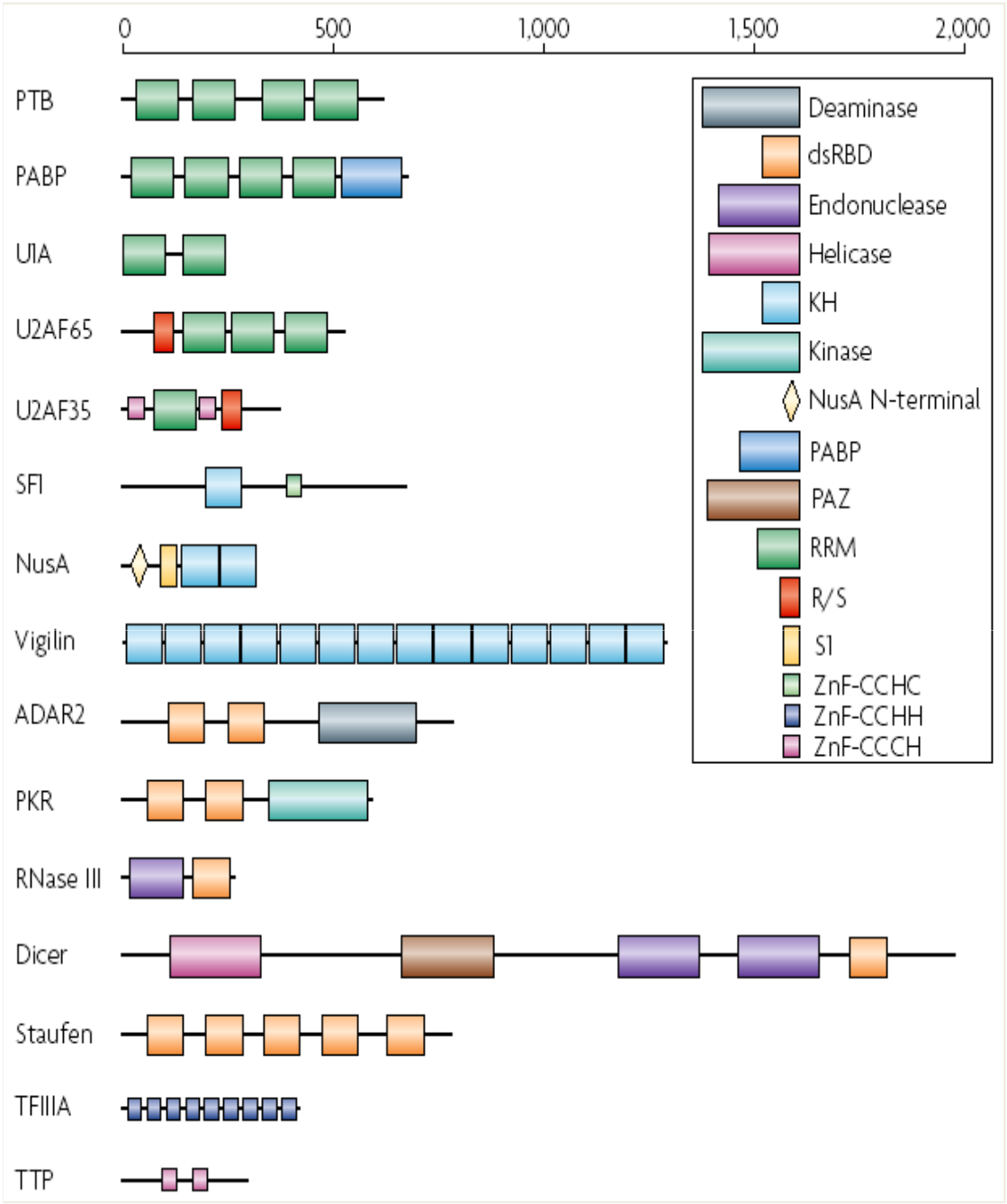
Complexity of LEs is reflected by the specific structure and mode of interaction displayed by RNA binding proteins.

