Lecture 2.19

Processing to export

Alternative Splicing (Splice Lesson 7 to Lesson 18)
What we have learnt:

More than 90% of vertebrate genes (protein-coding) undergo Alternative Splicing

Alternative Splicing usually concerns 1 or 2 exons in a gene (median 8.8 exons), but exceptions exists (from none to several).

cis-element defining AS are two kinds:
1) 5'-ss, 3'-ss, branching site and polypyrimidine tract are common to all introns but their "goodnes score" can vary, it is generally high for constitutive exons and low for alternative exons (exceptions exist).
2) exonic and intronic sequences that either enhance or repress splicing sites utilization: ESE, ISE, ESS, ISS, when intronic usually surrounding exons.

trans-factors regulating AS are different kinds:
1) snNRPs and associated proteins, that recognize type 1) *cis*-elements
2) SR and hnRNP RNA binding proteins, that have a role in both connecting elements of the constitutive splicing (exon definition and intron definition models) and in AS regulation by binding to ESE, ISE, ESS, ISS.
3) tissue-specific splicing regulators, exemplified by the Nova-1 and Nova-2 proteins

In parallel with transcription factors, regulators are not "per se" repressors or activators, but their activity is more or less context- and position-dependent (exemplified by Nova).

RNA – protein interaction
gene expression and splicing regulation

Mutually exclusive exons ?

suggestions from bioinformatic studies

(the case of Dscam gene in Drosophila)

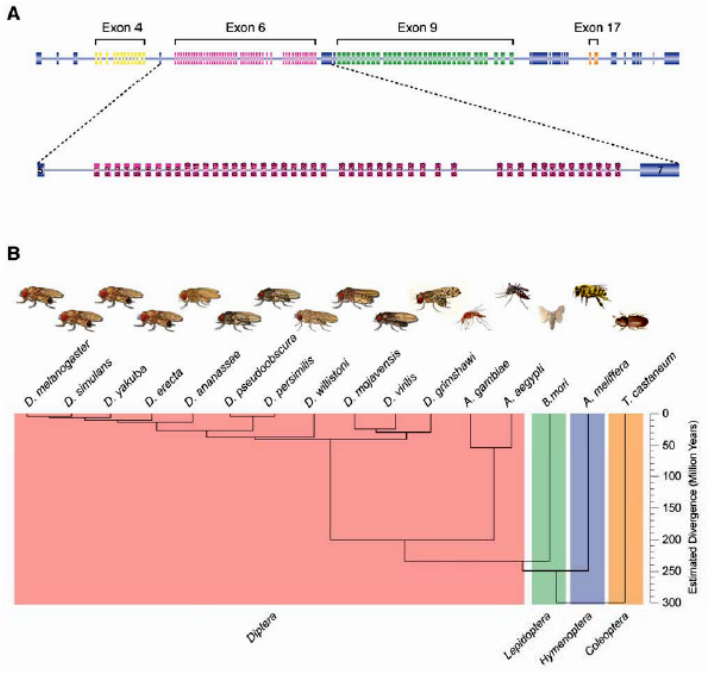
Cell, Vol. 123, 65–73, October 7, 2005, Copyright ©2005 by Elsevier Inc. DOI 10.1016/j.cell.2005.07.028

Mutually Exclusive Splicing of the Insect *Dscam* Pre-mRNA Directed by Competing Intronic RNA Secondary Structures

Brenton R. Graveley*
Department of Genetics and Developmental Biology
University of Connecticut Health Center
263 Farmington Avenue
Farmington, Connecticut 06030

Summary

Drosophila Dscam **encodes 38,016 distinct axon guidance receptors through the mutually exclusive alternative splicing of 95 variable exons. Importantly, known mechanisms that ensure the mutually exclusive splicing of pairs of exons cannot explain this phenomenon in *Dscam*. I have identified two classes of conserved elements in the *Dscam* exon 6 cluster, which contains 48 alternative exons—the docking site, located in the intron downstream of constitutive exon 5, and the selector sequences, which are located upstream of each exon 6 variant. Strikingly, each selector sequence is complementary to a portion of the docking site, and this pairing juxtaposes one, and only one, alternative exon to the upstream constitutive exon. The mutually exclusive nature of the docking site:selector sequence interactions suggests that the formation of these competing RNA structures is a central component of the mechanism guaranteeing that only one exon 6 variant is included in each *Dscam* mRNA.**



Two conserved sequence emerged: the 1° is the “docking site”, that preceedes exon block.

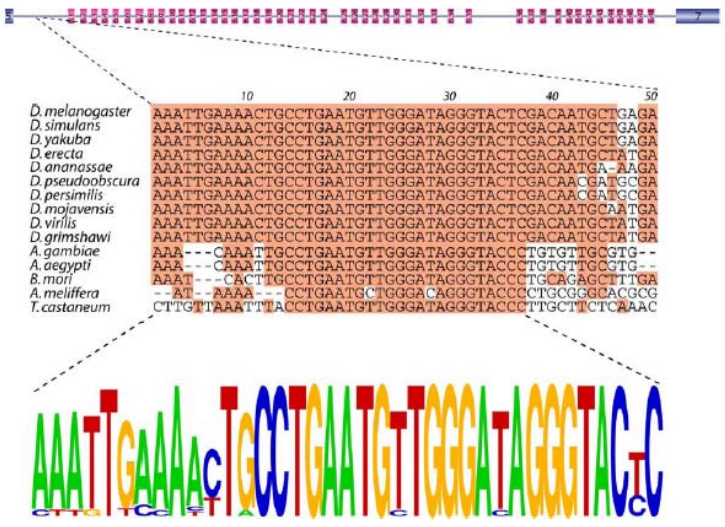


Figure 2. The Docking Site

The nucleotide sequence alignment of the docking sites of 15 insects. The most common nucleotide at each position is shaded. The docking site consensus is represented as a pictogram (bottom). The height of each letter represents the frequency of each nucleotide at that position.

Two conserved sequence emerged: second is the “selector”, that preceedes each exons.

