Alternative splicing control 2

The most significant 4 slides from last lecture:

Alternative (skipped?) events		Total events (×10³)	Number detected (×10³)	Both isoforms detected	Number tissue-regulated	% Tissue-regulated (observed)	% Tissue-regulated (estimated)
		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

■ Constitutive exon or region   ▬ Body read   ┅ Junction read   pA Polyadenylation site

□ Alternative exon or extension   Inclusive/extended isoform   Exclusive isoform   Both isoforms

2

introne-10-10^5 nucleotidi · esone-circa 200 nucleotidi · introne-10-10^5 nucleotidi

U1 · U2 · U1 · U2 · hnRNP · 3′ · proteine che legano poli-A

complessi hnRNP

U2AF$_{65}$ · 35 · SR · SR · SR · U1 · U2 · A · U2AF$_{65}$ · 35 · SR · SR · SR · U1

Y Y Y Y · A G · G U · Y Y Y Y · A G · G U

ESE · ESE

introne · esone · introne · esone · introne
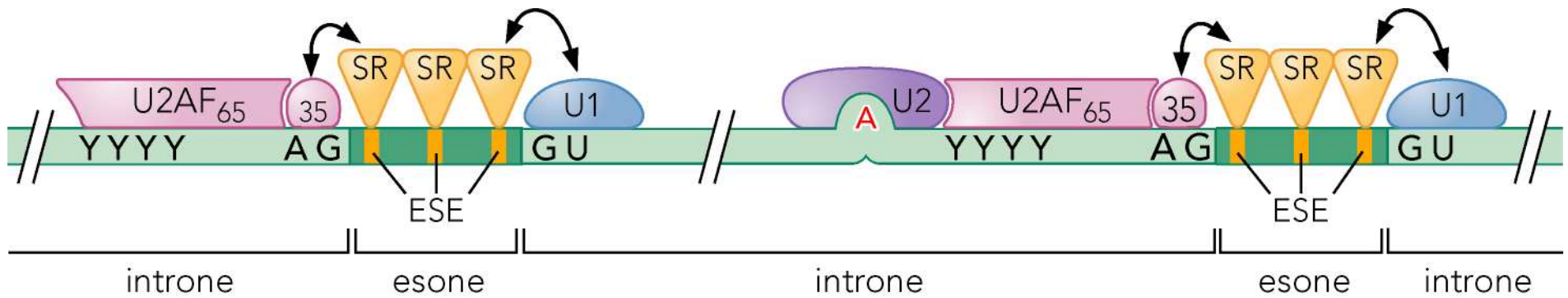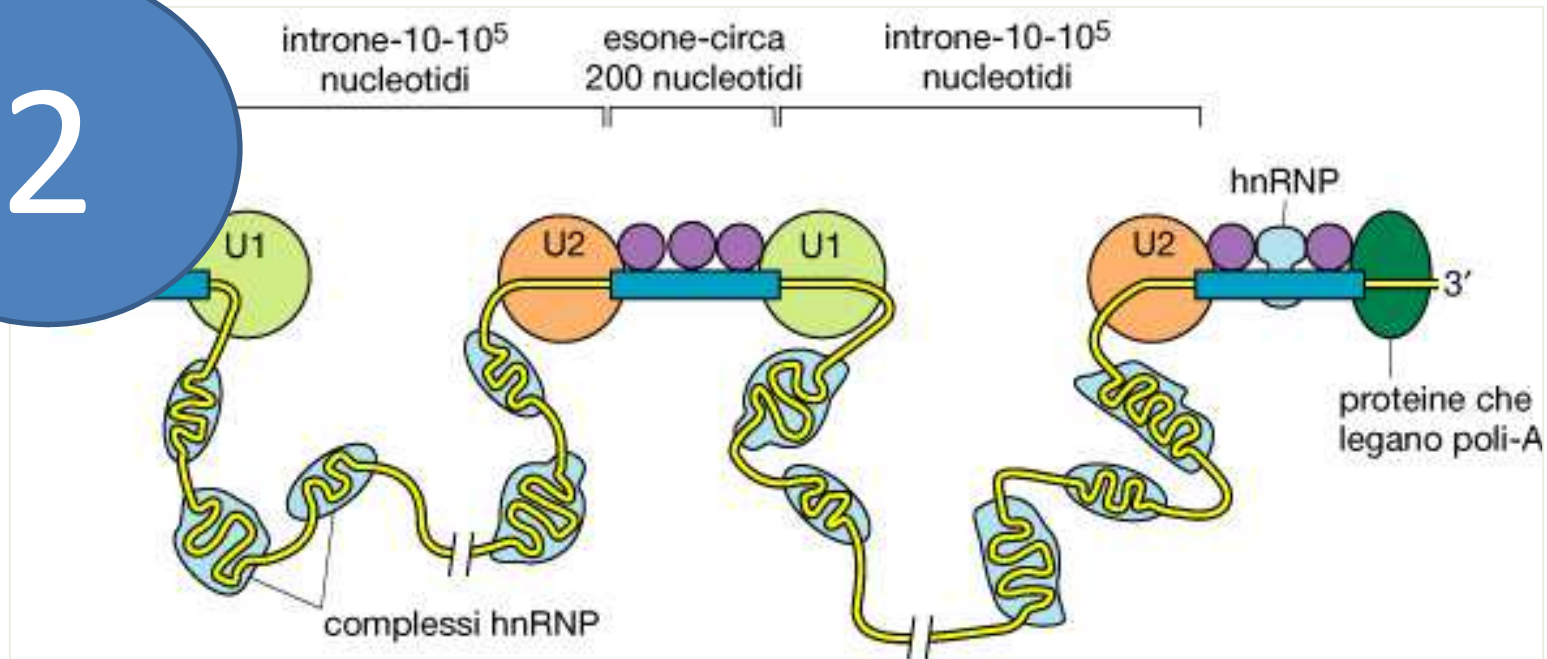
Regulatory sequences are found primarily close to the 5'-ss and 3'-ss i.e.

ISS   ESS   ISS

ESE   ISE
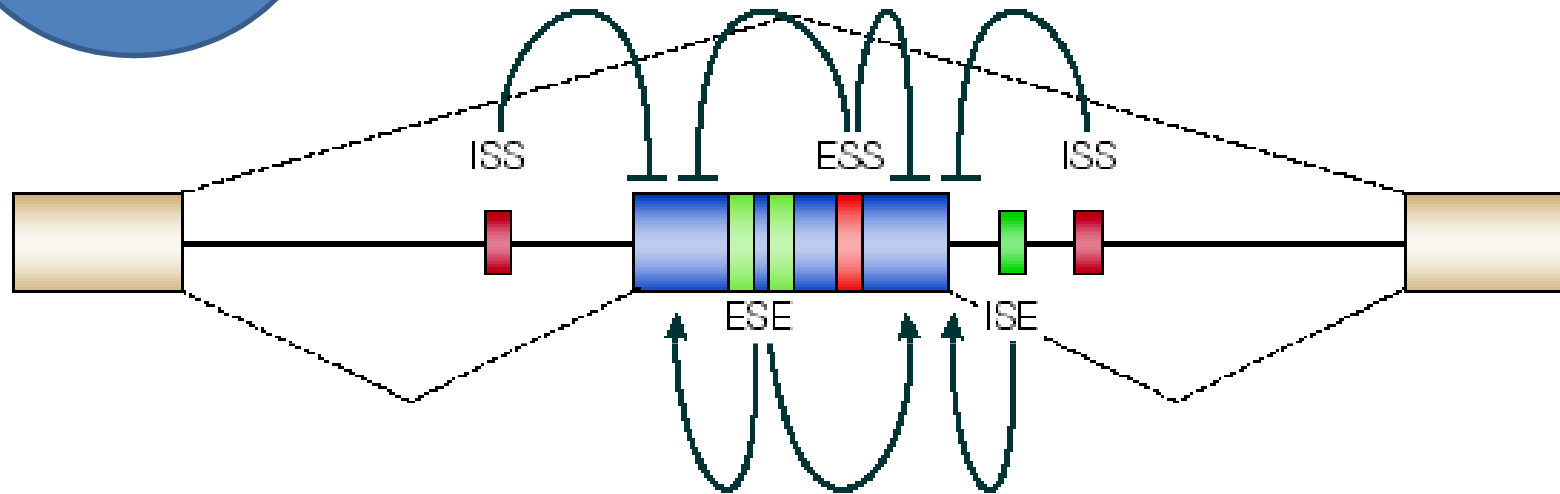
Figure 1 | **Elementary alternative splicing events and regulatory elements. A** | In addition to the splice-site consensus sequences, a number of auxiliary elements can influence alternative splicing. These are categorized by their location and activity as exon splicing enhancers and silencers (ESEs and ESSs) and intron splicing enhancers and silencers (ISEs and ISSs). Enhancers can activate adjacent splice sites or antagonize silencers, whereas silencers can repress splice sites or enhancers. Exon inclusion or skipping is determined by the balance of these competing influences, which in turn might be determined by relative concentrations of the cognate RNA-binding activator and repressor proteins.

From: Matlin et al. (2005), Nature Rev Mol Cell Biol, 6: 386.

Table 2 | **Tissue-specific alternative splicing factors**

| Name	Other names	Binding domain	Binding motif	Tissue expression	Target genes
nPTB		RRM	CUCUCU	Neurons, myoblasts and testes	BIN1, GLYRA2, ATP2B1, MEF2, NASP, SPAG9 and SRC
NOV			YCAY	Neurons of the hindbrain and spinal cord	GABRG2, GLYRA2 and NOVA1
NOV			YCAY	Neurons of the cortex, hippocampus and dorsal spinal cord	KCNJ, APLP2, GPHN, JNK2, NEO, GRIN1 and PLCB4
FOX1		RRM	(U)GCAUG	Muscle, heart and neurons	ACTN, EWSR1, FGFR2, FN1 and SRC
FOX2	RBM9	RRM	(U)GCAUG	Muscle, heart and neurons	EWS, FGFR2, FN1 and SRC
RBM35a	ESRP1	RRM	GU rich	Epithelial cells	FGFR2, CD44, CTNND1 and ENAH
RBM35b	ESRP2	RRM	GU rich	Epithelial cells	FGFR2, CD44, CTNND1 and ENAH
TIA1	mTIA1	RRM	U rich	Brain, spleen and testes	MYPT1, CD95, CALCA, FGFR2, TIAR, IL8, VEGF, NF1 and COL2A1
TIAR	TIAL1 and mTIAR	RRM	U rich	Brain, spleen, lung, liver and testes	TIA1, CALCA, TIAR, NF1 and CD95
SLM2	KHDRBS3 and TSTAR	KH	UAAA	Brain, tests and heart	CD44 and VEGFA
Quaking	QK and QKL	KH	ACUAAY[…]UAAY	Brain	MAG and PLP
HUB	HUC, HUD and ELAV2	RRM	AU rich	Neurons	CALCA, CD95 and NF1
MBNL	NA	CCCH zinc finger domain	YGCU(U/G)Y	Muscles, uterus and ovaries	TNTT2, INSR, CLCN1 and TNNT3
CELF1	BRUNOL2	RRM	U and G rich	Brain	TNTT2 and INSR
ETR3	CELF2 and BRUNOL3	RRM	U and G rich	Heart, skeletal muscle and brain	TNTT2, TAU and COX2
CELF4	BRUNOL4	RRM	U and G rich	Muscle	MTMR1 and TNTT2
CELF5	BRUNOL5 and NAPOR	RRM	U and G rich	Heart, skeletal muscle and brain	ACTN, TNTT2 and GRIN1
CELF6	BRUNOL6	RRM	U and G rich	Kidney, brain and testes	TNTT2

A2BP1, ataxin 2-binding protein 1; ACTN, α-actinin; APLP2, amyloid-β precursor-like protein 2; ATP2B1, ATPase, Ca²⁺ transporting, plasma membrane 1; BIN1, bridging integrator 1; CALCA, calcitonin-related polypeptide-α; CELF, CUGBP- and ETR3-like factor; CLCN1, chloride channel 1; COL2A1, collagen, type II, α1; COX2, cytochrome c oxidase II; CTNND1, catenin δ1, EWSR1, Ewing sarcoma breakpoint region 1; FGFR2, fibroblast growth factor receptor 2; FN1, fibronectin 1; GABRG2, GABA A receptor, γ2; GLYRA2, glycine receptor, α2 subunit; GPHN, gephyrin; GRIN1, glutamate receptor, ionotropic, NMDA 3B; IL8, interleukin-8; INSR, insulin receptor; JNK2, Jun N-terminal kinase 2; KCNJ, potassium inwardly-rectifying channel, subfamily; KHDRBS3, KH domain-containing, RNA-binding, signal transduction-associated protein 3; MAG, myelin associated glycoprotein; MBNL, muscleblind; MEF2, myocyte enhancing factor 2; MTMR1, myotubularin-related protein 1; NASP, nuclear autoantigenic sperm protein; NEO, neogenin; NF1, neurofibromin 1; NOVA, neuro-oncological ventral antigen; PLCB4, phospholipase C β4; PLP, proteolipid protein; PTB, polypyrimidine-tract binding protein; RBM, RNA-binding protein; RRM, RNA recognition motif; SLM2, SAM68-like mammalian protein 2; SPAG9, sperm associated antigen 9; TIA1, T cell-restricted intracellular antigen 1; TIAR, TIA1-related protein; TNTT2, troponin T type 2; VEGF, vascular endothelial growth factor.

model: competition between SR proteins and hnRNP. (hnRNP A1 multimerizes)
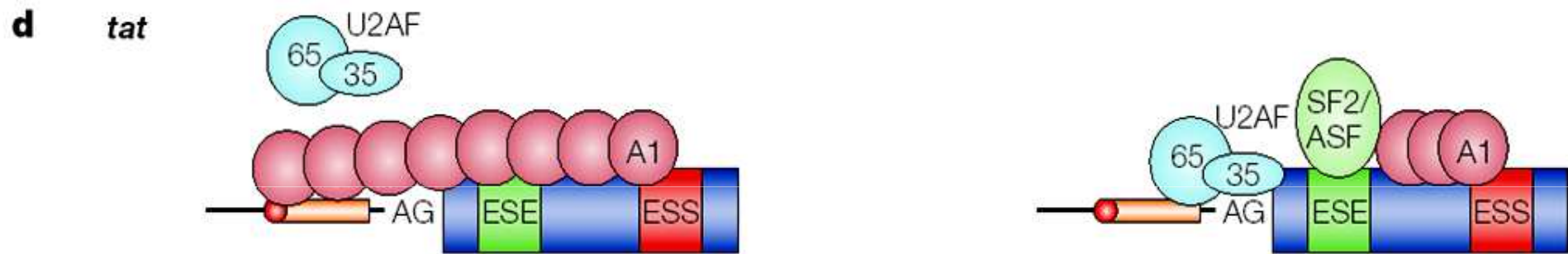
d | Inclusion of exon 3 of HIV1 *tat* pre-mRNA is determined by the nuclear ratio of specific heterogeneous nuclear ribonucleoprotein (hnRNP) and SR proteins. Propagative multimerization of hnRNPA1 from a high-affinity exon splicing silencer (ESS) is sterically blocked by the interaction of SF2/ASF with the upstream ESE. In this case, ESE function requires the RRM domains but not the RS domain of SF2/ASF.

model: an RNA-binding protein interacts with a small sequence in the pre-mRNA using a RRM domain and interacts with U1snRNP with another domain

b *FAS*

b | A weak 5' splice site in the *FAS* transcript is enhanced by TIA1 binding to a downstream intron splicing enhancer (ISE). TIA1 cooperatively promotes the interaction of U1 small nuclear ribonucleoprotein particles (snRNPs) with the pre-mRNA.

model: binding of a protein to a intronic sequence overlapping polypyrimidine competes with U2AF65 binding and inhibits splicing

c | Repression of the non-sex-specific *tra* 3' splice site involves the interaction of SXL with an intron splicing silencer (ISS) embedded in the polypyrimidine tract and the prevention of U2AF binding. This leads to selection of the downstream female-specific 3' splice site.