# RNA-binding proteins: modular design for efficient function

Review

*Bradley M. Lunde*[‡§], *Claire Moore*[¶] *and Gabriele Varani*[*‡]

Abstract | Many RNA-binding proteins have modular structures and are composed of multiple repeats of just a few basic domains that are arranged in various ways to satisfy their diverse functional requirements. Recent studies have investigated how different modules cooperate in regulating the RNA-binding specificity and the biological activity of these proteins. They have also investigated how multiple modules cooperate with enzymatic domains to regulate the catalytic activity of enzymes that act on RNA. These studies have shown how, for many RNA-binding proteins, multiple modules define the fundamental structural unit that is responsible for biological function.




**Many RNA-binding proteins have a modular structure.** Representative examples from some of the most common RNA-binding protein families, as illustrated here, demonstrate the variability in the number of copies (as many as 14 in vigilin) and arrangements that exist. This variability has direct functional implications. For example, Dicer and RNase III both contain an endo-nuclease catalytic domain followed by a double-stranded RNA-binding domain (dsRBD). So, both proteins recognize dsRNA, but Dicer has evolved to interact specifically with RNA species that are produced through the RNA interference pathway through additional domains that recognize the unique structural features of these RNAs. Different domains are represented as coloured boxes.

These include the RNA-recognition motif (RRM; by far the most common RNA-binding protein module), the K-homology (KH) domain (which can bind both single-stranded RNA and DNA), the dsRBD (a sequence-independent dsRNA-binding module) and RNAbinding zinc-finger (ZnF) domains. Enzymatic domains and less common functional modules are also shown. PABP, poly(A)-binding protein; PTB, polypyrimidine-tract binding; R/S, Arg/Ser-rich domain; SF1, splicing factor-1; TTP, tristetraprolin; U2AF, U2 auxiliary factor.