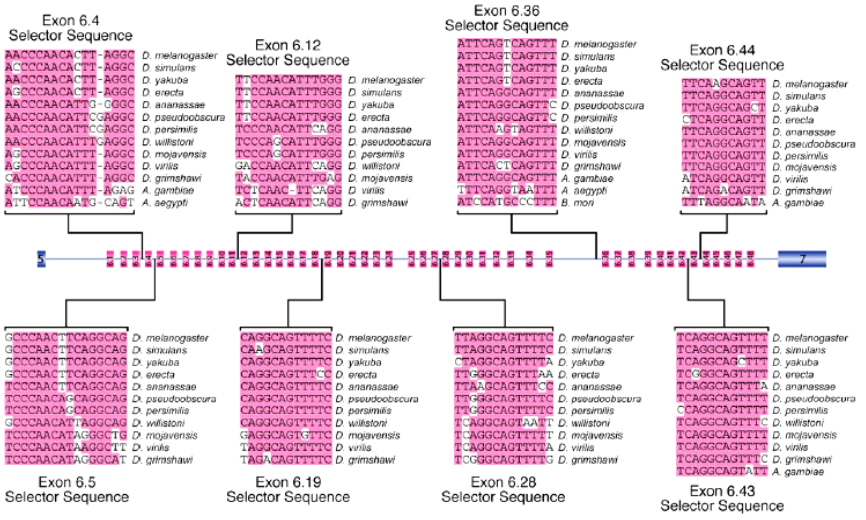


Figure 3. Conservation of Selector Sequences

Alignment of eight of the selector sequences and their locations with the exon 6 cluster are depicted. The most common nucleotides at each position are shaded.

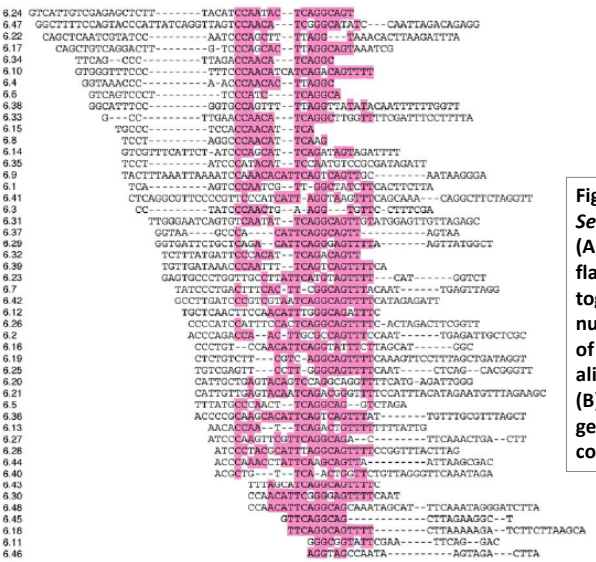
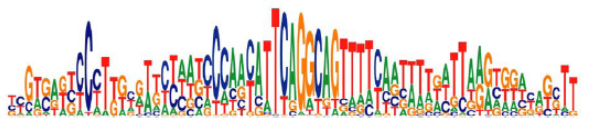


Figure 4. The *D. melanogaster Selector Sequence Consensus*
(A) The 48 selector sequences and flanking sequence were aligned together. The most frequent nucleotides in the central portion of the alignment are highlighted.
(B) The alignment was used to generate a selector sequence consensus.

B



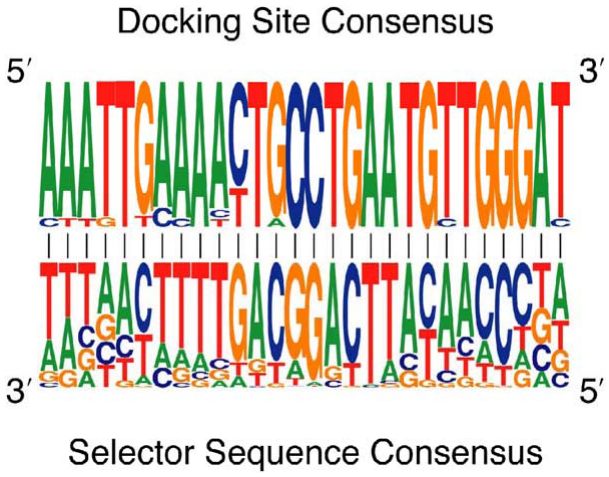


Figure 5. The Docking Site and Selector Sequences Consensus Are Complementary
The docking site consensus sequence is complementary to the central 28 nucleotides of the selector sequence consensus. The most frequent nucleotide at each position of the selector sequence is complementary to the docking site.

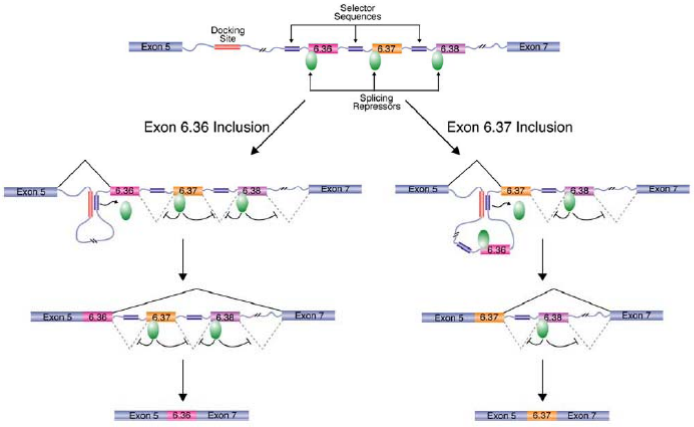


Figure 7. Model for the Mechanism of *Dscam* Exon 6 Mutually Exclusive Splicing

A model of the *Dscam* exon 6 cluster is depicted in which only variable exons 6.36, 6.37, and 6.38 are shown. A key component of this model is that a splicing repressor functions to prevent the exon 6 variants from being spliced together (green oval). In order for an exon 6 variant to be included in the *Dscam* mRNA, the selector sequence upstream of the exon must interact with the docking site. For example, if exon 6.36 is to be included (left), the selector sequence upstream of exon 6.36 will interact with the docking site. Likewise, if exon 6.37 is to be included, the selector sequence upstream of exon 6.37 will interact with the docking site. By some unknown mechanism, the docking site:selector sequence interaction inactivates the splicing repressor on the downstream exon and, consequently, activates the splicing of the downstream exon 6 variant to exon 5. Subsequently, the exon that is joined to exon 5 can only be spliced to constitutive exon 7 because the remaining exon 6 variants are actively repressed by the splicing repressor. As a result, only one exon 6 variant is included in the mRNA.

The model may explain how an exon is stochastically selected for splicing

however

RNA transcripts in a cell are always spliced in the same way.... how can the machinery “remember” which selector was used the first time?

An explanation may come from **nucleotide positioning** studies.
(see below)

Evolution of AS

Regulation of regulators

definition

Constitutive exon:	an exon that is found in every mRNA from that particular gene
Alternative exon:	an exon that may or may not be present in different mRNAs from that gene

phylogenetic conservation

Constitutive exons: usually quite high

Alternative exons: not so high, it is thought one of the major issues of evolution of organism complexity, at least in vetrebrates. Hortologue comparison defines CAS (conserved AS) and non-CAS exons.