model: expression of a tissue-specific paralogue of the PTB (polypyrimidine tract binding protein allows intron definition
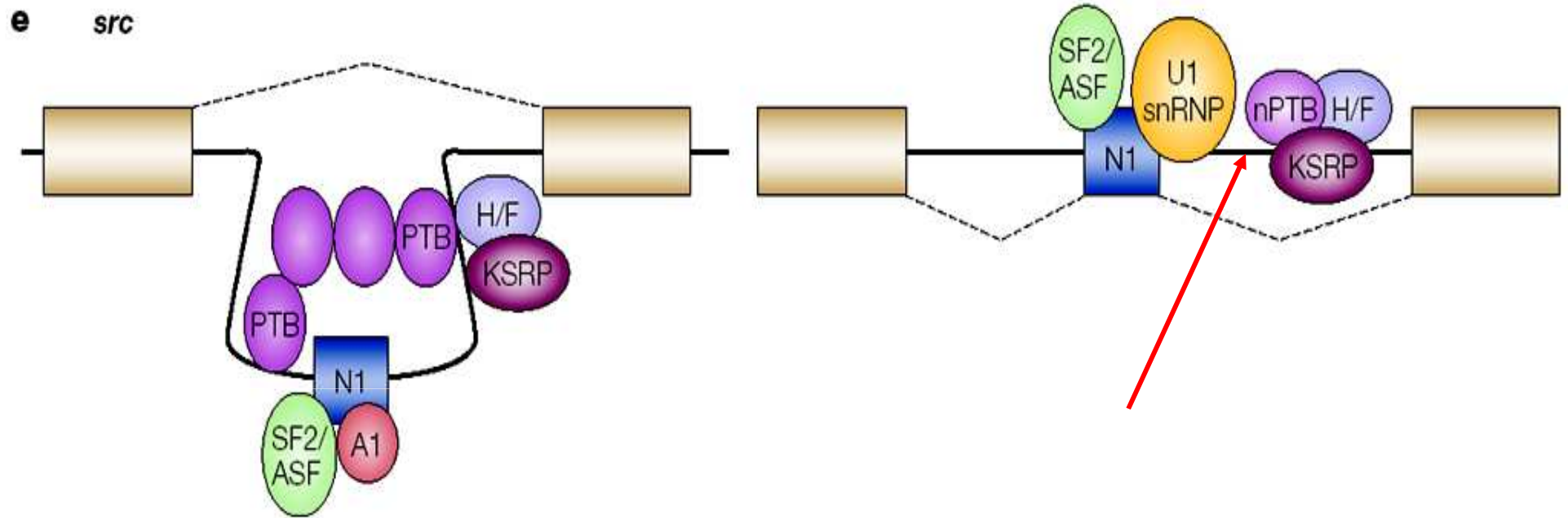
e | The regulation of N1 exon splicing in the *src* transcript provides an example of combinatorial control by cooperation and antagonism between numerous positively and negatively acting factors. In non-neuronal cells (left), N1 is excluded, whereas in neurons (right), it is included in the mature mRNA. Constitutive exons are shown as beige boxes, whereas alternative exons are shown as blue boxes. KSRP, KH-type splicing regulatory protein; nPTB, neural polypyrimidine tract binding protein.
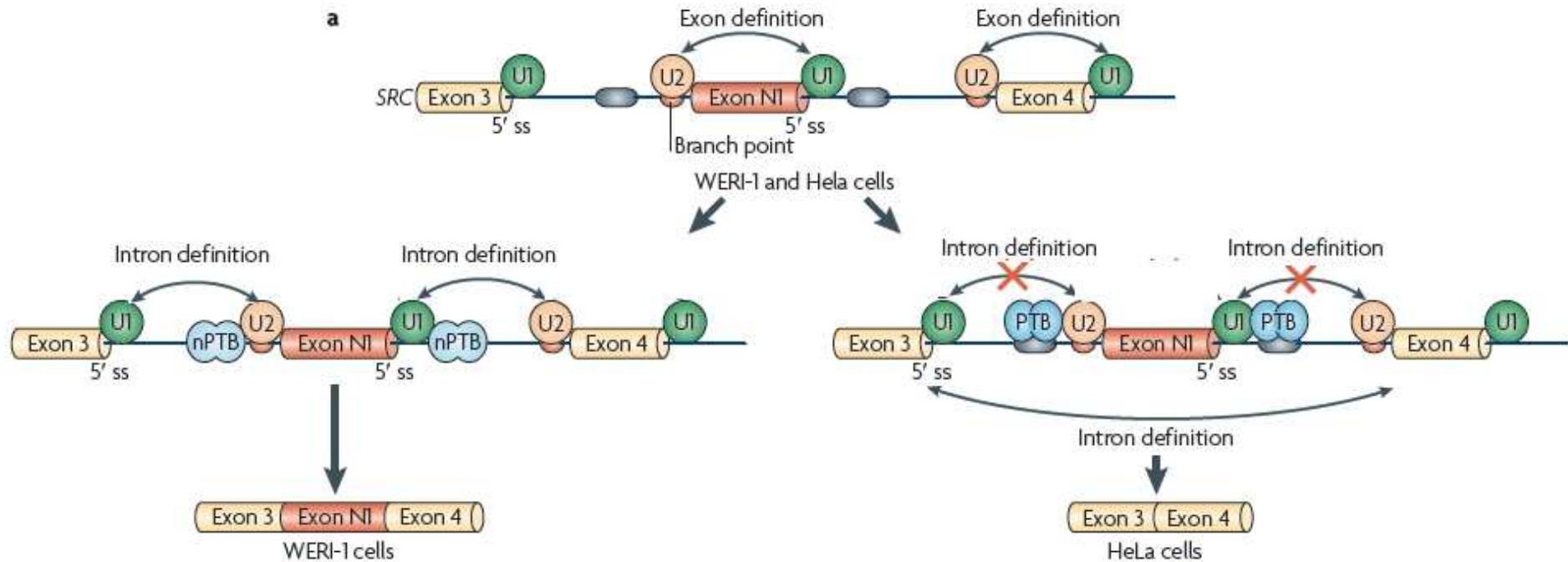
Figure 3 | Mechanisms of alternative splicing regulation at the transition from exon definition to intron definition. a | Polypyrimidine-tract binding protein (PTB) inhibits the inclusion of SRC exon N1 by inhibiting the interactions of the U1 and U2 small nuclear ribonucleoprotein particles (snRNPs) and intron definition. In both WERI-1 and HeLa cells, the N1 exon is defined by the binding of U1 snRNP to the 5′ splice site (ss) and of U2 snRNP to the branch point. In WERI-1 cells, in the absence of PTB, U1 and U2 snRNPs bound to the N1 exon interact with the U2 and U1 snRNP on adjacent constitutive exons, respectively, thereby allowing efficient spliceosome assembly on introns flanking exon N1. In HeLa cells, PTB binds to sequences flanking exon N1 and prevents the cross-intron interactions that occur in WERI- cells, thereby excluding exon N1.

from Chen & Manley 2009. Nat Rev Mol Cell Biol., 10:741., modified

Slide 10

YUN1 Your User Name, 1/9/2011

ARTICLES

An RNA map predicting Nova-dependent splicing regulation

Jernej Ule[1,2]*†, Giovanni Stefani[1,2]*†, Aldo Mele[1,2], Matteo Ruggiu[1,2], Xuning Wang[3], Bahar Taneri[4]†, Terry Gaasterland[4]†, Benjamin J. Blencowe[5] & Robert B. Darnell[1,2]

Nova proteins are neuron-specific alternative splicing factors. We have combined bioinformatics, biochemistry and genetics to derive an RNA map describing the rules by which Nova proteins regulate alternative splicing. This map revealed that the position of Nova binding sites (YCAY clusters) in a pre-messenger RNA determines the outcome of splicing. The map correctly predicted Nova's effect to inhibit or enhance exon inclusion, which led us to examine the relationship between the map and Nova's mechanism of action. Nova binding to an exonic YCAY cluster changed the protein complexes assembled on pre-mRNA, blocking U1 snRNP (small nuclear ribonucleoprotein) binding and exon inclusion, whereas Nova binding to an intronic YCAY cluster enhanced spliceosome assembly and exon inclusion. Assays of splicing intermediates of Nova-regulated transcripts in mouse brain revealed that Nova preferentially regulates removal of introns harbouring (or closest to) YCAY clusters. These results define a genome-wide map relating the position of a cis-acting element to its regulation by an RNA binding protein, namely that Nova binding to YCAY clusters results in a local and asymmetric action to regulate spliceosome assembly and alternative splicing in neurons.

Cross-link (U.V.)
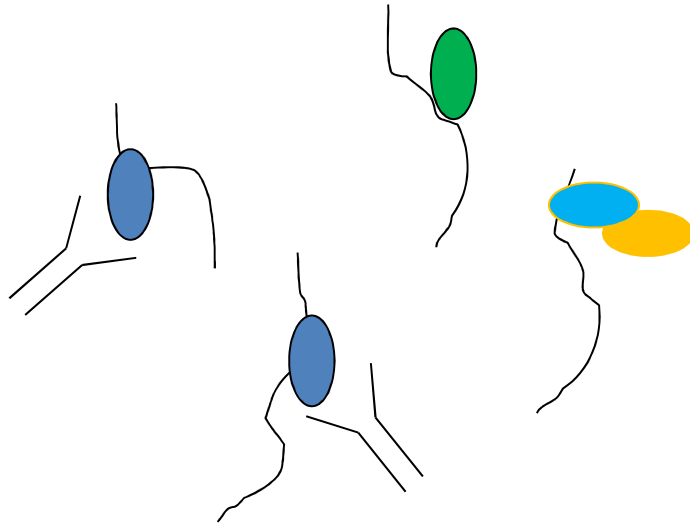
Nova: the first vertebrate tissue-specific splicing factors (neurons) (Nova1 – Nova2)

CLIP (cross-link immunoprecipitation) (Assignment BMG11)

IMPT

RNA identified on microarrays or sequenced

48 targets identified in previous studies

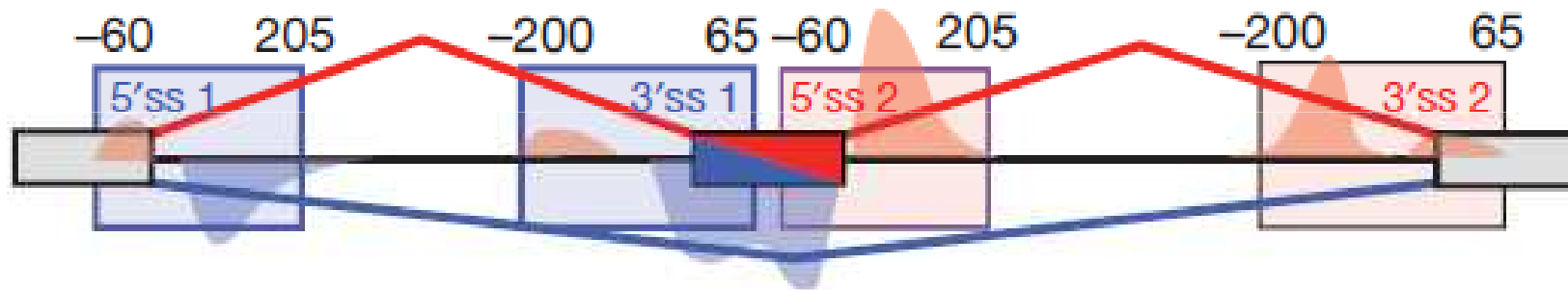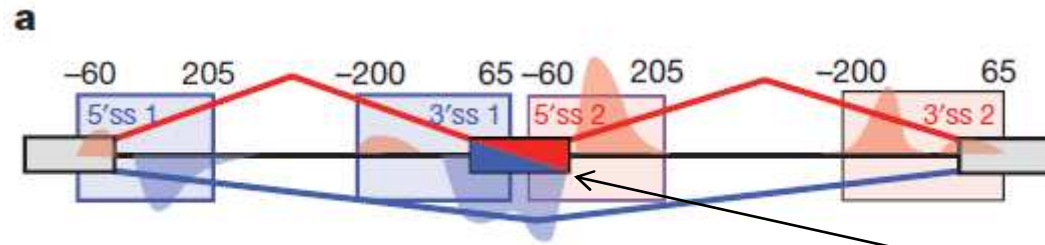Clustering of "YCAY" Nova recognition sequences in 48 Nova-regulated exons

Figure 1 | Definition of the Nova–RNA binding map. a, A generic premRNA showing the four regions that define the Nova–RNA binding map (the start and end of each region is labelled by a nucleotide distance to the splice site). Peaks demonstrate the positions of Nova-dependent splicing enhancers (red) or silencers (blue).

a

5'ss 1 | 3'ss 1 | 5'ss 2 | 3'ss 2
−60 | 205 | −200 | 65 −60 | 205 | −200 | 65

Alternative exon

b

P = 0.035 (NESE1)   *P* = 0.027 NISS1   (NISE1)   *P* = 0.047 NISS2/NESS1

5'ss 1   3'ss 1

Constitutive   Alternative

P = 0.012 NESS2   *P* < 0.001 NISE2   *P* = 0.002 NISE3

5'ss 2   3'ss 2

Inclusion
Skipping
Control

Alternative   Const.

YCAY cluster score — Position (nucleotides)
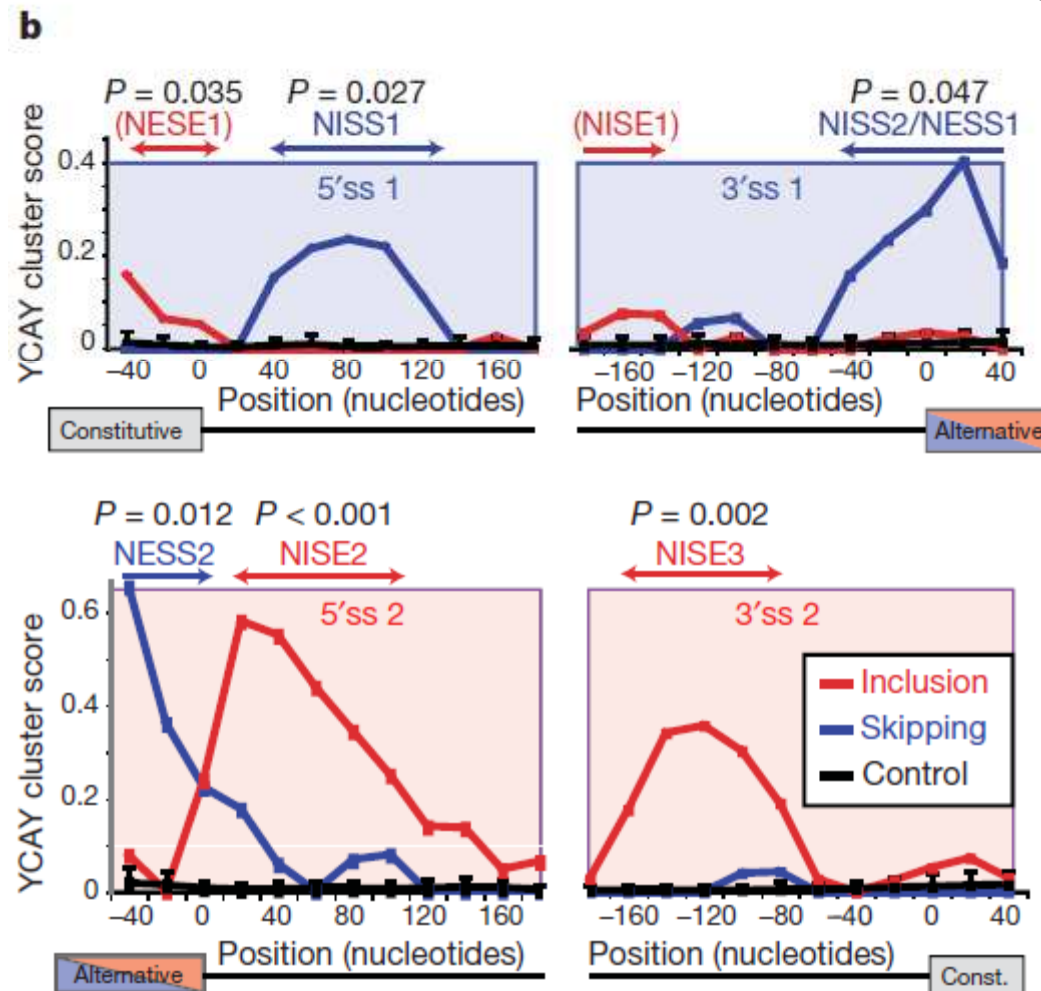
Scoring YACY clusters (median 28-base long) in 200-bp boxes at 5'- and 3'- SS
Control: genome-wide positions
Nova-independent

Using these observations, Aa. construct a "score" of Nova map that can predict, for each exon, the probability of being regulated (enhanced or silenced) by Nova1 & 2. This prediction was then verified on a group of exons of different genes, by comparing exon inclusion/exclusion between wt mice and Nova1,2$^{-/-}$ mice:

In K.O. mice brain

Red are intronic sites

Blue are exonic sites or adjacent to ss

Expression ratio
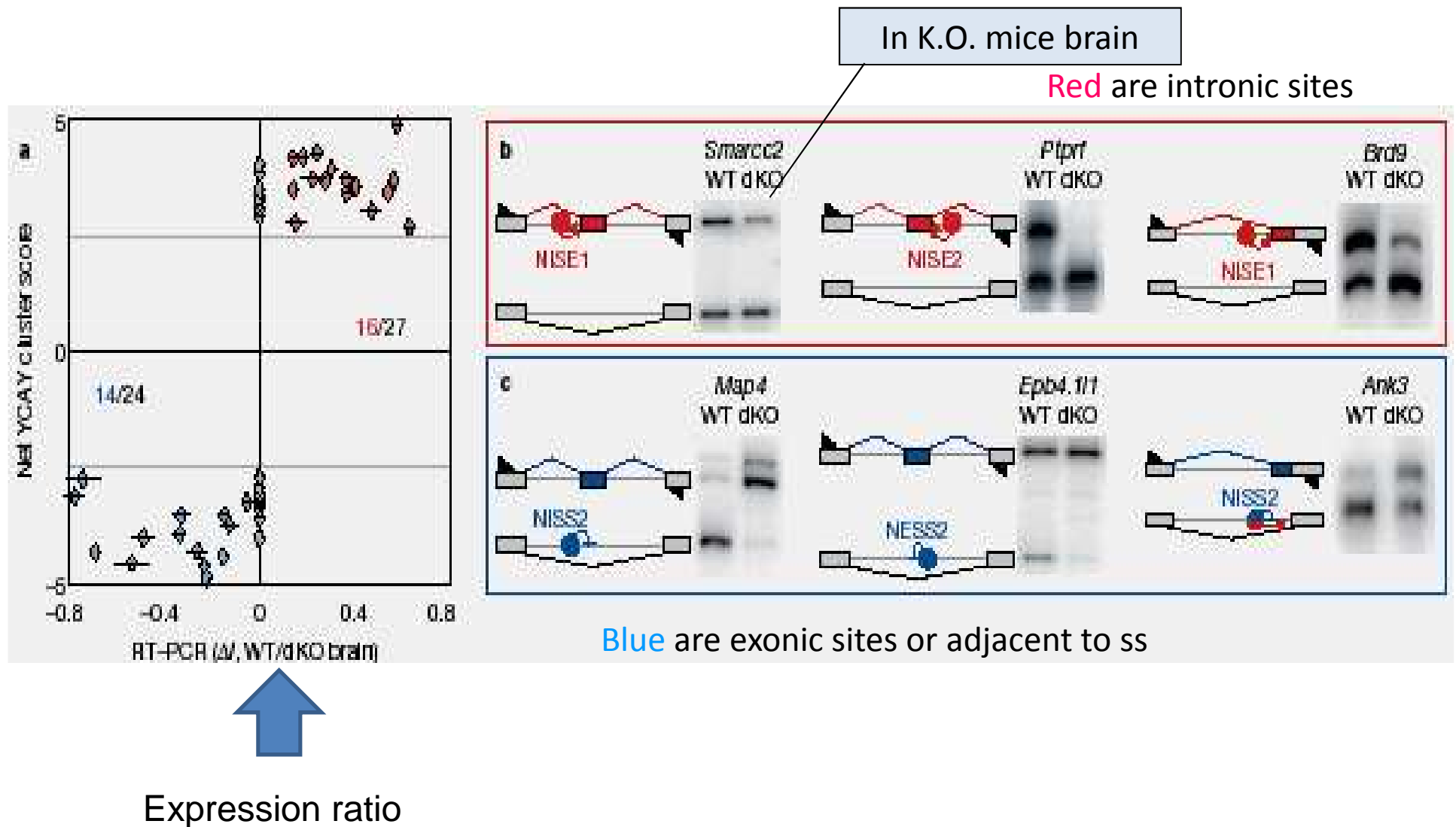
Genome-wide scanning with the Nova-score definition finds out 51 new candidate exons
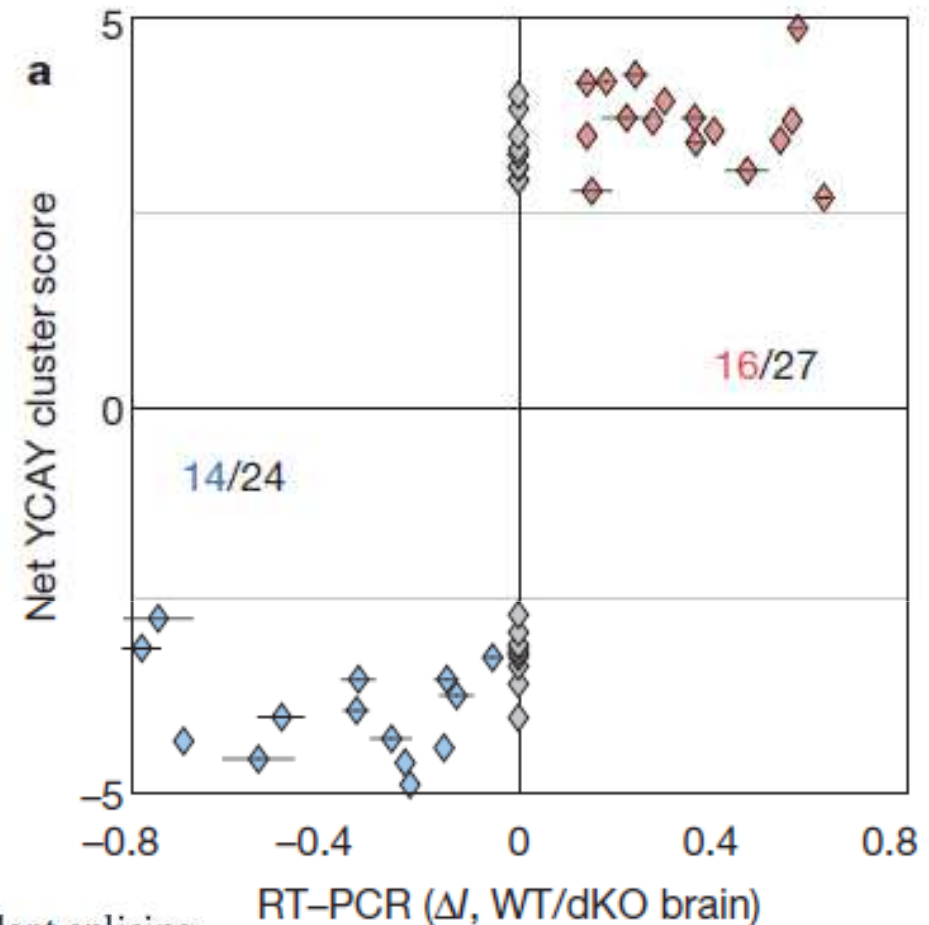
To test whether the RNA map can predict Nova-dependent splicing regulation *de novo*, we calculated a net YCAY cluster score by subtracting Nova silencer from enhancer cluster scores (Supplementary Fig. 4). Using a stringent scoring method ($|$net YCAY cluster score$| > 2.7$), we identified 51 candidate Nova-regulated alternative exons in a genomic database of bioinformatically predicted alternative exons (B.T. and T.G., personal communication). Ten previously validated Nova-regulated exons were among these top predictions.

Examples of predicted Nova-regulated exons: analysis in brain tissues from Nova1$^{-/-}$/ Nova2 $^{-/-}$ double K.O. mice.
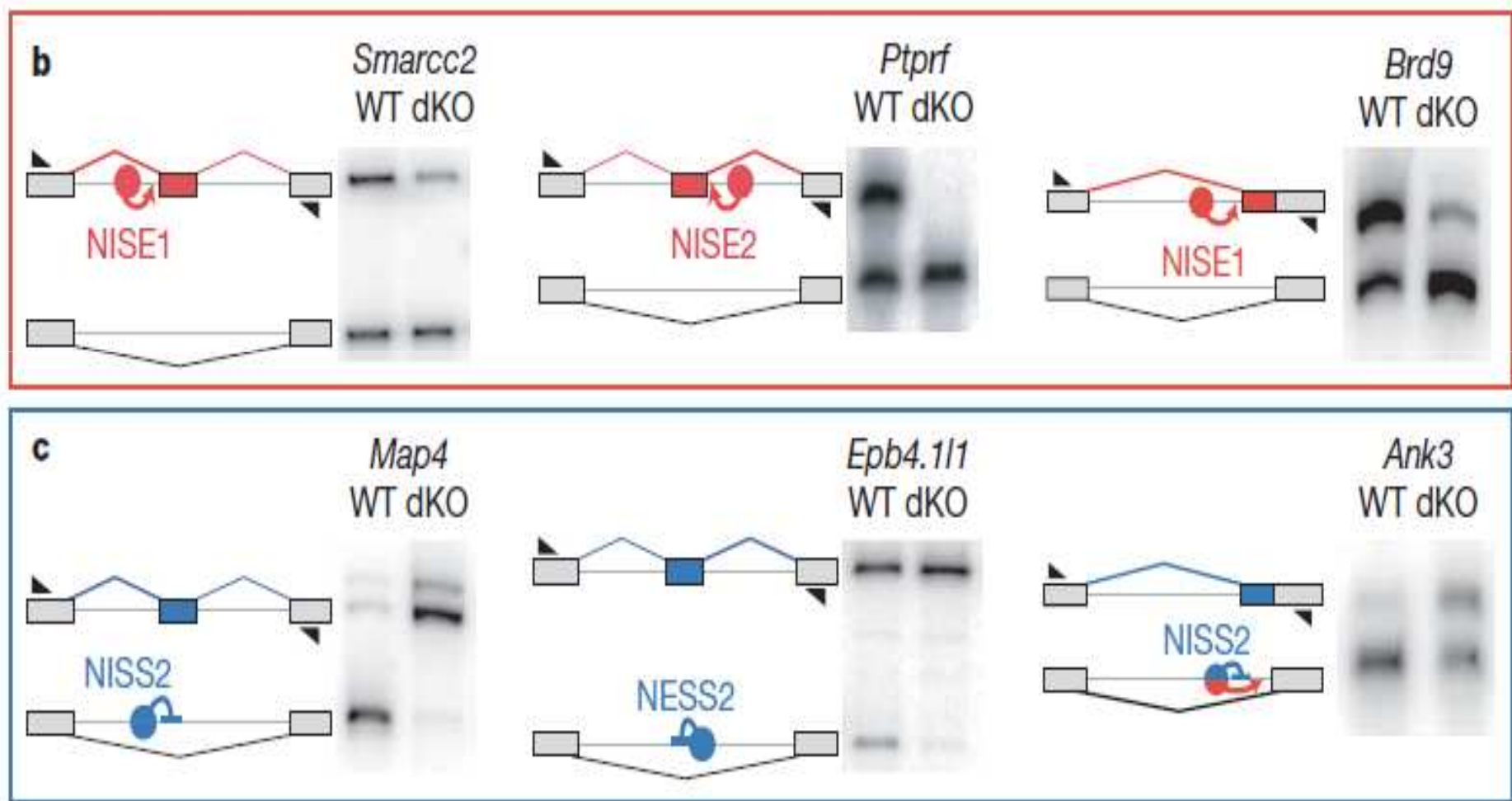
Mechanism: by studying the different steps in vitro on selected alternative exons, the authors demonstrated that Nova regulates assembly of the early spliceosomal complex

ESRP1 and ESRP2 Are Epithelial Cell-Type-Specific Regulators of FGFR2 Splicing

Claude C. Warzecha,[1,2] Trey K. Sato,[3] Behnam Nabet,[1] John B. Hogenesch,[3,5] and Russ P. Carstens[1,2,4,*]
[1]Renal Division, Department of Medicine
[2]Cell and Molecular Biology Graduate Group
[3]Department of Pharmacology, Institute for Translational Medicine and Therapeutics
[4]Department of Genetics
University of Pennsylvania School of Medicine, Philadelphia, PA 19104, USA
[5]Penn Genome Frontiers Institute, University of Pennsylvania, Philadelphia, PA 19104, USA

SUMMARY

Cell-type-specific expression of epithelial and mesenchymal isoforms of Fibroblast Growth Factor Receptor 2 (FGFR2) is achieved through tight regulation of mutually exclusive exons IIIb and IIIc, respectively. Using an application of cell-based cDNA expression screening, we identified two paralogous epithelial cell-type-specific RNA-binding proteins that are essential regulators of *FGFR2* splicing. Ectopic expression of either protein in cells that express FGFR2-IIIc caused a switch in endogenous FGFR2 splicing to the epithelial isoform. Conversely, knockdown of both factors in cells that express FGFR2-IIIb by RNA interference caused a switch from the epithelial to mesenchymal isoform. These factors also regulate splicing of *CD44*, *p120-Catenin* (*CTNND1*), and *hMena* (*ENAH*), three transcripts that undergo changes in splicing during the epithelial-to-mesenchymal transition (EMT). These studies suggest that Epithelial Splicing Regulatory Proteins 1 and 2 (ESRP1 and ESRP2) are coordinators of an epithelial cell-type-specific splicing program.
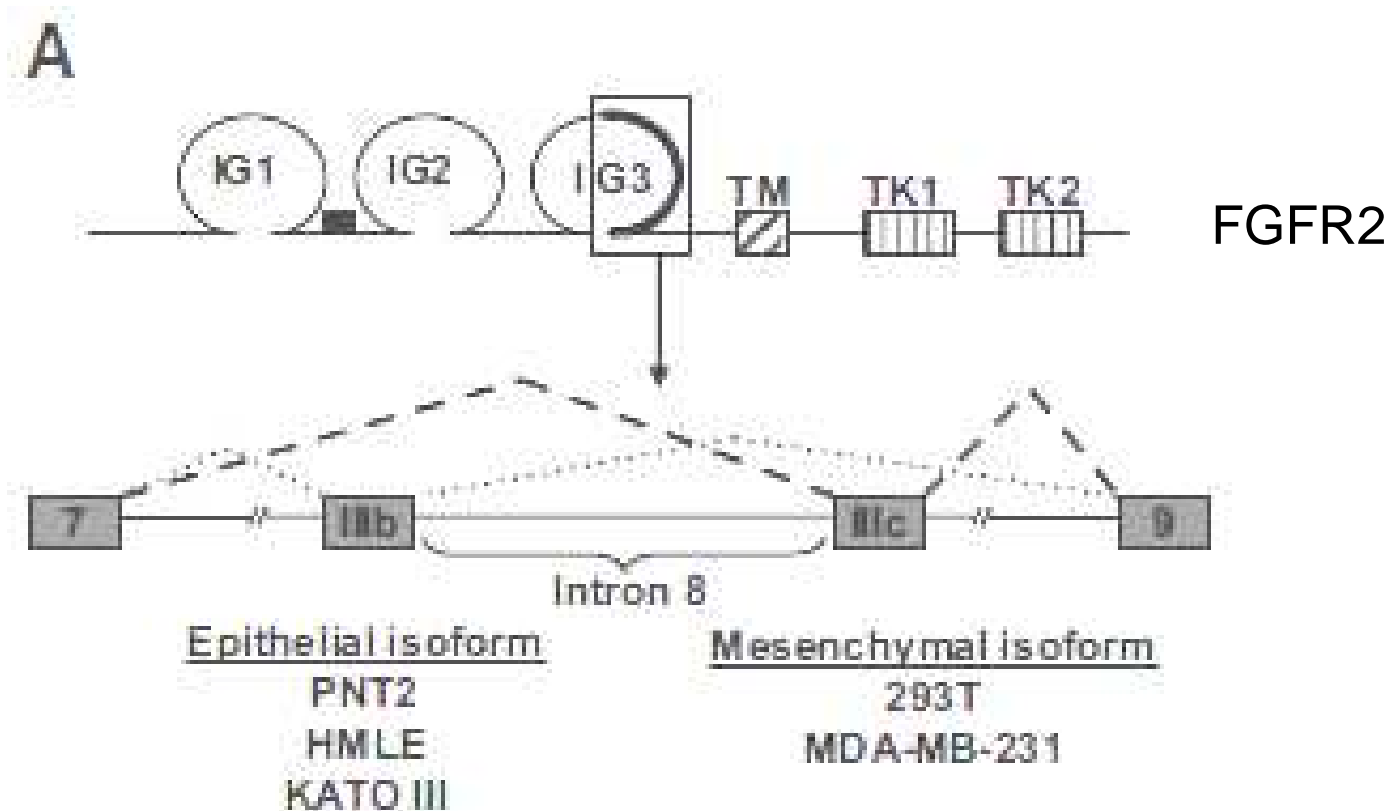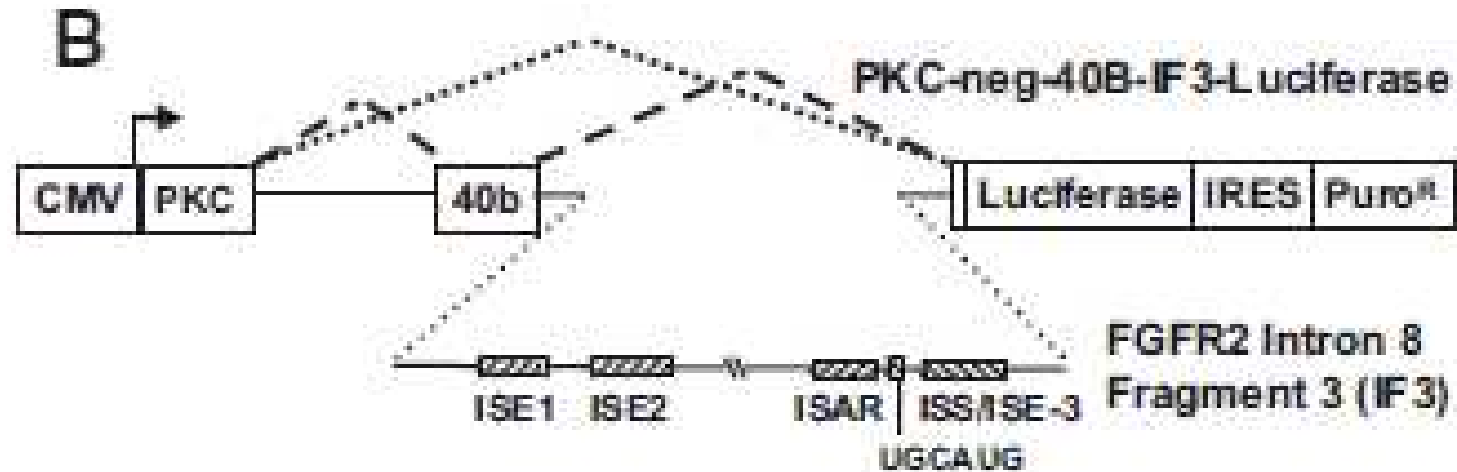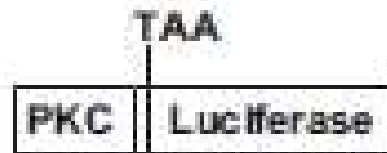
Figure 1. Identification of FGFR2 Splicing Regulators Rbm35a, Esrp1, and Rbm35b, Esrp2, in a High-Throughput cDNA Expression Screen