Multiple RRM in the same peptide

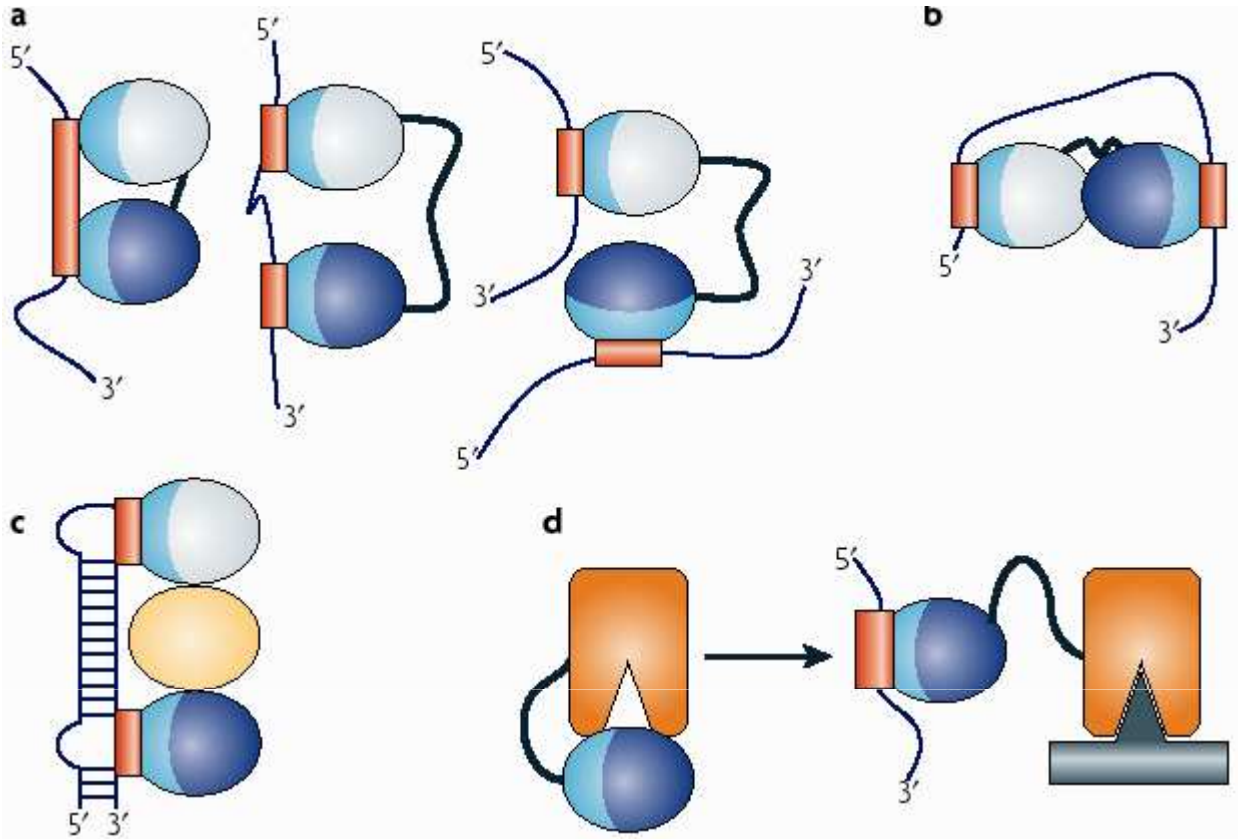



Figure 2 | **RNA-binding modules are combined to perform multiple functional roles.** RNA-binding domains (RBDs) function in various ways. **a** | They recognize RNA sequences with a specificity and affinity that would not be possible for a single domain or if multiple domains did not cooperate. Multiple domains combine to recognize a long RNA sequence (left), sequences separated by many nucleotides (centre), or RNAs that belong to different molecules altogether (right). **b** | RBDs can organize mRNAs topologically by interacting simultaneously with multiple RNA sequences. **c** | Alternatively, they can function as spacers to properly position other modules for recognition. **d** | They can combine with enzymatic domains to define the substrate specificity for catalysis or to regulate enzymatic activity. The RNA-binding modules are represented as ellipses with their RNA-binding surfaces coloured in light blue, and the corresponding binding sites in the RNA coloured in red; individual domains are coloured differently.