How new AS exons arise ?

•duplication followed by alternatization of constitutive exons

•exonization of noncoding sequences (example the Alu SINE)

•exon shuffling between separate genes

How prevalent is functional alternative splicing in the human genome?☆

Rotem Sorek[1,2], Ron Shamir[3] and Gil Ast[1]

[1]Department of Human Genetics, Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv 69978, Israel
[2]Compugen, 72 Pinchas Rosen Street, Tel Aviv 69512, Israel
[3]School of Computer Science, Tel Aviv University, Tel Aviv 69978, Israel

Comparative analyses of ESTs and cDNAs with genomic DNA predict a high frequency of alternative splicing in human genes. However, there is an ongoing debate as to how many of these predicted splice variants are functional and how many are the result of aberrant splicing (or 'noise'). To address this question, we compared alternatively spliced cassette exons that are conserved between human and mouse with EST-predicted cassette exons that are not conserved in the mouse genome. Presumably, conserved exon-skipping events represent functional alternative splicing. We show that conserved (functional) cassette exons possess unique characteristics in size, repeat content and in their influence on the protein. By contrast, most non-conserved cassette exons do not share these characteristics. We conclude that a significant portion of cassette exons evident in EST databases is not functional, and might result from aberrant rather than regulated splicing.

Sorek et al., (2004), Trends in Genetics, 20: 68

Numerous studies have shown that alternative splicing is prevalent in mammalian genomes. Using ESTs and cDNAs aligned to the genomic sequence, these studies estimate that between ~~35% and 59%~~ of all human genes undergo alternative splicing [1,2]. However, it is not clear how many of the splice variants predicted from ESTs are functional and how many represent aberrant splicing ('noise') or EST artefacts (such as genomic contamination) [3–5]. An mRNA variant can be defined as being 'functional' if it is required during the life-cycle of the organism and activated in a regulated manner.

Old data! RNA-seq shows that maybe almost all genes do AS

What portion of the observed splice variants represents functional alternative splicing? We employed a comparative genomics approach to address this question, by compiling a dataset of exon-skipping events (cassette exons) that are conserved between human and mouse. The conservation of such events in both human and mouse species, which diverged from their common ancestor 75–110 million years ago, suggests the functional importance of these exons.

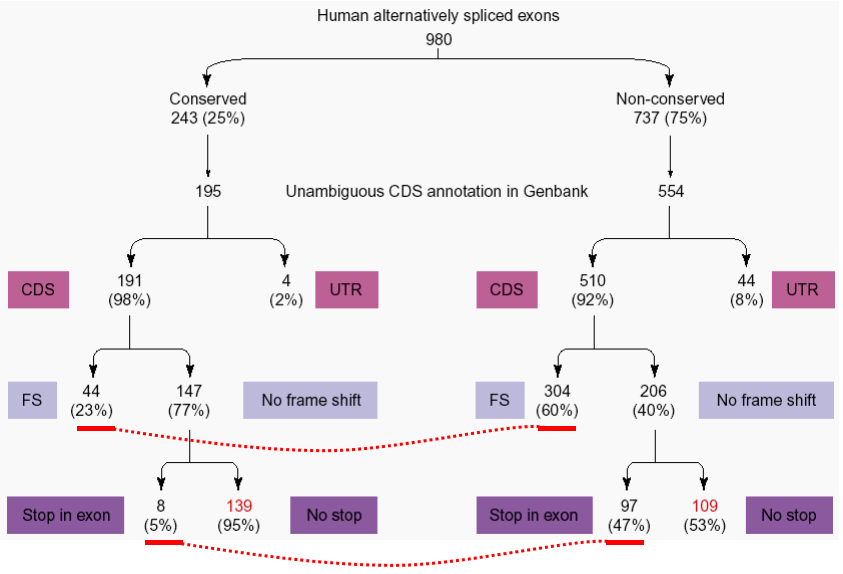
Sorek et al., (2004), Trends in Genetics, 20: 68

Dataset: 980 EST-predicted human alternatively spliced cassette exons.

From these 980 exons, 243 (25%) were also found to be alternatively spliced in mouse ['conserved alternatively spliced exons' (CAS exons)].

The remaining 737 (75%) are non-conserved alternatively spliced exons' (non-CAS exons).

Sorek et al., (2004), Trends in Genetics, 20: 68



Sorek et al., (2004), Trends in Genetics, 20: 68

Table 1. Features differentiating between conserved alternatively spliced exons and non-conserved alternatively spliced exons

Features	Conserved alternatively spliced exons	Non-conserved alternatively spliced exons[a]	*P* value[b]
Average size	87	116	$P < 10^{-6}$
Percentage of exons that are a multiple of three	77% (147/191)	40% (206/510)	$P < 10^{-5}$
Percentage of exons that are 'peptide cassettes'[c]	73% (139/191)	21% (109/510)	$P < 10^{-15}$
Percentage of exon insertions that result in a longer protein (from a total of exons that are not a multiple of three)	61% (27/44)	8% (25/304)	$P < 10^{-9}$
Percentage of exon insertions that result in a protein <100 amino acids (from a total of exons that are not a multiple of three)	9% (4/44)	30% (91/304)	$P < 0.02$
Average supporting expressed sequences	9	2.2	$P < 10^{-6}$
Percentage of exons that contain repetitive elements[d]	<0.5% (1/243)	26% (191/737)	$P < 10^{-20}$

[a]Non-conserved alternatively spliced exons are human exons that were found to be alternatively spliced in human ESTs but were not found in the mouse genome.
[b]The *P* value was calculated using Fisher's exact test, except for the 'average size' and 'average support', for which *P* values were calculated using student's *t* test.
[c]A 'peptide cassette' exon is defined such as it neither causes a frame shift nor contains a stop codon, so that the effect of its insertion or deletion on the translated protein is a local insertion or deletion of a peptide.
[d]Exons were aligned to a database of repetitive elements, and 'hits' with e-value $<10^{-10}$ were considered positive.