(A) Schematic of the FGFR2 protein (top) and the pre-mRNA in the region encoding exons IIIb and IIIc. IG, immunoglobulin-like domains; TM, transmembrane domain; TK, tyrosine kinase domains.
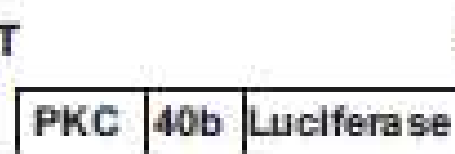
B

PKC-neg-40B-IF3-Luciferase

| CMV | PKC | | 40b | | | | Luciferase | IRES | Puro^R |

FGFR2 Intron 8
Fragment 3 (IF3)

ISE1 ISE2   ISAR   ISS/ISE-3

UGCAUG

Splicing Outcomes:

TAA

Mesenchymal cells:

| PKC | | Luciferase |

Exon 40B Skipped:
Negative luminescence

TAA

Epithelial cells or 293T
cells expressing
epithelial splicing
regulator:

| PKC | 40b | Luciferase |

Exon 40B inclusion:
Positive luminescence

(B) Schematic of the reporter minigene and screening strategy. An FGFR2 intron 8 fragment (IF3) required for IIIb inclusion in epithelial cells is positioned downstream of a heterologous exon.   Auxiliary intronic cis-elements are indicated by hatched boxes.

High-throughput retrotransfection of 293T
cells stably expressing the heterologous
luciferase reporter in duplicate.
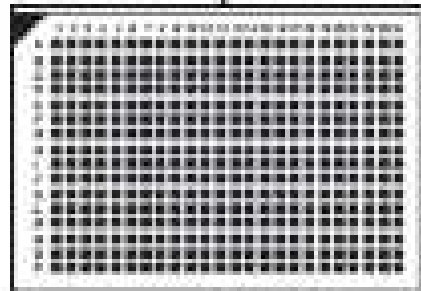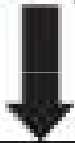
Prearrayed cDNA matrix
 in 384 well plates containing
>14,000 human or mouse
cDNAs in pCMV-Sport6.

Incubate at 37°C for
~60 hours; assay for
increased luciferase
activity.
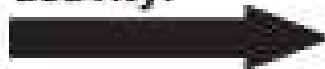
Analyze data, pick
clones, sequence
confirm cDNA
identity, and validate.
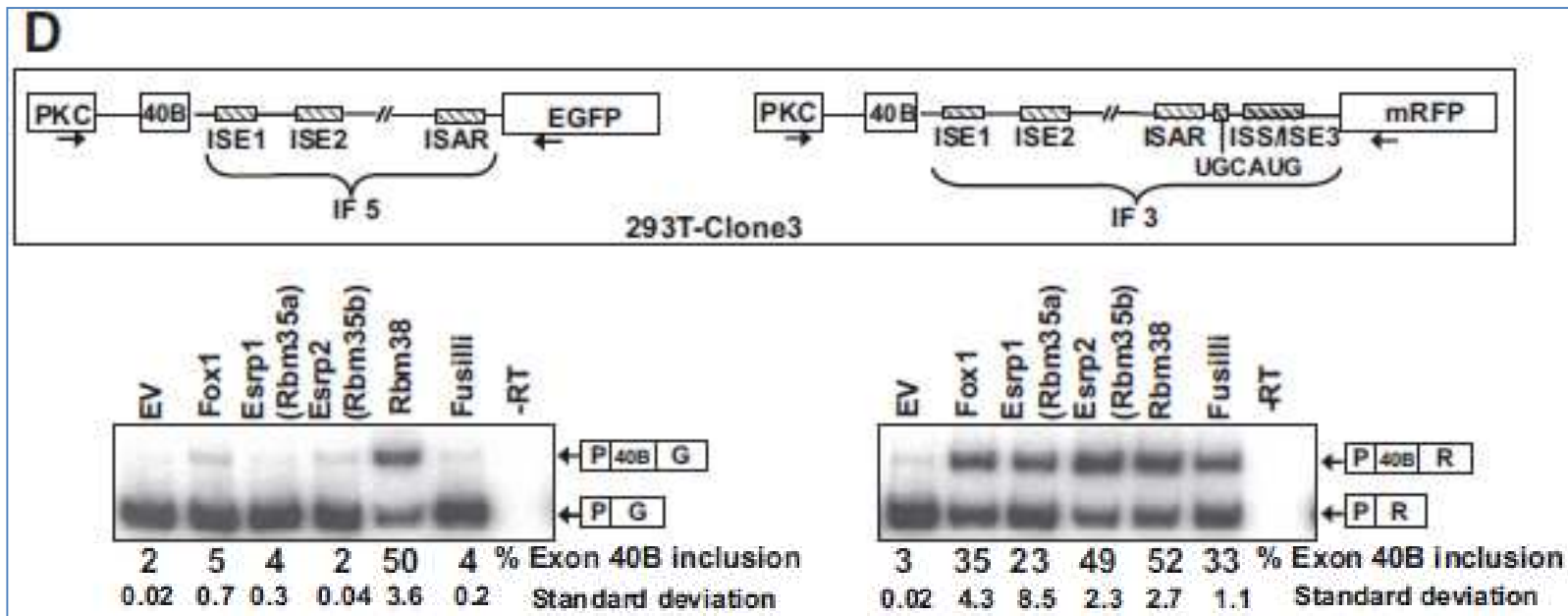
(D) Schematic of the minigenes stably expressed in 293T-clone 3 (top). Cells were transiently transfected with empty vector (EV) or cDNAs for Fox-1, Rbm35a (Esrp1), Rbm35b (Esrp2), Rbm38, or Fusilli, and exon inclusion was determined by RT-PCR. Average percentages of exon 40B inclusion with standard deviations compiled from three experiments are indicated below a representative gel. The Rbm35a, Rbm35b, and Rbm 38 cDNAs represented here are the MGC clones from the screening collection.

Great!
So we do have specific splicing factors regulating tissue-specific alternative splicing events....

However... there are other puzzling results

1) The regulation of Dscam mutually exclusive exon choice in neurons

2) The effect of promoter type and enhancers (!) on weak exon inclusion

3) Co-transcriptionality of splicing

4) chromatin effects on splicing (?)

Robust discrimination between self and non-self neurites requires thousands of Dscam1 isoforms

Daisuke Hattori[1], Yi Chen[1], Benjamin J. Matthews[2], Lukasz Salwinski[3], Chiara Sabatti[4], Wesley B. Grueber[5] & S. Lawrence Zipursky[1]

Down Syndrome cell adhesion molecule (Dscam) genes encode neuronal cell recognition proteins of the immuno-globulin superfamily. In Drosophila, Dscam1 generates 19,008 different ectodomains by alternative splicing of three exon clusters, each encoding half or a complete variable immunoglobulin domain. Identical isoforms bind to each other, but rarely to isoforms differing at any one of the variable immunoglobulin domains. Binding between isoforms on opposing membranes promotes repulsion6. Isoform diversity provides the molecular basis for neurite self-avoidance. Self-avoidance refers to the tendency of branches from the same neuron (self-branches) to selectively avoid one another12. To ensure that repulsion is restricted to self branches, different neurons express different sets of isoforms in a biased stochastic fashion. Genetic studies demonstrated that Dscam1 diversity has a profound role in wiring the fly brain. Here we show how many isoforms are required to provide an identification system that prevents non-self branches from inappropriately recognizing each other. Using homologous recombination,we generated mutant animals encoding 12, 24, 576 and 1,152 potential isoforms. Mutant animals with deletions encoding 4,752 and 14,256 isoforms14 were also analysed. Branching phenotypes were assessed in three classes of neurons. Branching patterns improved as the potential number of isoforms increased, and this was independent of the identity of the isoforms. Although branching defects in animals with 1,152 potential isoforms remained substantial, animals with 4,752 isoforms were indistinguishable from wild-type controls. Mathematical modelling studies were consistent with the experimental results that thousands of isoforms are necessary to ensure acquisition of unique Dscam1 identities in many neurons. We conclude that thousands of isoforms are essential to provide neurons with a robust discrimination mechanism to distinguish between self and non-self during self-avoidance.

Mutually Exclusive Splicing of the Insect *Dscam* Pre-mRNA Directed by Competing Intronic RNA Secondary Structures
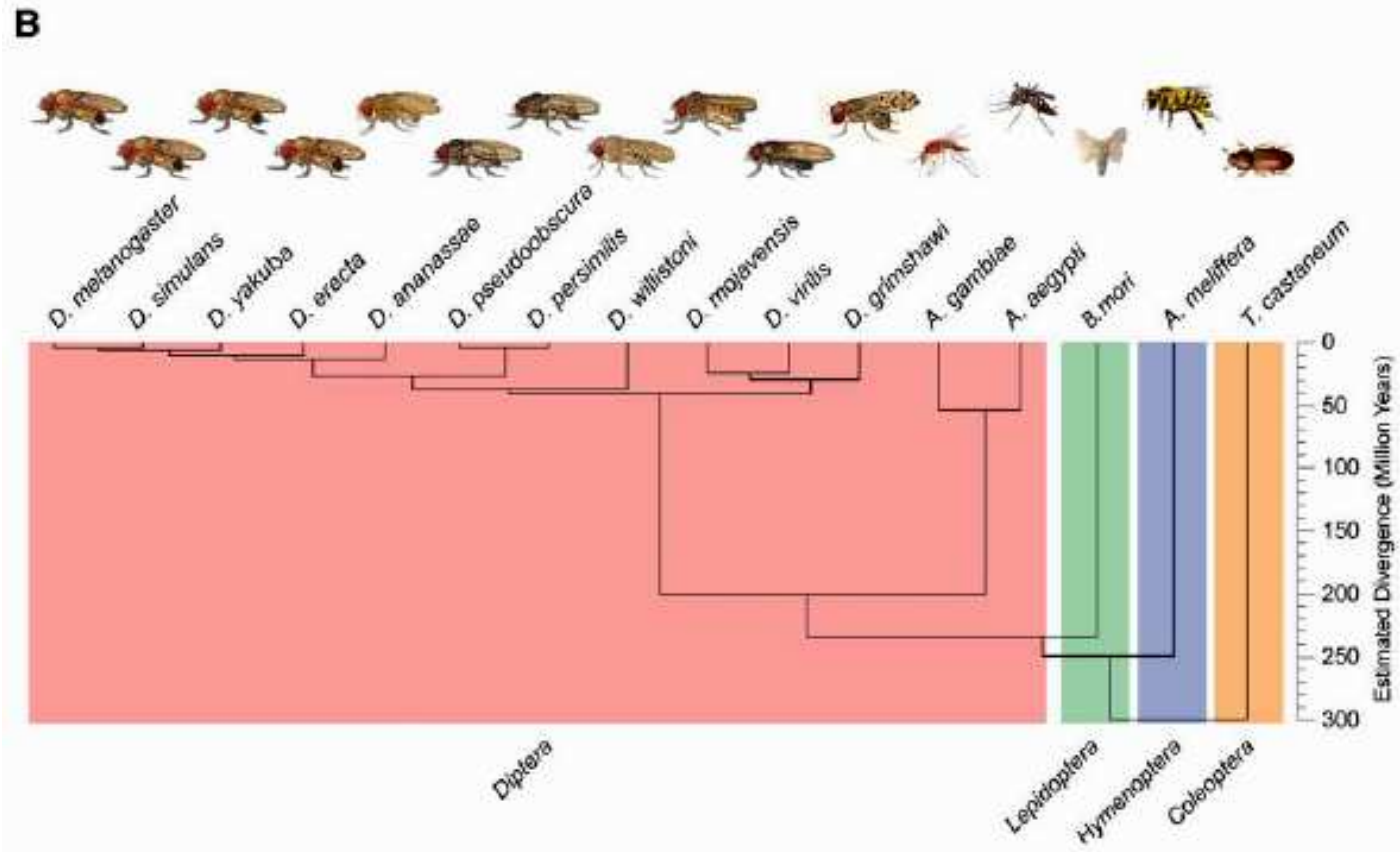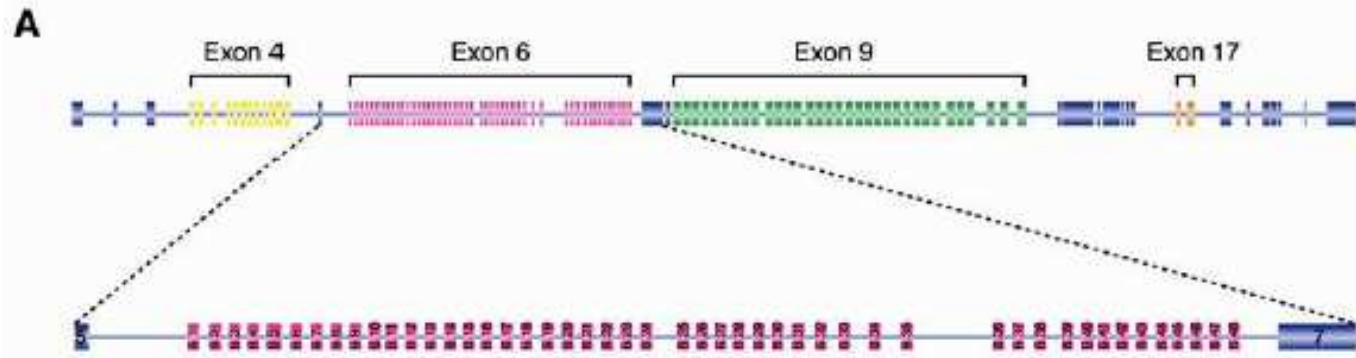
Brenton R. Graveley*
Department of Genetics and Developmental Biology
University of Connecticut Health Center
263 Farmington Avenue
Farmington, Connecticut 06030

Assigned paper

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.