# Heteromeric Multimerization of RBPs mediate recognition of complex response elements

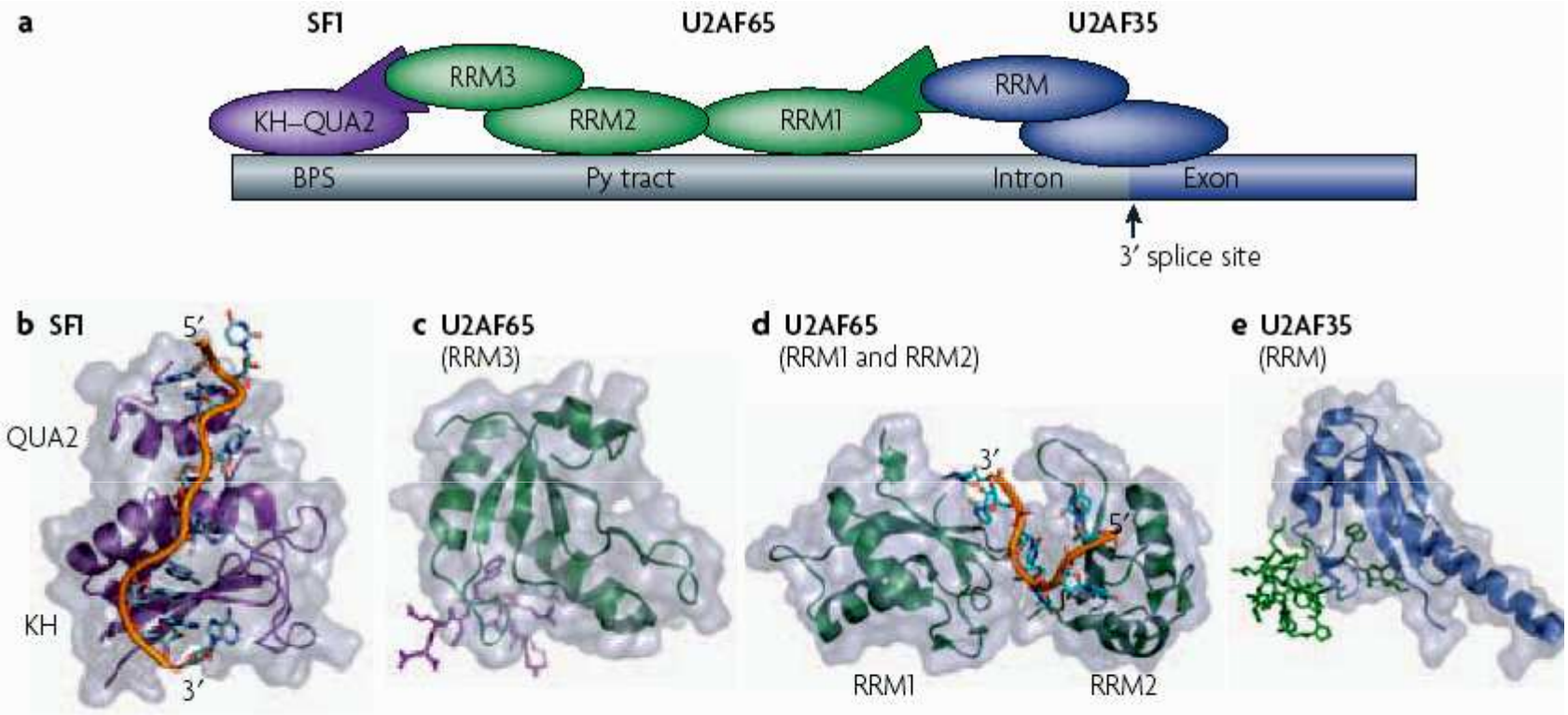



Figure 4 | **Protein–protein interactions and protein-RNA interactions define the site of spliceosomal assembly.** **a** | Schematic of the interactions between various proteins and RNA at the splicing site. The structures of some of the key domains that are involved in these interactions are shown in panel **b**. In the RNA, the branch-point sequence (BPS), pyrimidine tract (Py tract), and the 3′ splice site are labelled with the intron shown in grey and the exon in dark blue. **b** | Splicing factor-1 (SF1) recognizes the BPS through its K-homology (KH)–Quaking homology-2 (QUA2) domains, which creates an extended KH domain that can recognize the full BPS sequence RNA[42]. **c** | This interaction is strengthened by protein–protein interactions between the N terminus of SF1 and the non-canonical RNA-recognition motif-3 (RRM3) of U2 auxiliary factor subunit-65 ( U2AF65)[35]. **d** | RRM3 is bound to the pyrimidine tract through its first two canonical RRMs, RRM1 and RRM2 (REF. 10). **e** | Last, the U2AF65 interaction is also aided by protein–protein interactions between its N terminus and the non-canonical RRM of U2AF35 bound at the 3′ splice site[33]. The protein and peptide structures are colour coded as in panel **a**.

Another approach that has been used was the identification of RNA sequences that are preferred by specific proteins suspected to regulate localization.