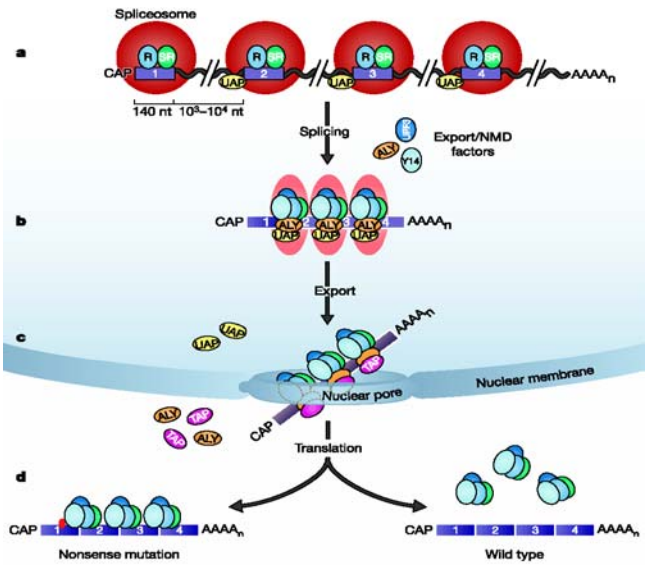
Sorek et al., (2004), Trends in Genetics, 20: 68

Insertion of exons with in-frame stop codons will result in premature termination of translation (truncated protein)

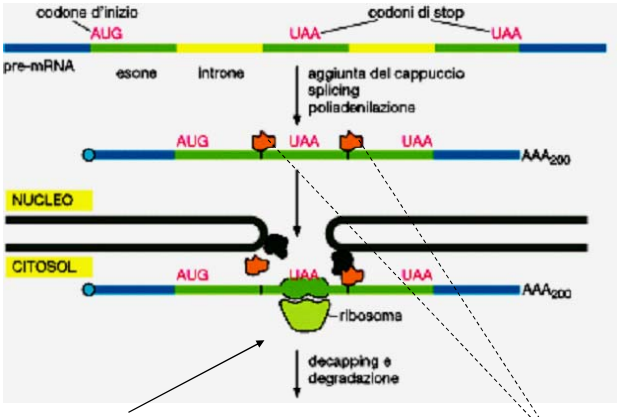
However, even when nonsense mutations arise (e.g. in cancer cells), truncated proteins are rarely found in cells.

How does it happens ?

Nonsense mediated Decay



Non-sense mediated decay of mRNA



Translation test

Exon-junction complex

Further processing: more “signals” in the RNA sequence ?

capping

editing

transport from nucleus to cytoplasm (export)

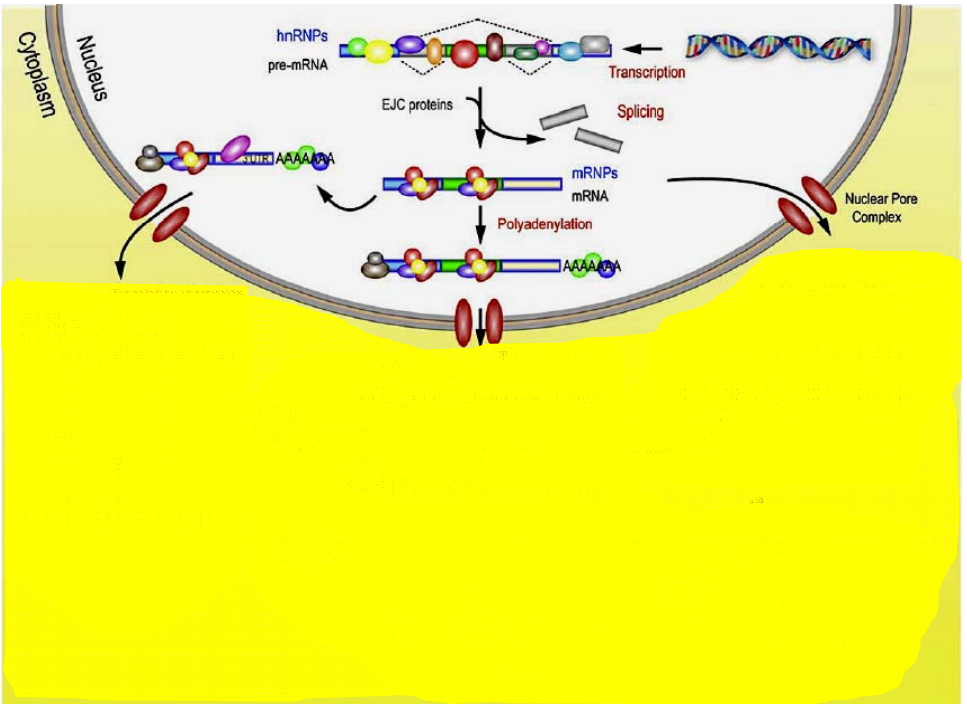
nonsense mediated decay (NMD)

cytoplasmatic localization

Translational initiation

Stability – degradation

Deadenylation – cytoplasmatic polyadenylation



These events are:

- Specific – they depend on RNA sequences
- Due to interation with specific RNA-binding proteins
- Regulated