A

Exon 4   Exon 6   Exon 9   Exon 17

B

D. melanogaster
D. simulans
D. yakuba
D. erecta
D. ananassae
D. pseudoobscura
D. persimilis
D. willistoni
D. mojavensis
D. virilis
D. grimshawi
A. gambiae
A. aegypti
B. mori
A. mellifera
T. castaneum

Estimated Divergence (Million Years)

0
50
100
150
200
250
300

Diptera   *Lepidoptera*   *Hymenoptera*   *Coleoptera*

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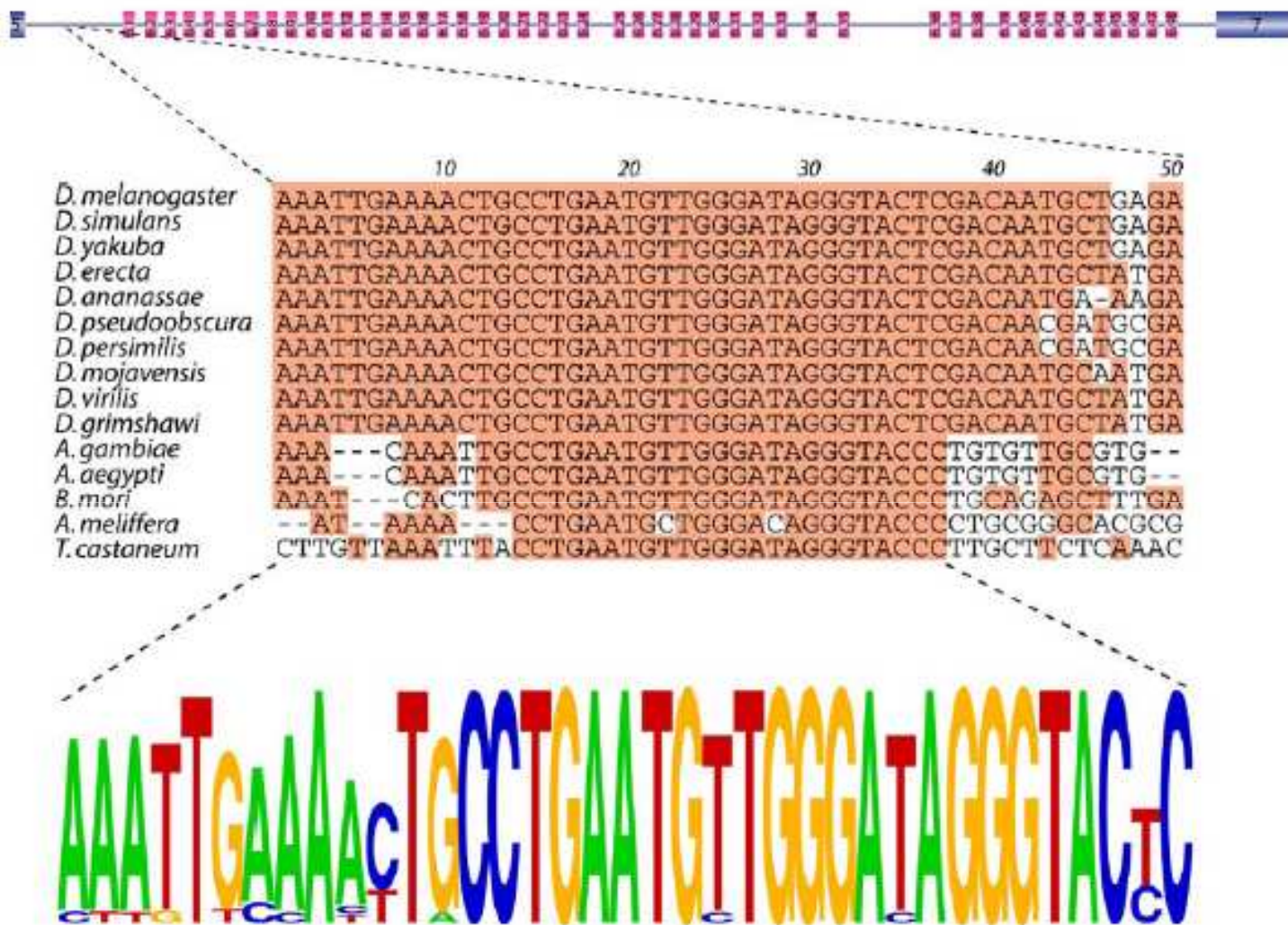

Figure 2. The Docking Site

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.
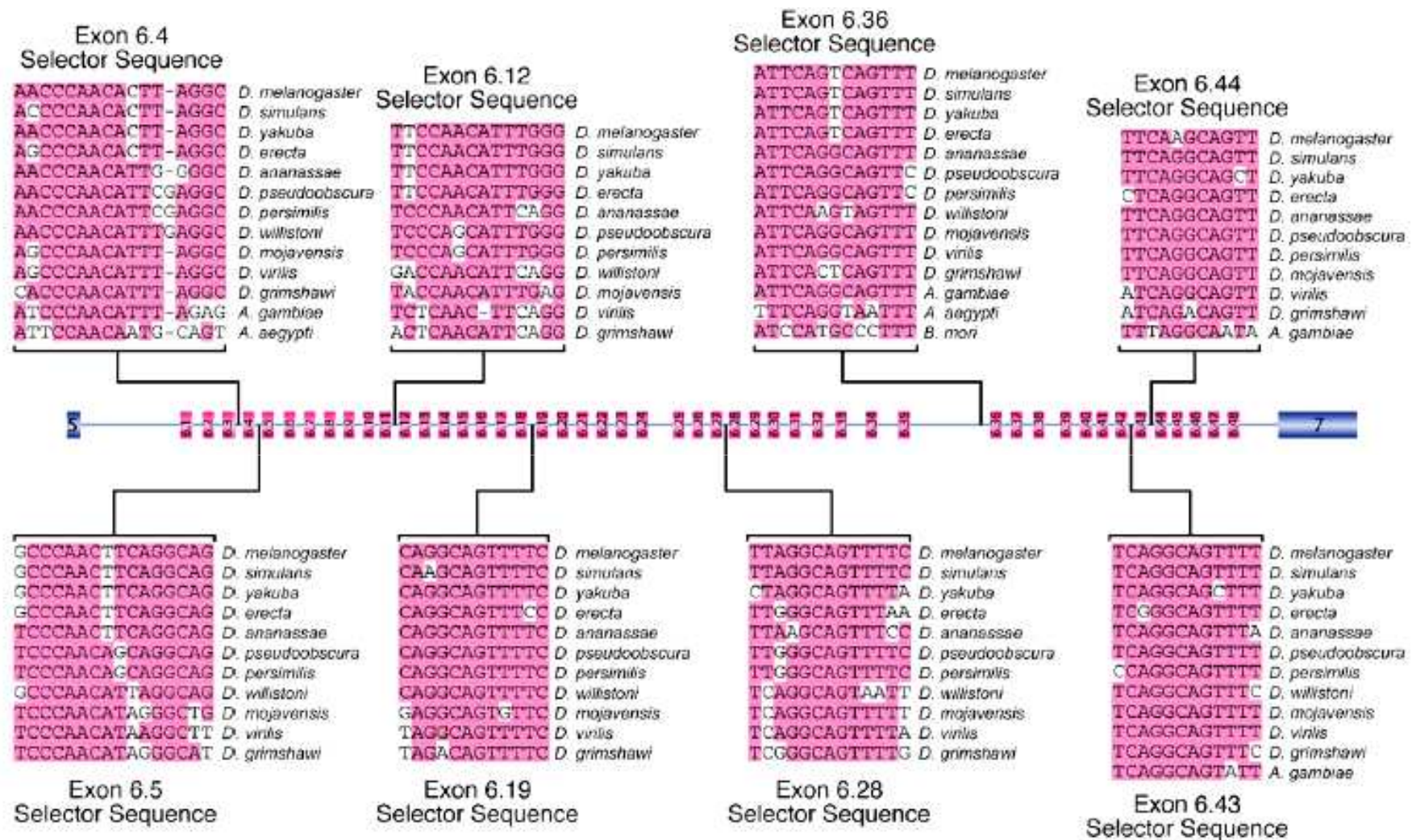

Figure 3. Conservation of Selector Sequences

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.
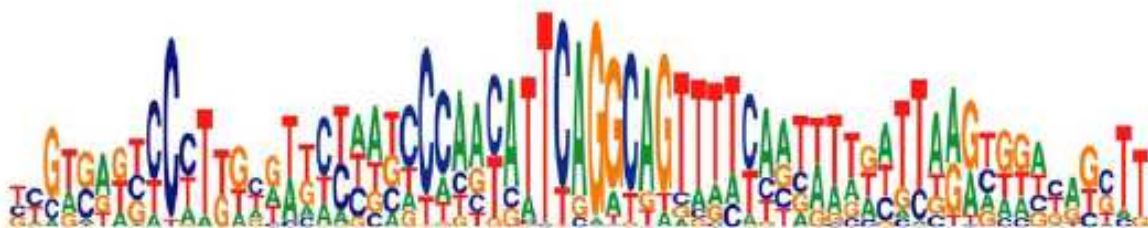
Figure 4. The *D. melanogaster* Selector Sequence Consensus
(A) The 48 selector sequences and flanking sequence were aligned together. The most frequent nucleotides in the central portion of the
alignment are highlighted.
(B) The alignment was used to generate a selector sequence consensus.
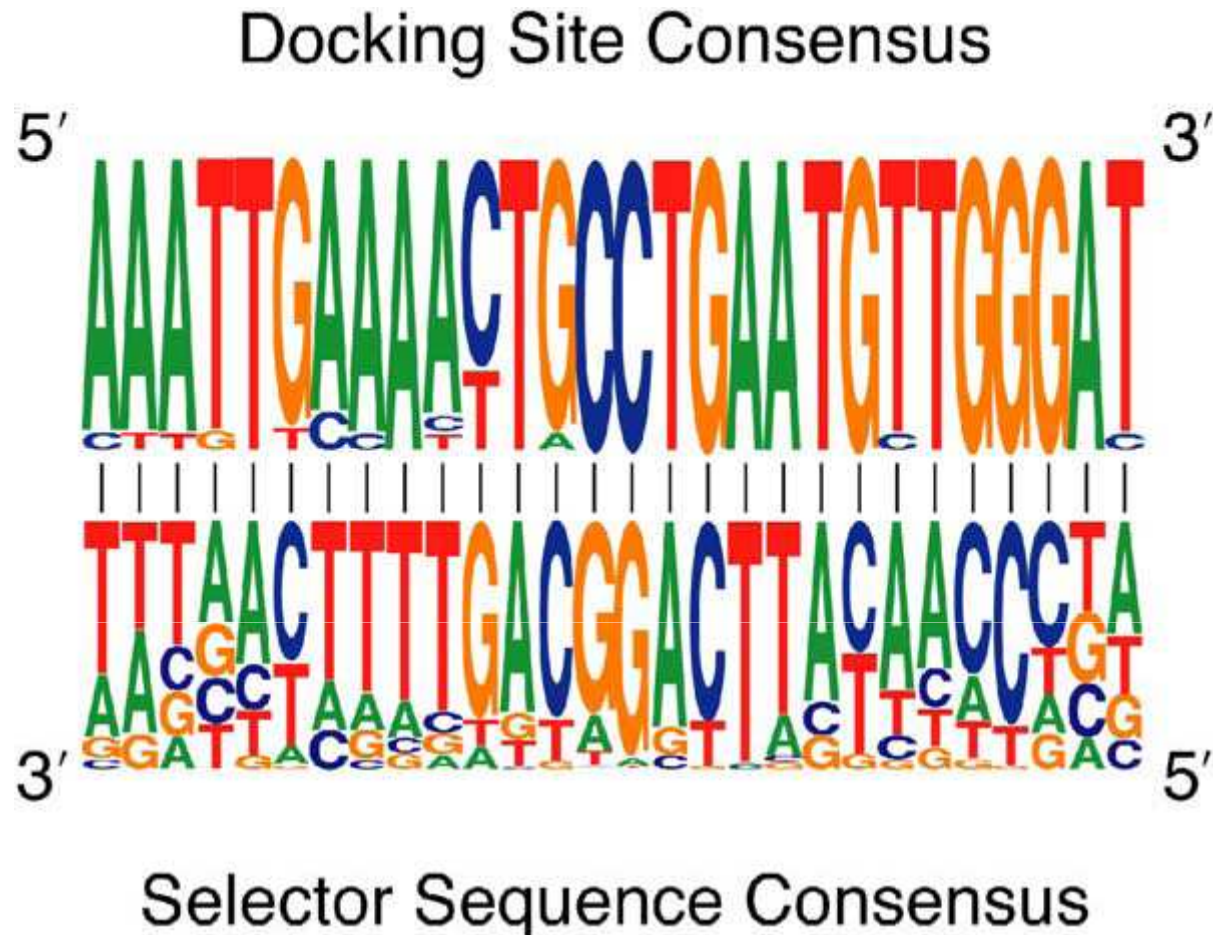
Figure 5. The Docking Site and Selector Sequences Consensus Are Complementary
The docking site consensus sequence is complementary to the central 28 nucleotides of the selector sequence consensus. The most frequent nucleotide at each position of the selector sequence is complementary to the docking site.
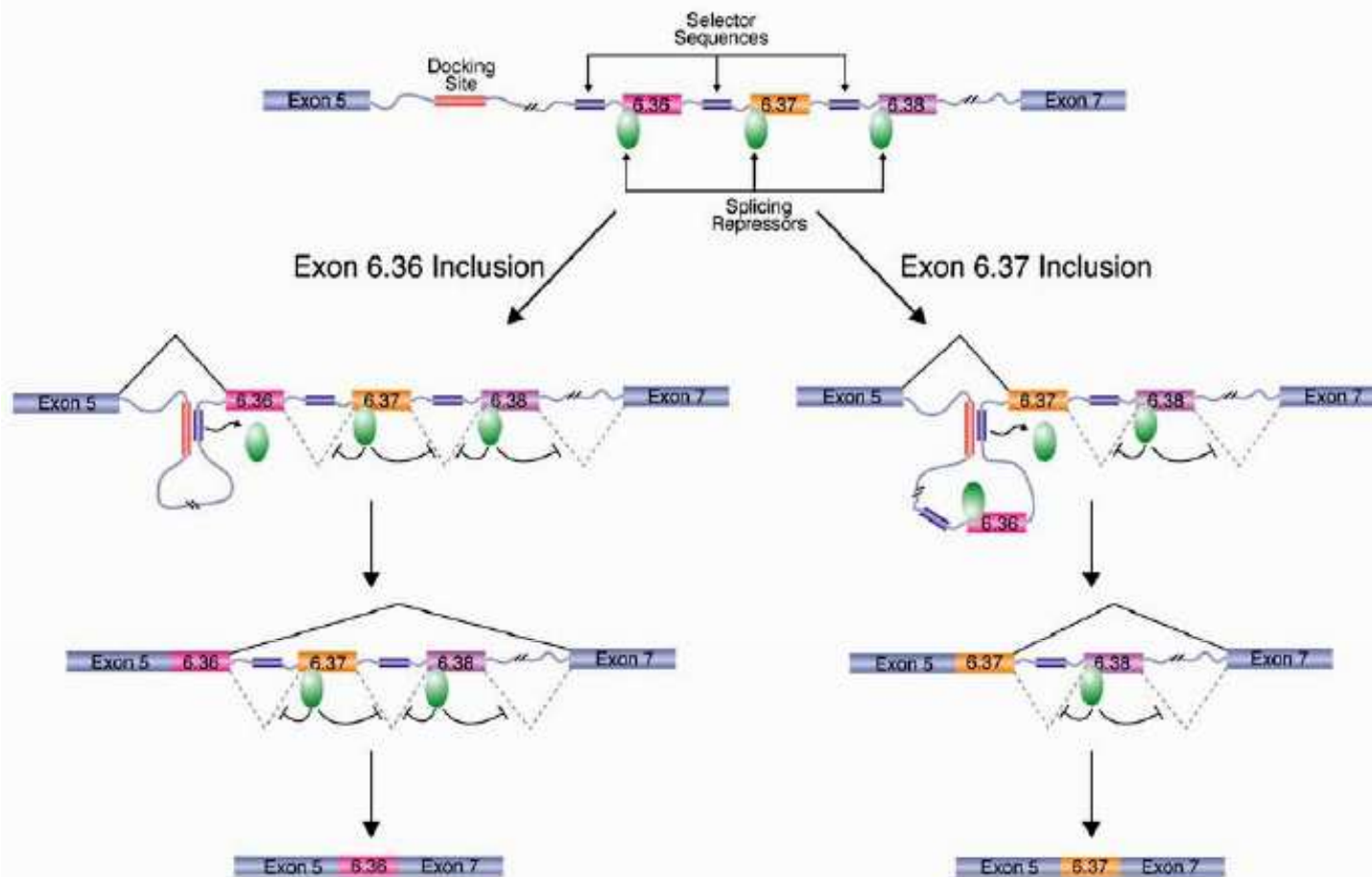
Figure 7. Model for the Mechanism of *Dscam* Exon 6 Mutually Exclusive Splicing

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 given cell are always spliced in the same way…. how can the machinery "remember" which selector was used the first time?