Methods to identify RNA elements:

RNA SELEX

RNA recombination and selection

Comparison of regulatory sequences

Immunoprecipitation or CLIP followed by microarrays or RNA-Seq

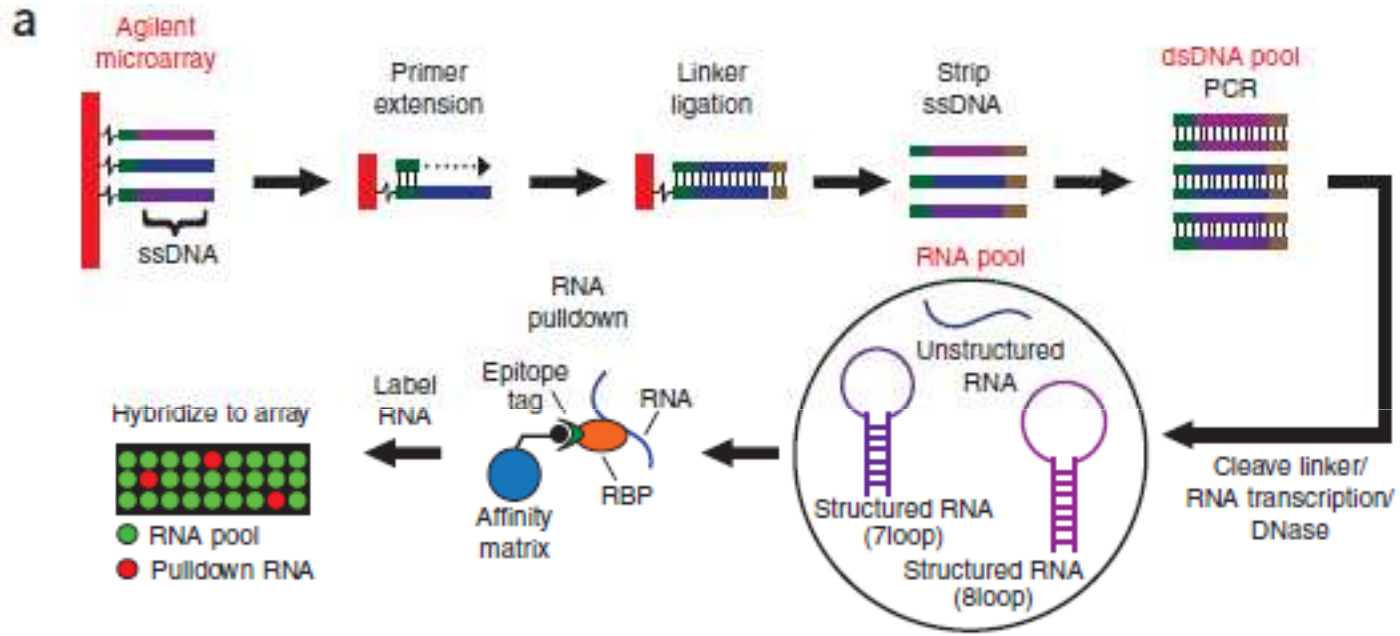
RNAcompete (see next paper)

# Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins

Debashish Ray[1,4], Hilal Kazan[2,4], Esther T Chan[3], Lourdes Peña Castillo[1], Sidharth Chaudhry[3], Shaheynoor Talukder[1], Benjamin J Blencowe[1,3], Quaid Morris[1–3] & Timothy R Hughes[1,3]

**Metazoan genomes encode hundreds of RNA-binding proteins (RBPs) but RNA-binding preferences for relatively few RBPs have been well defined[1]. Current techniques for determining RNA targets, including *in vitro* selection and RNA co-immunoprecipitation[2–5], require significant time and labor investment. Here we introduce RNAcompete, a method for the systematic analysis of RNA binding specificities that uses a single binding reaction to determine the relative preferences of RBPs for short RNAs that contain a complete range of k-mers in structured and unstructured RNA contexts. We tested RNAcompete by analyzing nine diverse RBPs (HuR, Vts1, FUSIP1, PTB, U1A, SF2/ASF, SLM2, RBM4 and YB1). RNAcompete identified expected and previously unknown RNA binding preferences. Using *in vitro* and *in vivo* binding data, we demonstrate that preferences for individual 7-mers identified by RNAcompete are a more accurate representation of binding activity than are conventional motif models. We anticipate that RNAcompete will be a valuable tool for the study of RNA-protein interactions.**