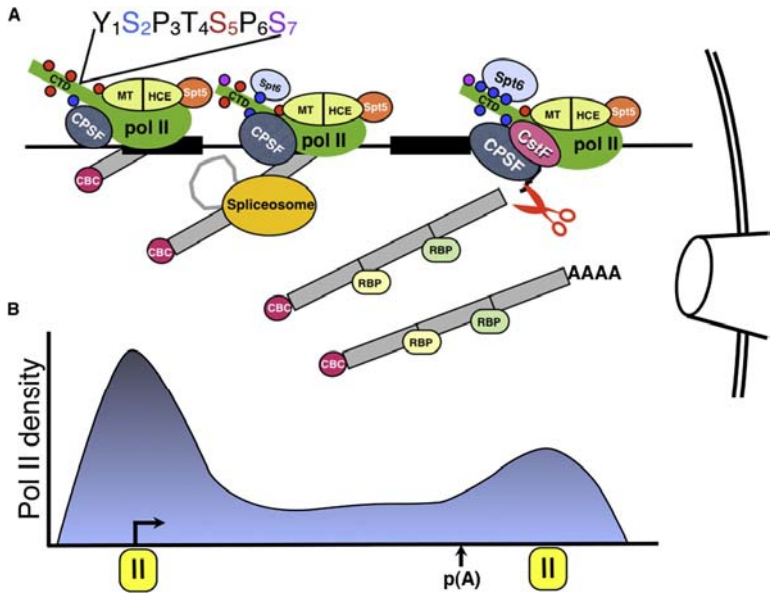
Molecular Cell
Review

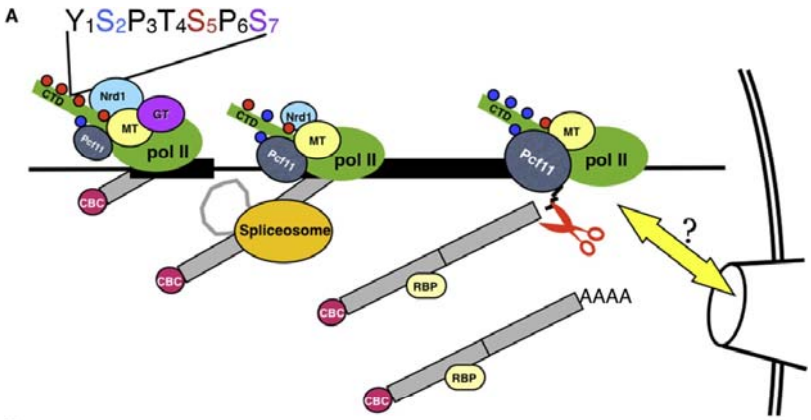
"Cotranscriptionality": The Transcription Elongation Complex as a Nexus for Nuclear Transactions

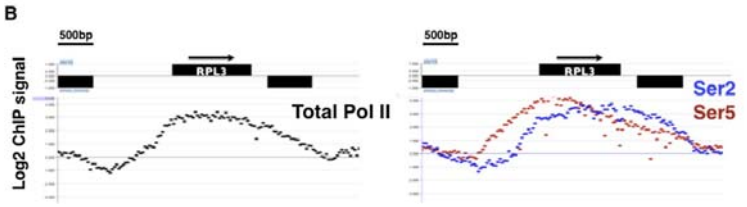
Roberto Perales[1] and David Bentley[1,*]
[1]Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, UCHSC, MS8101, P.O. Box 6511, Aurora CO, 80045, USA
*Correspondence: david.bentley@uchsc.edu
DOI 10.1016/j.molcel.2009.09.018

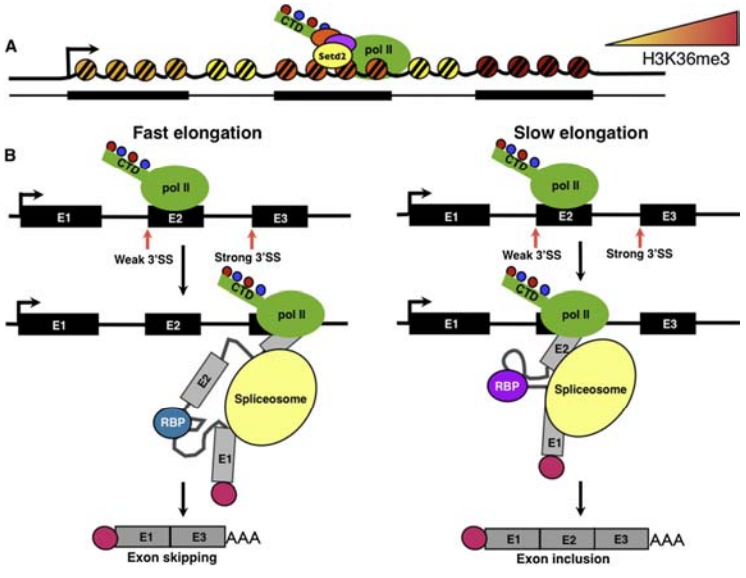
Much of the complex process of RNP biogenesis takes place at the gene cotranscriptionally. The target for RNA binding and processing factors is, therefore, not a solitary RNA molecule but, rather, a transcription elongation complex (TEC) comprising the growing nascent RNA and RNA polymerase traversing a chromatin template with associated passenger proteins. RNA maturation factors are not the only nuclear machines whose work is organized cotranscriptionally around the TEC scaffold. Additionally, DNA repair, covalent chromatin modification, "gene gating" at the nuclear pore, Ig gene hypermutation, and sister chromosome cohesion have all been demonstrated or suggested to involve a cotranscriptional component. From this perspective, TECs can be viewed as potent "community organizers" within the nucleus.

review




A
Y1S2P3T4S5P6S7
CTD
Nrd1
GT
MT
Pcf11
pol II
CBC
Spliceosome
RBP
AAAA
?


B
500bp
RPL3
Log2 ChIP signal
Total Pol II
Ser2
Ser5

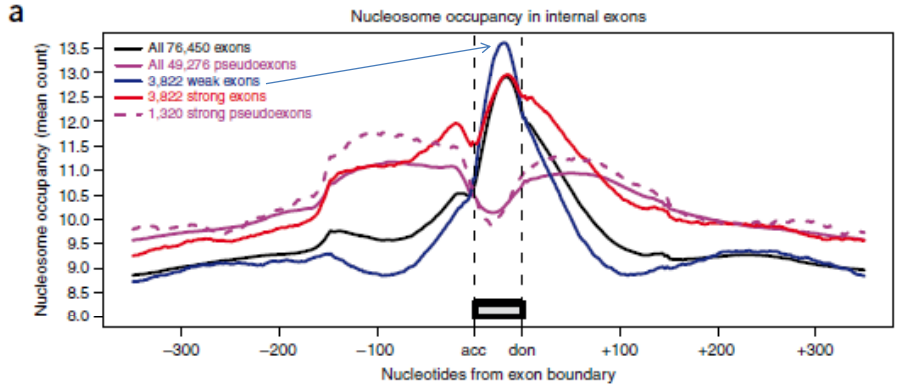

A
CTD
pol II
Setd2
H3K36me3
B
Fast elongation
Slow elongation
E1
E2
E3
Weak 3'SS
Strong 3'SS
RBP
Spliceosome
AAA
Exon skipping
Exon inclusion

nature
structural &
molecular biology

Nucleosome positioning as a determinant of exon recognition

Hagen Tilgner[1,3], Christoforos Nikolaou[1,3], Sonja Althammer[1], Michael Sammeth[1], Miguel Beato[1], Juan Valcárcel[1,2] & Roderic Guigó[1]

Chromatin structure influences transcription, but its role in subsequent RNA processing is unclear. Here we present analyses of high-throughput data that imply a relationship between nucleosome positioning and exon definition. First, we have found stable nucleosome occupancy within human and *Caenorhabditis elegans* exons that is stronger in exons with weak splice sites. Conversely, we have found that pseudoexons—intronic sequences that are not included in mRNAs but are flanked by strong splice sites—show nucleosome depletion. Second, the ratio between nucleosome occupancy within and upstream from the exons correlates with exon-inclusion levels. Third, nucleosomes are positioned central to exons rather than proximal to splice sites. These exonic nucleosomal patterns are also observed in non-expressed genes, suggesting that nucleosome marking of exons exists in the absence of transcription. Our analysis provides a framework that contributes to the understanding of splicing on the basis of chromatin architecture.