Second, inclusion of weak exons may depend on the ability of promoters to regulate the "speed" of RNA Polymerase II

Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation

Sebastián Kadener*, Juan Pablo Fededa*, Michael Rosbash†, and Alberto R. Kornblihtt*‡

*Laboratorio de Fisiología y Biología Molecular, Departamento de Fisiología, Biología Molecular y Celular, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón II (C1428EHA), Buenos Aires, Argentina; and †Howard Hughes Medical Institute, Brandeis University, Waltham, MA 02454

Promoters and enhancers are cis-acting elements that control gene transcription via complex networks of protein–DNA and protein–protein interactions. Whereas promoters deal with putting in place the RNA polymerase, both enhancers and promoters can control transcriptional initiation and elongation. We have previously shown that promoter structure modulates alternative splicing, strengthening the concept of a physical and functional coupling between transcription and splicing. Here we report that the promoter effect is due to the control of RNA pol II elongation. We found that the simian virus 40 (SV40) transcriptional enhancer, inserted in fibronectin (FN) minigene constructs transfected into mammalian cells, controls alternative splicing by inhibiting inclusion of the FN extra domain I (EDI) exon into mature mRNA. Deletion analysis of enhancer subdomains and competitions *in vivo* with excess of specific enhancer DNA subfragments demonstrate that the "minimal" enhancer, consisting of two 72-bp repeats, is responsible for the splicing effect. The 72-bp repeat region has been reported to promote RNA pol II elongation. When transcription is driven by the α-globin promoter linked to the SV40 enhancer, basal EDI inclusion and activation by the SR (Ser–Arg-rich) protein SF2/ASF are much lower than with other promoters. Deletion of only one of the two 72-bp repeats not only provokes higher EDI inclusion levels but allows responsiveness to SF2/ASF. These effects are the consequence of a decrease in RNA pol II elongation evidenced both by an increase in the proportions of shorter proximal over full length transcripts and by higher pol II densities upstream of the alternative exon detected by chromatin immunoprecipitation.
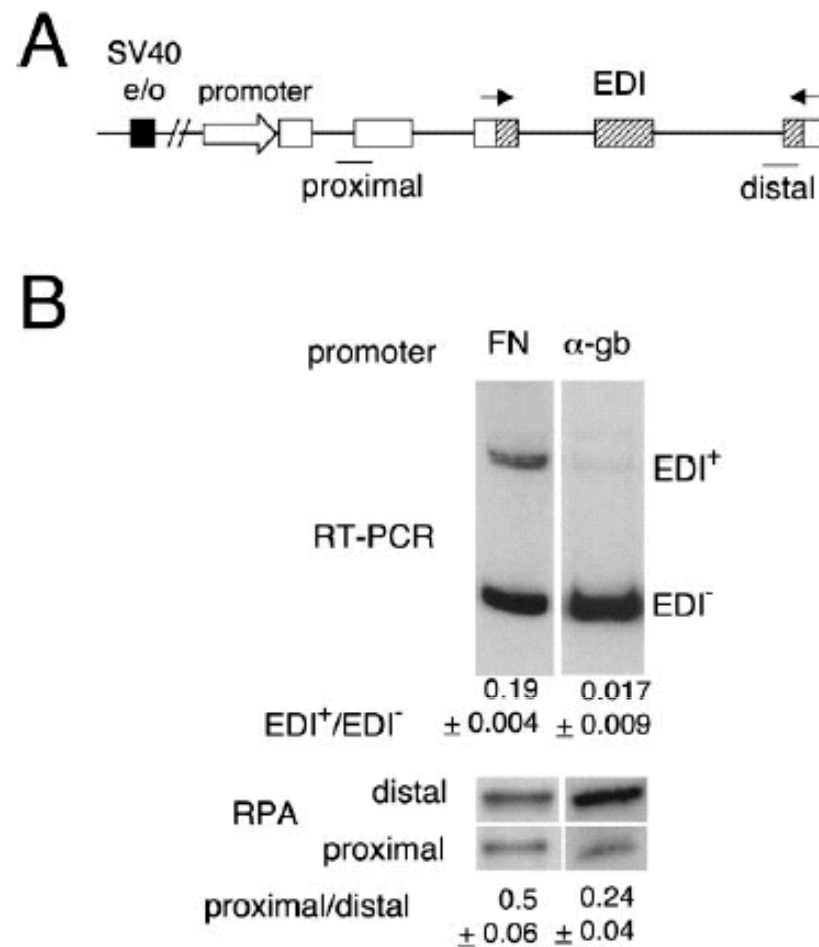
Fig. 1. FN and α-gb promoters elicit different alternative splicing ratios and pol II processivities. (A) Scheme of the minigenes transfected to assess alternative splicing. Open exons, human α-gb; dashed exons, human FN; black box, SV40 e/o; arrows, primers used to amplify the mRNA splicing variants by RT-PCR, and lines, proximal and distal probes used for RPA. (B Upper) Hep3B cells were transfected with 600 ng of pSVEDA/FN (FN promoter) or pSVEDA-Tot (α-gb promoter) plus 400 ng of pCMVβgal. RNA splicing variants were detected by radioactive RT-PCR and analyzed in 6% native polyacrylamide gels. Ratios between radioactivity in EDI⁺ bands and radioactivity in EDI⁻ bands are shown under each lane. (Lower) RPA with proximal and distal probes shown in A, to measure levels of short and long transcripts of transfected Hep3B cells. RT-PCR and RPA ratios correspond to at least three independent transfection experiments.
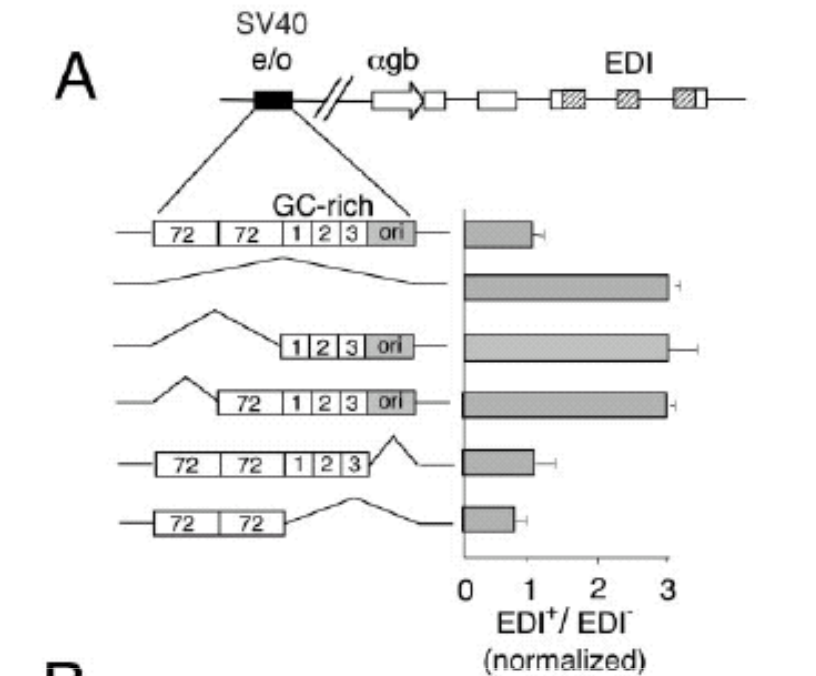
Fig. 3. (*A*) Deletion analysis of the SV40 eo with respect to alternative splicing of the EDI exon. Horizontal bars indicate normalized EDIEDI ratios of Hep3B cells transfected with a series of –gb promoter constructs carrying different internal deletions of the SV40 eo. Results correspond to the mean SD of at least three independent transfection experiments.
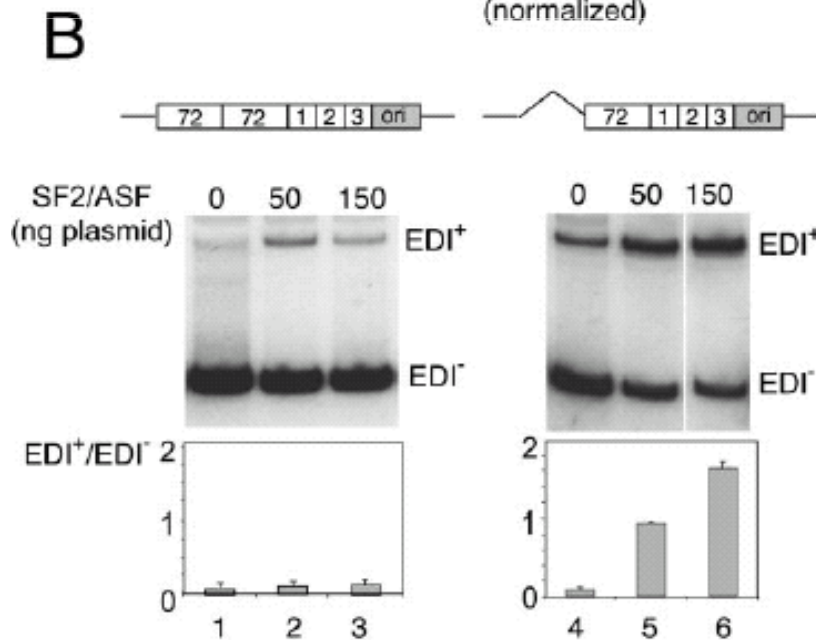
(*B*) Deletion of only one 72-bp repeat confers responsiveness to SF2ASF to the -gb promoter construct. Hep3B cells were transfected with pSVEDATot (lanes 1–3) or a variant lacking the distal 72-bp repeat of the SV40 enhancer (lanes 4–6) and cotransfected with the indicated amounts of a plasmid expressing SF2ASF (13). Transfections in lanes 1 and 4 contained 150 ng of empty DNA vector.

Similar results were obtained in Cos-7 and HeLa cells.
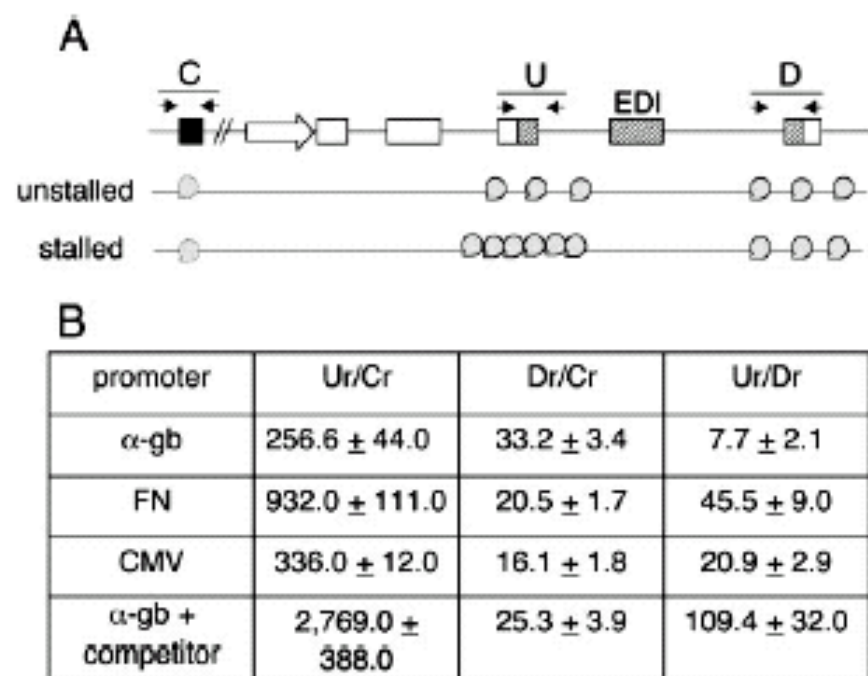
Fig. 5. ChIP with an Ab to RNA pol II. (A) Scheme of the minigenes transfected to assess pol II densities. Arrows indicate the pairs of primers used in real time PCRs to quantitatively amplify DNA that is bound to the immunoprecipitated pol II, at two regions mapping U, D of the EDI alternative exon, and at a third C region outside of the transcription unit. (B) Cells were transfected with α-gb, FN, or CMV promoter constructs and, where indicated, cotransfected with a 10-fold molar excess of a competitor plasmid carrying the SV40 e/o. After 48 h, cells were fixed with formaldehyde and treated for ChIP and real time PCR analysis as described in *Experimental Procedures*. Ur = Uim/Uin; Cr = Cim/Cin; Dr = Dim/Din where Uim, Cim, and Dim are the template DNA amounts recovered after *im*munoprecipitation by anti-pol II, and Uin, Cin, and Din are the *in*put DNA amounts, all estimated by real time PCR at regions U, C, and D, respectively. Results correspond to a representative transfection experiment of Cos-7 cells and show the mean $\pm$ SD of three real time PCR determinations.

These studies and the following in the decade suggest a model, that is called the "bump" model.

At specific points, there may be a (structurally determined?) obstacle to the advancement of RNA Pol II.

This allows the "weak" sites more opportunity to be spliced with the previous, just because the competing "good" site has not been transcribed yet

Very similar results were obtained using other transcriptional regulators: the inclusion of weak exons was found to be repressed in response to steroids if the reporter gene contained a SRE in its promoter (ERE, GRE, PRE).

Coordinate Regulation of Transcription and Splicing by Steroid Receptor Coregulators

Didier Auboeuf,[1] Arnd Hönig,[2]* Susan M. Berget,[2]
Bert W. O'Malley[1]†

Recent observations indicating that promoter identity influences alternative RNA-processing decisions have created interest in the regulatory interactions between RNA polymerase II transcription and precursor messenger RNA (pre-mRNA) processing. We examined the impact of steroid receptor–mediated transcription on RNA processing with reporter genes subject to alternative splicing driven by steroid-sensitive promoters. Steroid hormones affected the processing of pre-mRNA synthesized from steroid-sensitive promoters, but not from steroid-unresponsive promoters, in a steroid receptor–dependent and receptor-selective manner. Several nuclear receptor coregulators showed differential splicing effects, suggesting that steroid hormone receptors may simultaneously control gene transcription activity and exon content of the product mRNA by recruiting coregulators involved in both processes.

This was found to depend on steroid receptor interaction with "coactivators" that also function as splicing regulators.

Steroid Hormone Receptor Coactivation and Alternative RNA Splicing by U2AF^{65}-Related Proteins CAPERα and CAPERβ

Dennis H. Dowhan,[1] Eugene P. Hong,[2]
Didier Auboeuf,[1,3] Andrew P. Dennis,[1]
Michelle M. Wilson,[2] Susan M. Berget,[2]
and Bert W. O'Malley[1,*]
[1]Department of Molecular and Cellular Biology
[2]Department of Biochemistry and Molecular Biology
Baylor College of Medicine
One Baylor Plaza
Houston, Texas 77030

Increasing evidence indicates that transcription and pre-mRNA processing are functionally coupled to modulate gene expression. Here, we report that two members of the U2AF^{65} family of proteins, hCC1.3, which we call CAPERα, and a related protein, CAPERβ, regulate both steroid hormone receptor-mediated transcription and alternative splicing. The CAPER proteins coactivate the progesterone receptor in luciferase transcription reporter assays and alter alternative splicing of a calcitonin/calcitonin gene-related peptide minigene in a hormone-dependent manner. The importance of CAPER coactivators in the regulation of alternative RNA splicing of an endogenous cellular gene (VEGF) was substantiated by siRNA knockdown of CAPERα. Mutational analysis of CAPERβ indicates that the transcriptional and splicing functions are located in distinct and separable domains of the protein. These results indicate that steroid hormone receptor-regulated transcription and pre-mRNA splicing can be directly linked through dual function coactivator molecules such as CAPERα and CAPERβ.
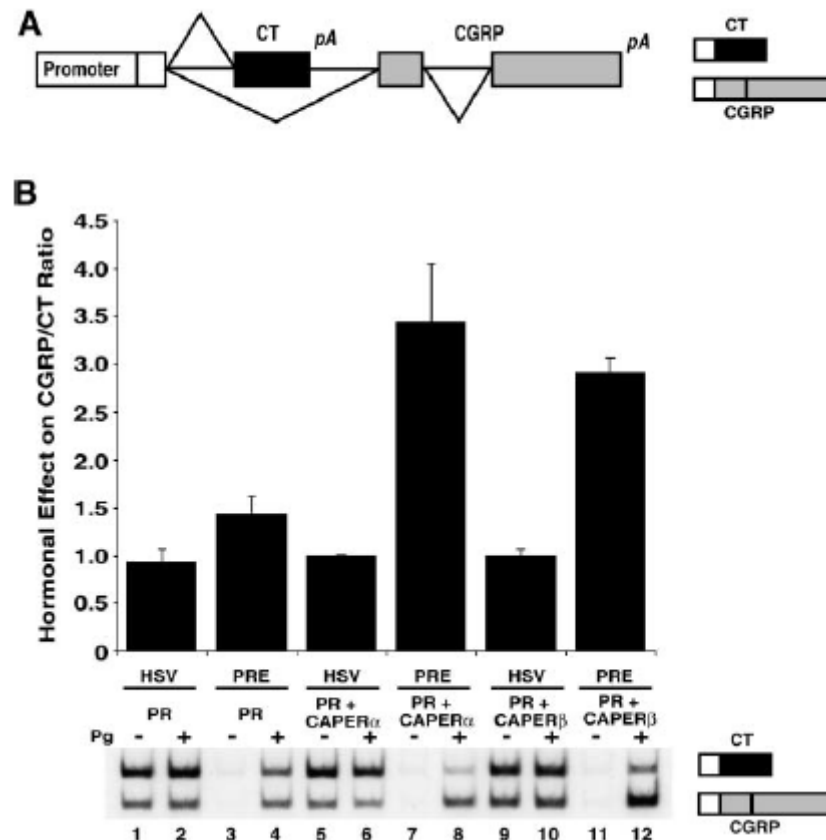
....mmm .... wait a moment!