7-base (7loop) and 8-base loop (8loop) sequences in the context RNA hairpins containing unique 10–base pair stems. RNAs in t unstructured category should be either linear or contain weak se ondary structures under our assay conditions; most RNAs intera intramolecularly under physiological conditions, and therefore it not possible to design a large and diverse population of entirely line RNAs. Moreover, constraints were introduced to minimize foldi of unstructured RNAs, misfolding of the structured RNAs, extensi base-pairing among any two RNAs and microarray cross-hybridizatic This resulted in reduction of desired coverage. The two sets unstructured RNAs in the final array design each contain 81% of possible 10-mers, but contain many instances of shorter k-mers. F example, each set contains at least 12 copies of all possible 8-mers, a at least 64 copies of all possible 7-mers, with the exception of thc containing a SapI restriction site (GCTCTTC/GAAGAGC) used pool synthesis (see below). It also contains 59% of all possible 8loo (75% of which are in duplicate), 99.4% of all possible 7loops (99.3 of which are in duplicate) and all possible loops of six bases or l (100% of which are in duplicate). Thus, the pool contains indepe


a
Agilent microarray
ssDNA
Primer extension
Linker ligation
Strip ssDNA
dsDNA pool PCR
Cleave linker/ RNA transcription/ DNase
RNA pool
Unstructured RNA
Structured RNA (7loop)
Structured RNA (8loop)
RNA pulldown
Epitope tag
RNA
RBP
Affinity matrix
Label RNA
Hybridize to array
RNA pool
Pulldown RNA