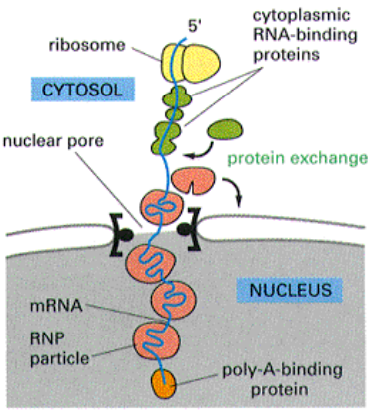
Molecular Cell
Article

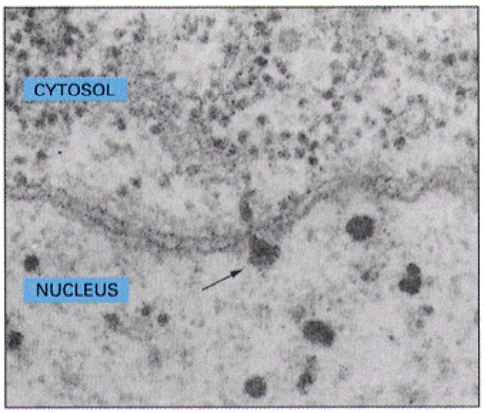
Biased Chromatin Signatures around Polyadenylation Sites and Exons

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
DOI 10.1016/j.molcel.2009.10.008

SUMMARY

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.





7012–7017 | PNAS | June 19, 2001 | vol. 98 | no. 13

Assembly and transport of a premessenger RNP particle

Bertil Daneholt*

Department of Cell and Molecular Biology, Medical Nobel Institute, Karolinska Institutet, Bi

Salivary gland cells in the larvae of the dipteran *Chironomus tentans* offer unique possibilities to visualize the assembly and nucleocytoplasmic transport of a specific transcription product. Each nucleus harbors four giant polytene chromosomes, whose transcription sites are expanded, or puffed. On chromosome IV, there are two puffs of exceptional size, Balbiani ring (BR) 1 and BR 2. A BR gene is 35–40 kb, contains four short introns, and encodes a 1-MDa salivary polypeptide. The BR transcript is packed with proteins into a ribonucleoprotein (RNP) fibril that is folded into a compact ring-like structure. The completed RNP particle is released into the nucleoplasm and transported to the nuclear pore, where the RNP fibril is gradually unfolded and passes through the pore. On the cytoplasmic side, the exiting extended RNP fibril becomes engaged in protein synthesis and the ensuing polysome is anchored to the endoplasmic reticulum. Several of the BR particle proteins have been characterized, and their fate during the assembly and transport of the BR particle has been elucidated. The proteins studied are all added cotranscriptionally to the pre-mRNA molecule. The various proteins behave differently during RNA transport, and the flow pattern of each protein is related to the particular function of the protein. Because the cotranscriptional assembly of the pre-mRNP particle involves proteins functioning in the nucleus as well as proteins functioning in the cytoplasm, it is concluded that the fate of the mRNA molecule is determined to a considerable extent already at the gene level.

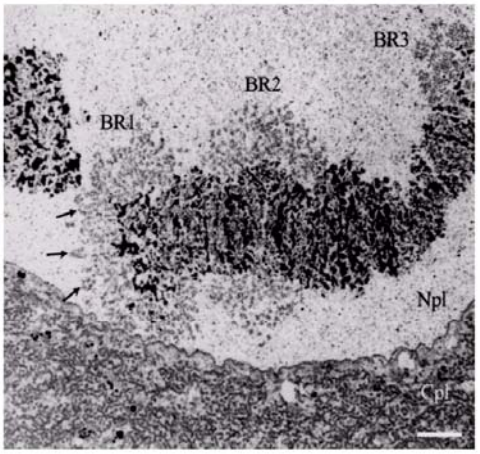


Fig. 1. Electron micrograph showing chromosome IV with its three giant puffs (BRs) in a salivary gland cell from *C. tentans*. The three BRs (BR1, BR2, and BR3) are indicated as well as the nucleoplasm (Npl) and cytoplasm (Cpl). The arrows mark a few prominent transcription loops (cf. Fig. 2*D*). (Bar equals 2 μm.)

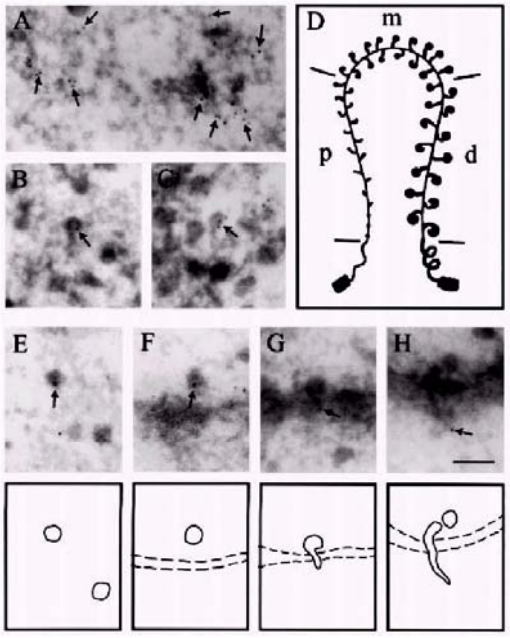


Fig. 2. Intracellular distribution of the cap-binding protein **CBP20** in *C. tentans* salivary gland cells studied by immunoelectron microscopy. The assembly of the BR RNP particle is shown in *A–D*: proximal portions of the BR gene are displayed in *A*, distal portions in *B* and *C*, and a schematic drawing of the BR gene in *D* (p, proximal; m, middle; d, distal portions of the gene). The fate of the released BR particles is shown in *E–H*: BR particles are present in the nucleoplasm (*E*), at the pore (*F*), and in an unfolded conformation when passing through the pore (*G* and *H*). Gold particles are marked by arrows and indicate the position of CBP20. It should be noted that gold particles are at the leading 59 end of the BR particle when it passes through the nuclear pore. (Bar equals 100 nm.) Modified from ref. 27; produced by permission of The Rockefeller University Press.