how does it come that a "coactivator" that is recruited to promoters (or enhancers) through interaction with a steroid receptor, then may regulate splicing, which happens far away from there ?
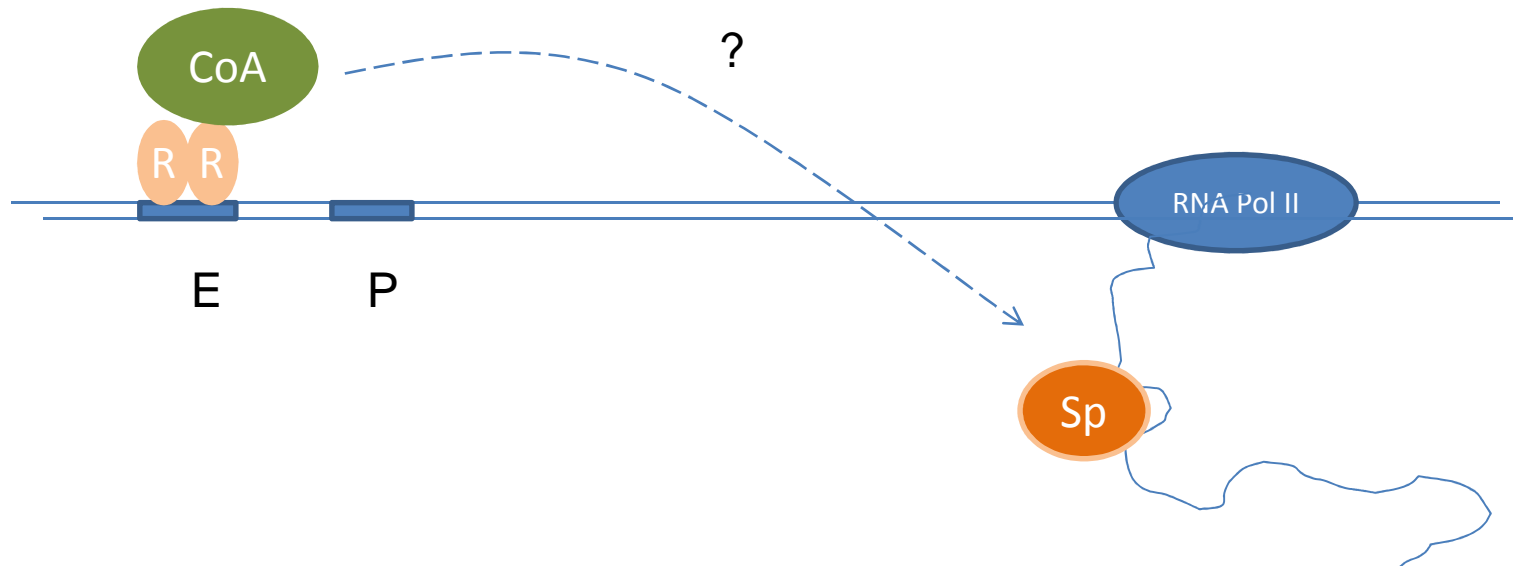
This is explained by "**co-transcriptionality**"

"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

$Y_1S_2P_3T_4S_5P_6S_7$

B

A

$Y_1S_2P_3T_4S_5P_6S_7$

B
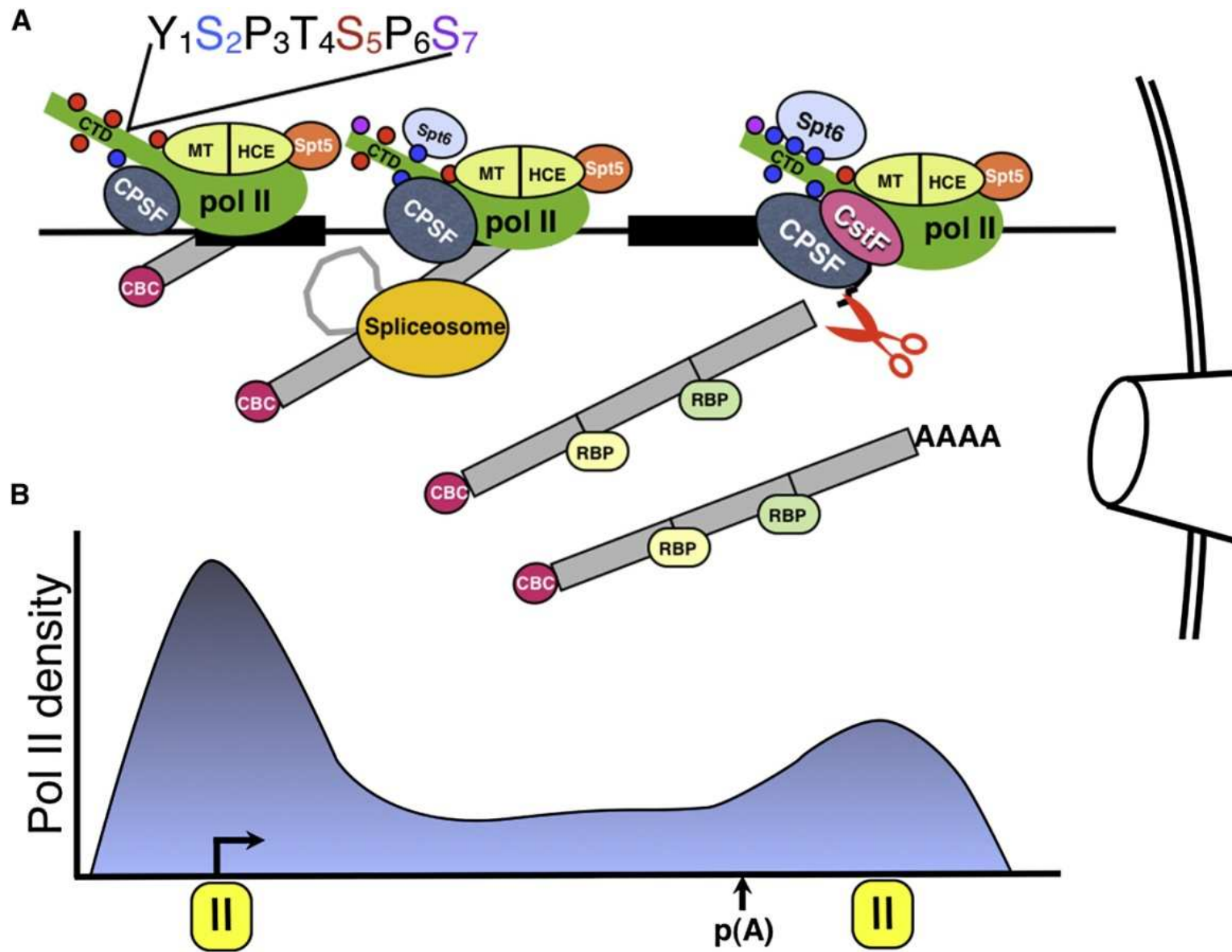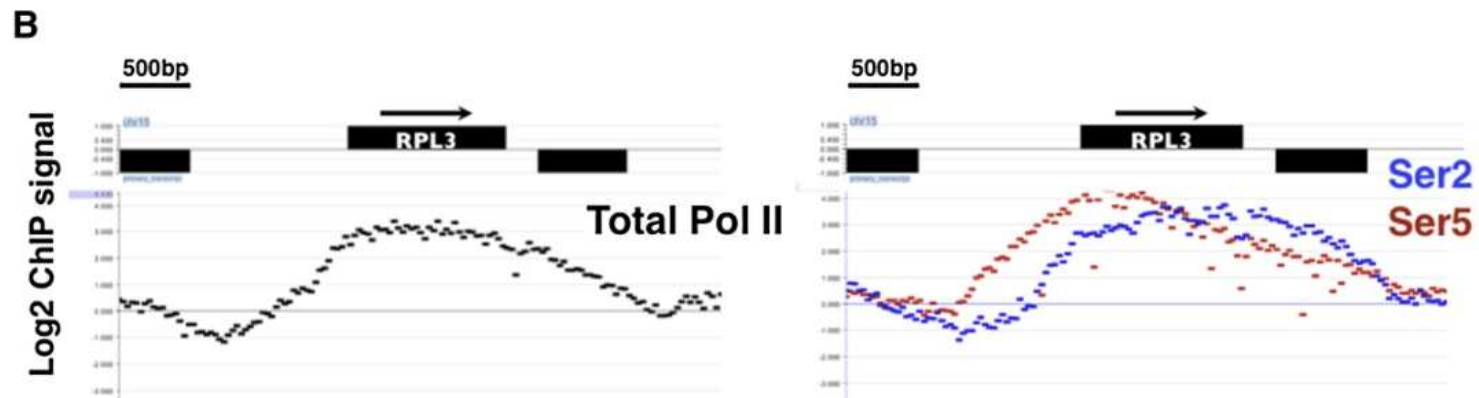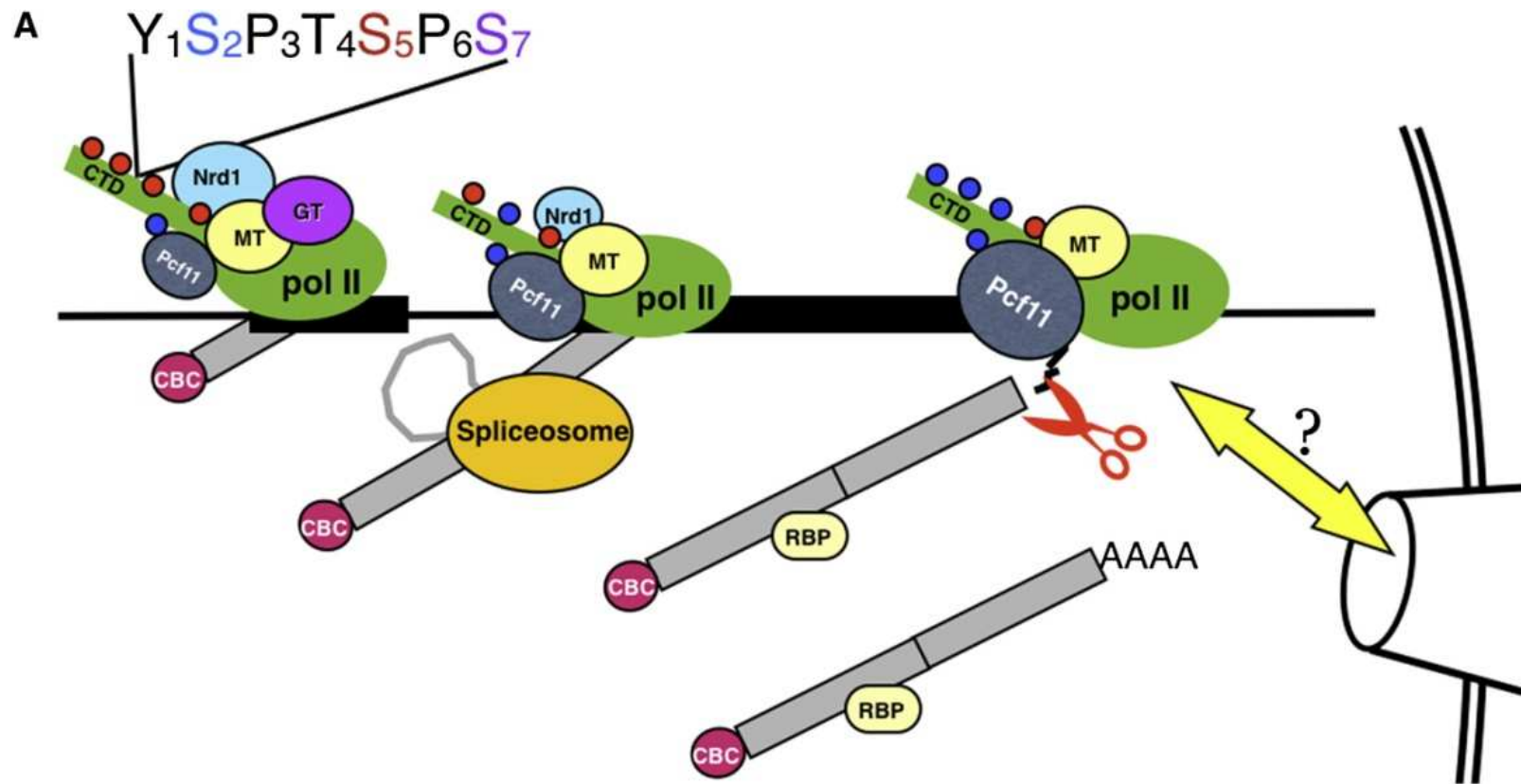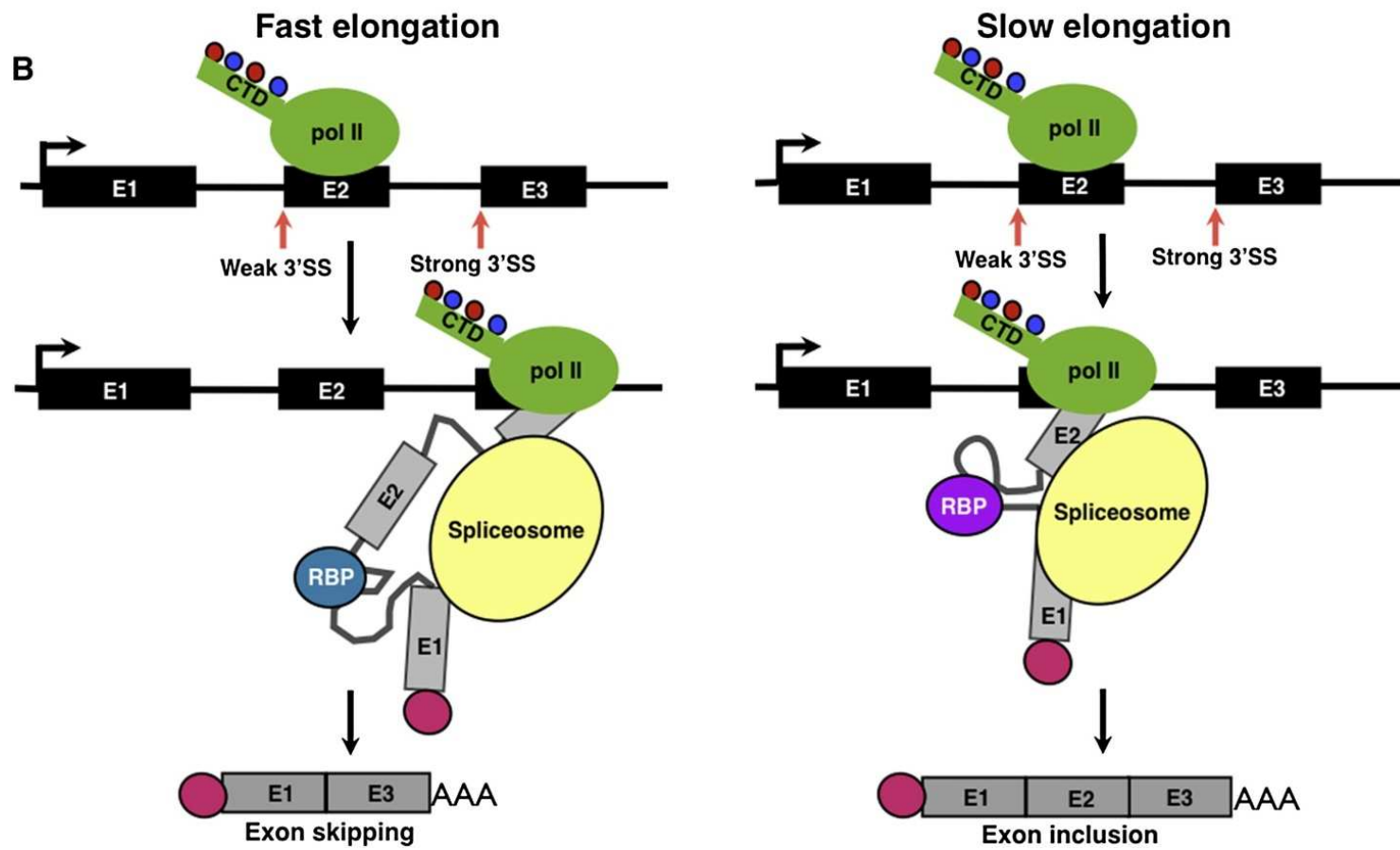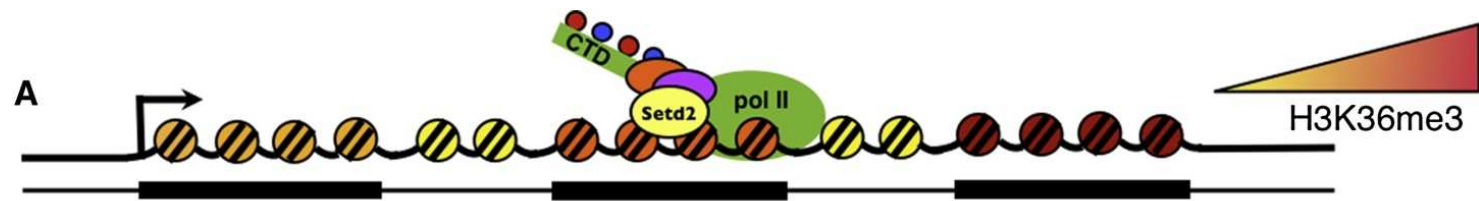
Two different but not exclusive models can be devised:

1) the speed counts or "bump" model – just takes into account the opportunity of a splice site to interact with the appropriae proteins.