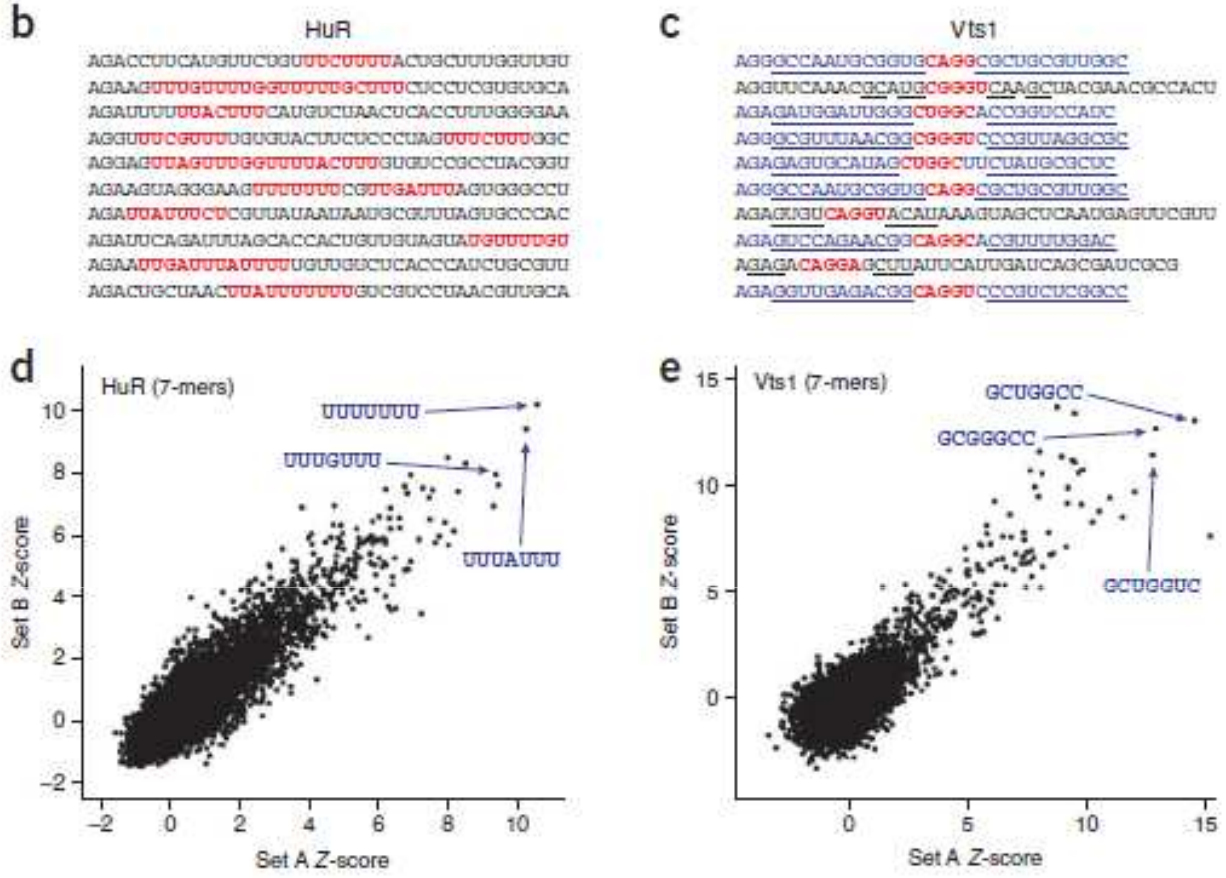



**Figure 1** The RNAcompete method and example data for HuR and Vts1. (**a**) Outline of the RNAcompete method. (**b**,**c**) The top ten binding sequences for HuR (**b**) and Vts1 (**c**). Red indicates primary sequences matching the known binding preference. Blue indicates designed stem loops. Sequences capable of base-pairing to form stem-loop structures are underlined in (**c**). (**d**,**e**) Correlations between robust average 7-mer scores (that is, excluding the top and bottom quartiles) from independent microarray probe sets (set A and set B) for HuR (**d**) and Vts1 (**e**), displayed as *Z*-scores (that is, both axes have a median of zero and s.d. of one).

| Protein | Domain(s) | Known motif/ binding site | Our highest-correlating motif | Top five 7-mers | Correlation 7-mers | Correlation motif |
|---|---|---|---|---|---|---|
| Vts1 | One SAM domain |  in stem-loop |  | GCUGGCC<br>GCGGGCC<br>CGCGGGC<br>GGCGGGC<br>CGCUGGC | 0.53 | 0.50 |
| SLM2 | One KH domain | NA |  | AGAUAAA<br>AUAAAAC<br>UAAAUAA<br>AAAUAAA<br>AUUAAAC | 0.66 | 0.35 |
| YB1 | Full-length; one cold-shock domain |  |  | CCUGGGG<br>GCCUGCC<br>CUGGGCU<br>GGUCUGC<br>CCCUGCG | 0.32 | 0.14 |
| RBM4 | Full-length; two RRM domains | NA |  | GCGGGCG<br>GCGCGGG<br>CGCGCGG<br>CGCGGGA<br>GCGCGCU | 0.56 | 0.48 |
| SF2/ASF | Two RRM domains |  |  | GCGAGCG<br>GAGCGAG<br>GAGCGGA<br>CGAGCGG<br>ACGCAGC | 0.62 | 0.48 |
| FUSIP1 | One RRM domain |  |  | AAAGAAA<br>GAAAGAA<br>AAAGAAG<br>AAAAGAA<br>AGAGAAA | 0.32 | 0.43 |
| HuR | Full-length; three RRM domains |  |  | UUUUUUU<br>UUUGUUU<br>UUAUUUU<br>UUUAUUU<br>UUGUUUU | 0.62 | 0.56 |
| U1A | One of two RRM domains |  |  | AUUGCAC<br>UUGCACA<br>UUGCACG<br>UUUGCAC<br>UUGCACU | 0.54 | 0.20 |
| PTB | Full-length; four RRM domains |  |  | CUUUCUU<br>UUUCUUU<br>UCUUUCU<br>UUACUUU<br>UCUAUCU | 0.39 | 0.30 |