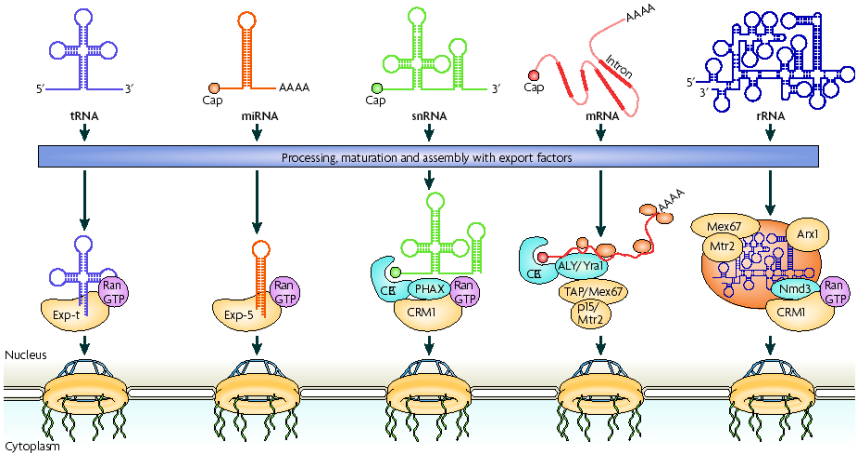


Figure 1 | **The different RNA export pathways.** The major RNA export routes are shown (tRNA, microRNA (miRNA), small nuclear (sn)RNA, mRNA, ribosomal (r)RNA). In each case, the primary RNA transcript is shown, as well as the transport-competent RNA after it has undergone processing, maturation and assembly with export factors (export adaptors are shown in blue, export receptors are shown in yellow). Prominent structural motifs in pre-RNAs are indicated (single/double-stranded RNA, loops, exons and introns, 5′ cap and 3′ poly(A) tail). For the mRNA export route, the names of both metazoan and yeast proteins are indicated, and mRNAs are shown with additional adaptor proteins and RNA-binding factors (orange ovals). In the case of rRNA, the general exporter in eukaryotes, CRM1, and two auxiliary exporters, Mex67–Mtr2 and Arx1, that have only been studied in yeast are depicted. CBC, cap-binding complex; Exp, exportin.

Kohler and Hurt, 2007

Meccanismi di esporto (esportine e ran)

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

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.

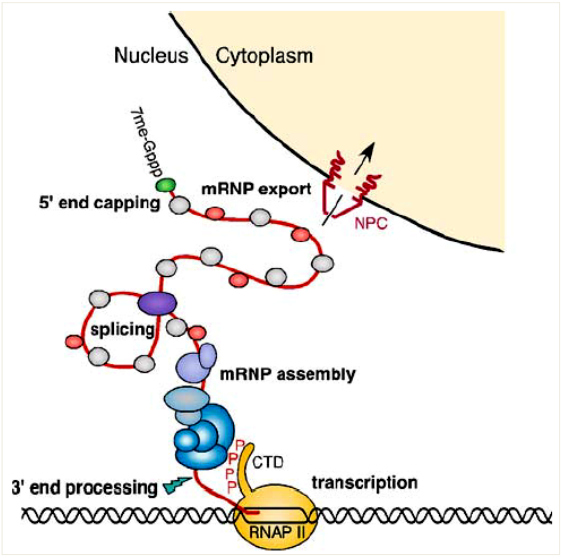
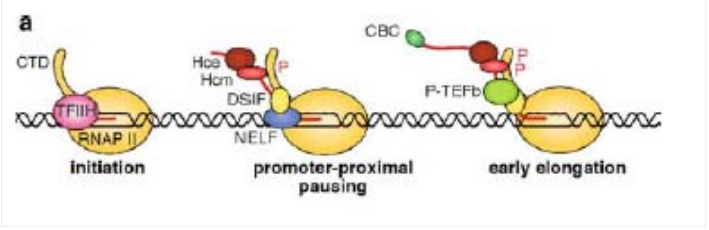
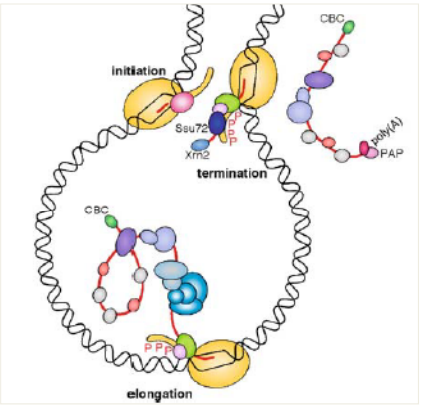
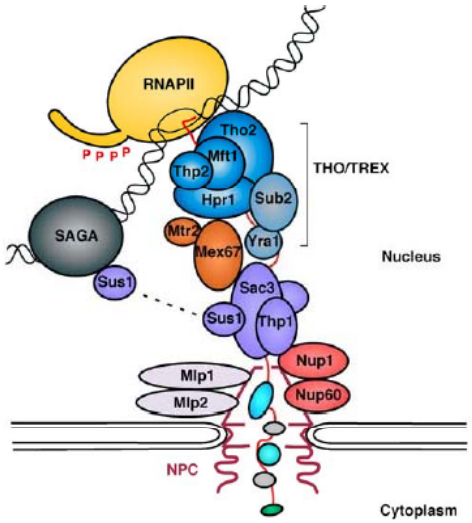


Fig. 1 Schematic view of the nuclear side of eukaryotic gene expression, from transcription to nuclear export. NPC Nuclear pore complex, CTD C-terminal domain of Rpb1, RNAPII RNA polymerase II.







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	Xm2		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.

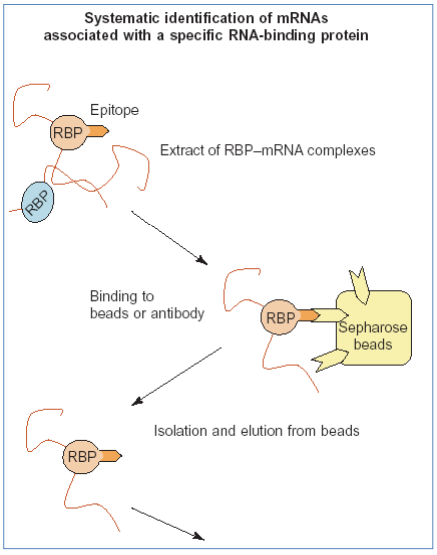
transport protein specificity ?

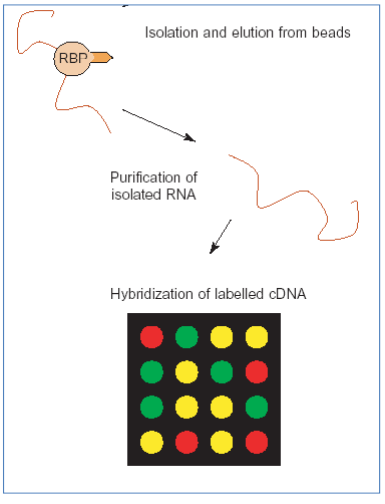
Genome-wide analysis of RNA–protein interactions illustrates specificity of the mRNA export machinery

Haley Hieronymus & Pamela A. Silver

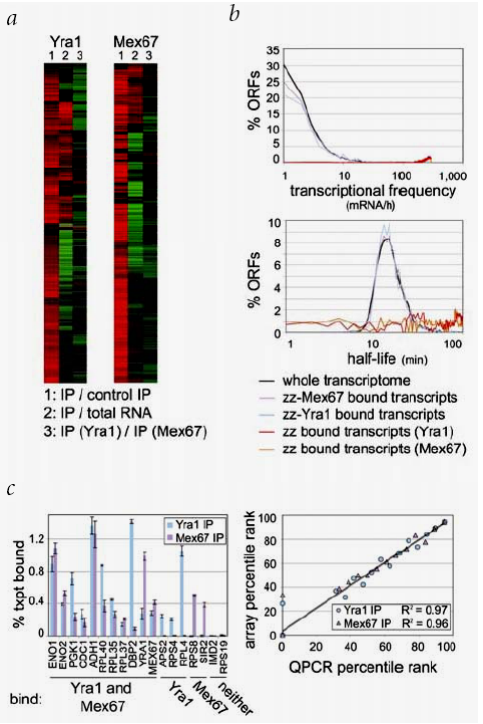
nature genetics • volume 33 • february 2003
33:155-161.

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.



Genome-wide identification of functionally distinct subsets of cellular mRNAs associated with two nucleocytoplasmic-shuttling mammalian splicing factors

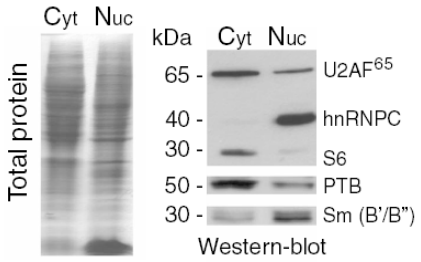
Margarida Gama-Carvalho[*], Nuno L Barbosa-Morais[*†], Alexander S Brodsky[‡§], Pamela A Silver[‡] and Maria Carmo-Fonseca[*]

Genome **Biology** 2006, **7**:R113 (doi:10.1186/gb-2006-7-11-r113)

Background: Pre-mRNA splicing is an essential step in gene expression that occurs co-transcriptionally in the cell nucleus, involving a large number of RNA binding protein splicing factors, in addition to core spliceosome components. Several of these proteins are required for the recognition of intronic sequence elements, transiently associating with the primary transcript during splicing. Some protein splicing factors, such as the U2 small nuclear RNP auxiliary factor (U2AF), are known to be exported to the cytoplasm, despite being implicated solely in nuclear functions. This observation raises the question of whether U2AF associates with mature mRNA-ribonucleoprotein particles in transit to the cytoplasm, participating in additional cellular functions.

Results: Here we report the identification of RNAs immunoprecipitated by a monoclonal antibody specific for the U2AF 65 kDa subunit (U2AF^{65}) and demonstrate its association with spliced mRNAs. For comparison, we analyzed mRNAs associated with the polypyrimidine tract binding protein (PTB), a splicing factor that also binds to intronic pyrimidine-rich sequences but additionally participates in mRNA localization, stability, and translation. Our results show that 10% of cellular mRNAs expressed in HeLa cells associate differentially with U2AF^{65} and PTB. Among U2AF^{65}-associated mRNAs there is a predominance of transcription factors and cell cycle regulators, whereas PTB-associated transcripts are enriched in mRNA species that encode proteins implicated in intracellular transport, vesicle trafficking, and apoptosis.

Conclusion: Our results show that U2AF^{65} associates with specific subsets of spliced mRNAs, strongly suggesting that it is involved in novel cellular functions in addition to splicing.



U2AF^{65} is localized in both cytoplasmatic and nuclear fractions

PTB (polypyrimidine-tract binding protein): idem

mRNA (polyA) bound to either U2AF65 or PTB immunoprecipitated, labeled and hybridized to microarrays

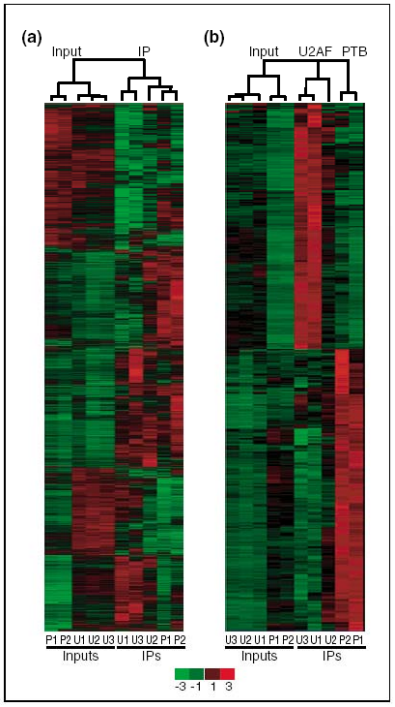


Figure 2 Clustering analysis of microarray data.

(a) Unsupervised clustering of the microarray dataset was performed with the dChip software using standard settings considering all nonredundant probes with positive hybridization signal. The dataset includes microarray hybridization results from input and immunoprecipitation (IP) samples from three experiments with anti-U2AF^{65} antibody (U1 to U3) and two experiments with anti-PTB antibody (P1 and P2). Sample clustering defines a tree with two first level branches corresponding to input and IP samples.

(b) Re-clustering analysis after clearing transcripts that were over-represented either in the inputs or in all immunoprecipitation samples. Sample clustering defines a tree with three first level branches corresponding to input, U2AF^{65}, and PTB immunoprecipitation samples. For clustering analysis, the probe signal intensities for each mRNA are standardized to have mean 0 and standard deviation 1 across all samples. The color scale for mRNAs is presented as follows: red represents expression level above mean expression of a gene across all samples, black represents mean expression; and green represents expression lower than the mean. Because of the standardization, probe signal intensities most likely fall within [-3, 3].
PTB, polypyrimidine tract binding protein; U2AF, U2 small nuclear RNP auxiliary factor.

from Gama-Carvalho 2006 Genome Biol 7:R113