2) the "differential splicing factor" model – RNA Pol II CTD aither carries different splicing factors that were associated 1) during assembly at promoters or b) in place, since they are associated to, for example, nucleosomes (?).

Chromatin organization marks exon-intron structure

Schraga Schwartz[1], Eran Meshorer[2] & Gil Ast[1]

An increasing body of evidence indicates that transcription and splicing are coupled, and it is accepted that chromatin organization regulates transcription. Little is known about the cross-talk between chromatin structure and exon-intron architecture. By analysis of genome-wide nucleosome-positioning data sets from humans, flies and worms, we found that exons show increased nucleosome-occupancy levels with respect to introns, a finding that we link to differential GC content and nucleosome-disfavoring elements between exons and introns. Analysis of genome-wide chromatin immunoprecipitation data in humans and mice revealed four specific post-translational histone modifications enriched in exons. Our findings indicate that previously described enrichment of H3K36me3 modifications in exons reflects a more fundamental phenomenon, namely increased nucleosome occupancy along exons. Our results suggest an RNA polymerase II–mediated cross-talk between chromatin structure and exon-intron architecture, implying that exon selection may be modulated by chromatin structure.

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.

nucleosome positioning assay
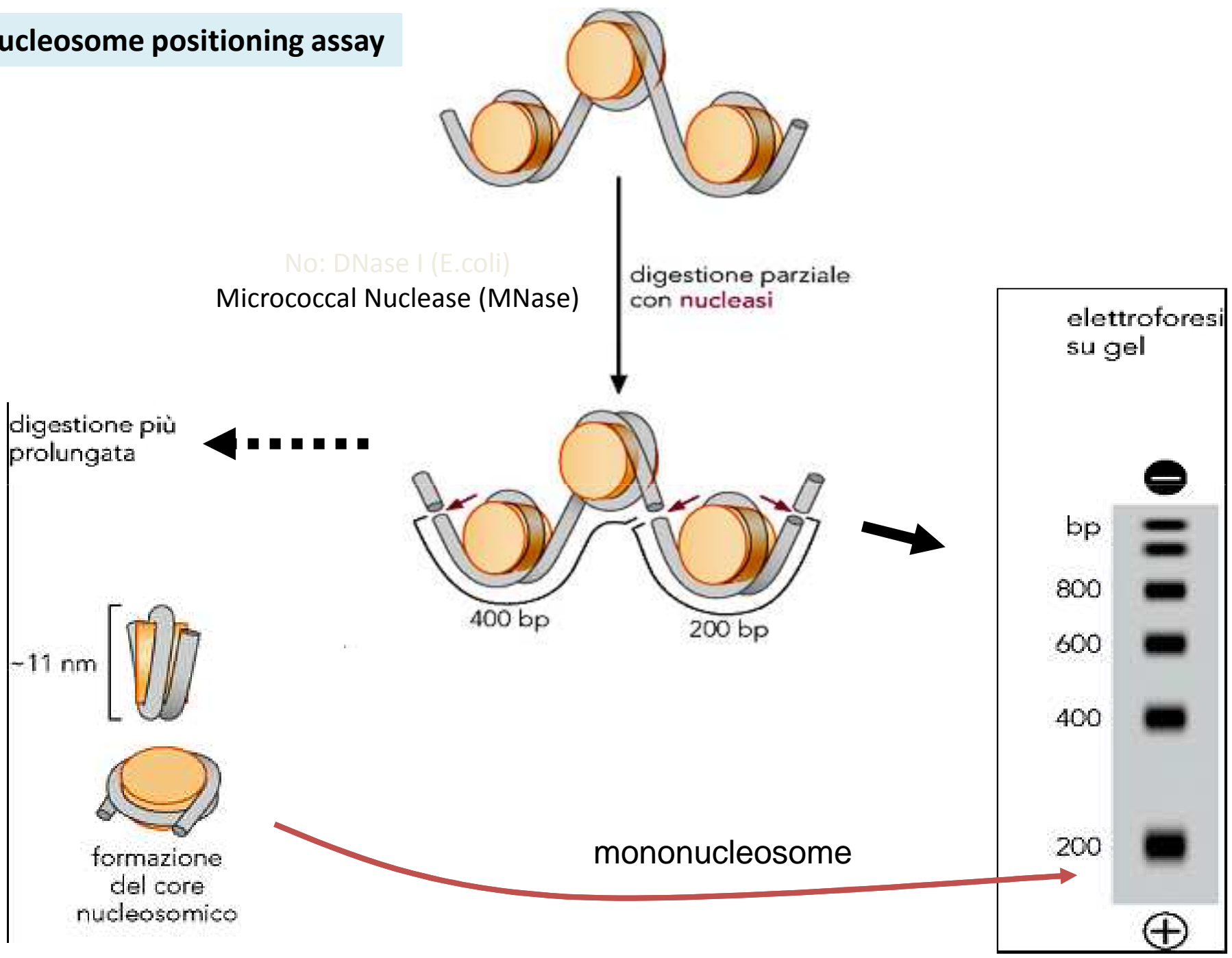
No: DNase I (E.coli)

Micrococcal Nuclease (MNase)

digestione parziale con **nucleasi**

digestione più prolungata

~11 nm

400 bp   200 bp

elettroforesi su gel

bp

800

600

400

200

formazione del core nucleosomico

mononucleosome

Nucleosome positioning analysis

1st step: digestion

MNasi

nuclei isolati

siti di taglio con MNasi

siti di taglio con RE
RE = enzimi di restrizione

purificazione del DNA tagliato
e trattamento con enzimi
di restrizione

nucleosomi casuali

nucleosomi posizionati
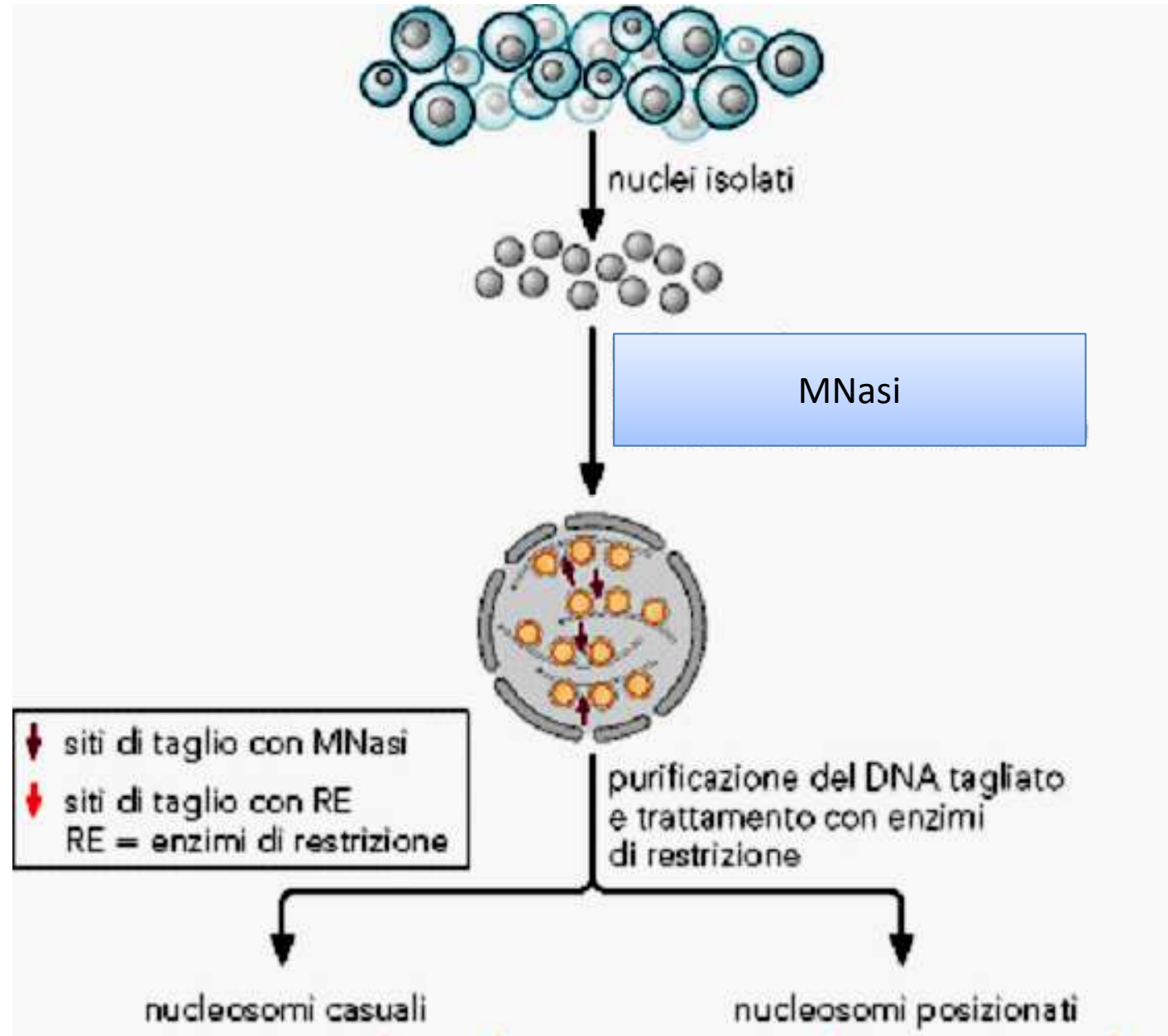
2 different cases:
mobile nucleosomes
versus
positioned nucleosomes
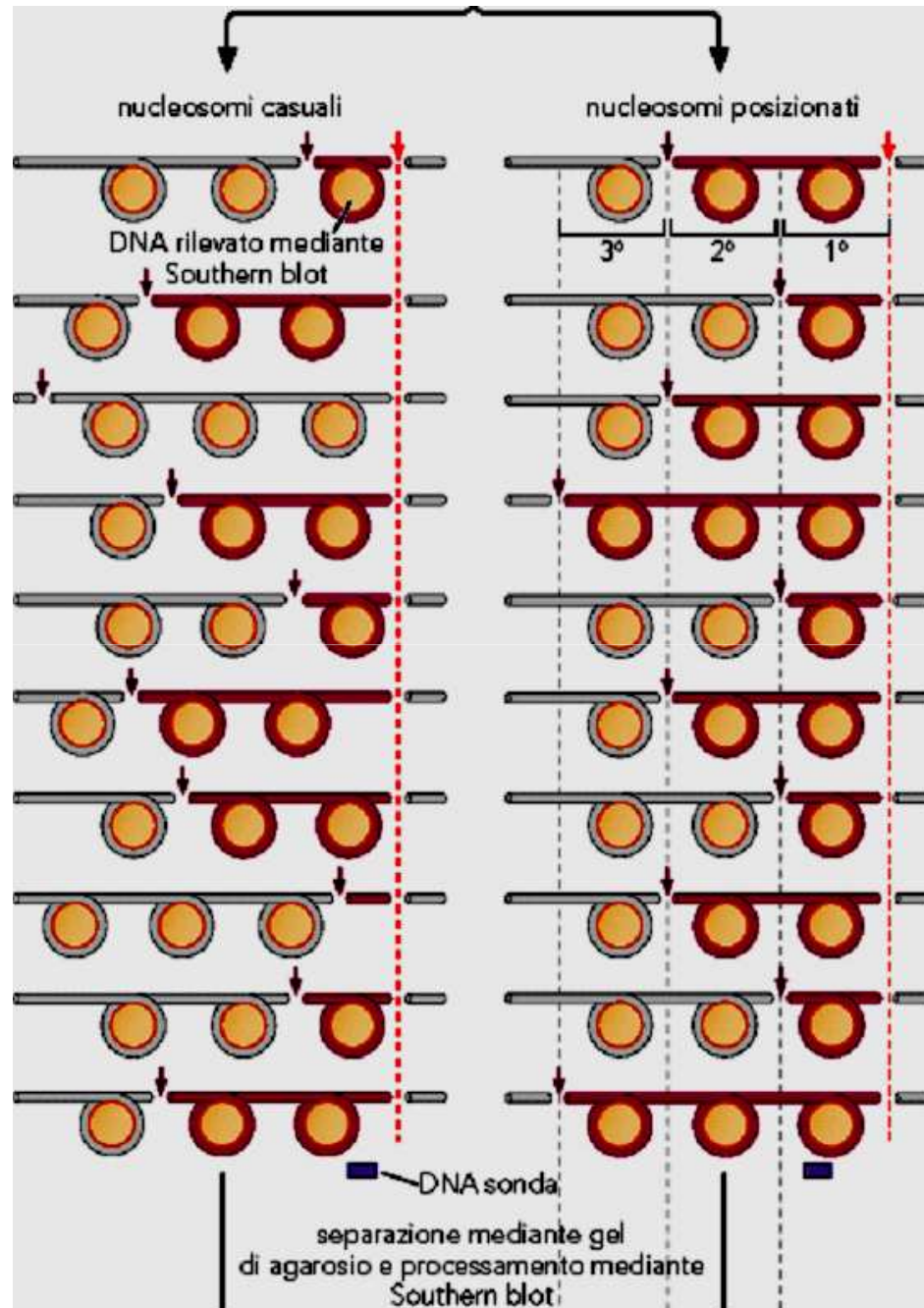
Analysis

mobile   versus   positioned

DNA sonda

separazione mediante gel
di agarosio e processamento mediante
Southern blot

sonda per
Southern blot

3°   2°   1°

taglio di molte posizioni a causa
della posizione casuale
dei nucleosomi

taglio in specifiche regioni
che si trovano fra i nucleosomi
posizionati

Pho2 binding sites

Pho4 is the P-sensitive inducer, whereas Pho2 is constitutive
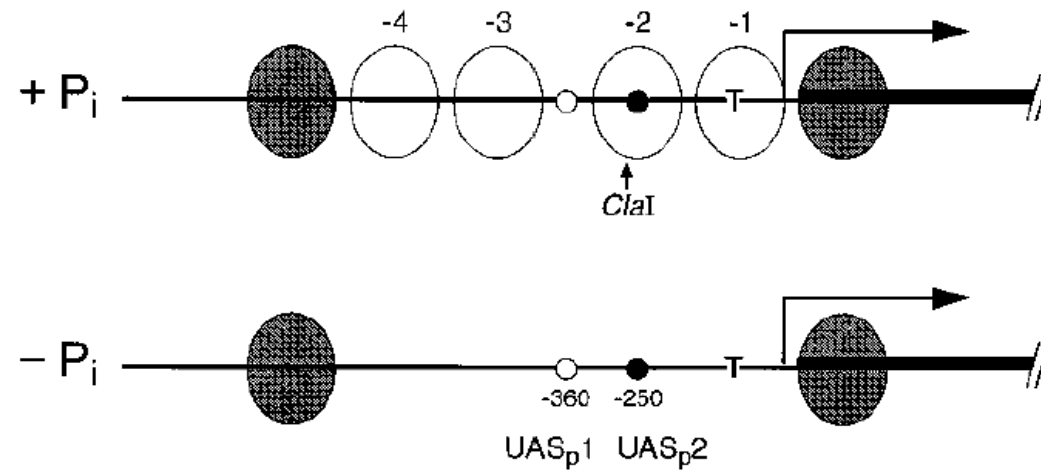
Figure 1. Chromatin Structure at the *PHO5* Promoter

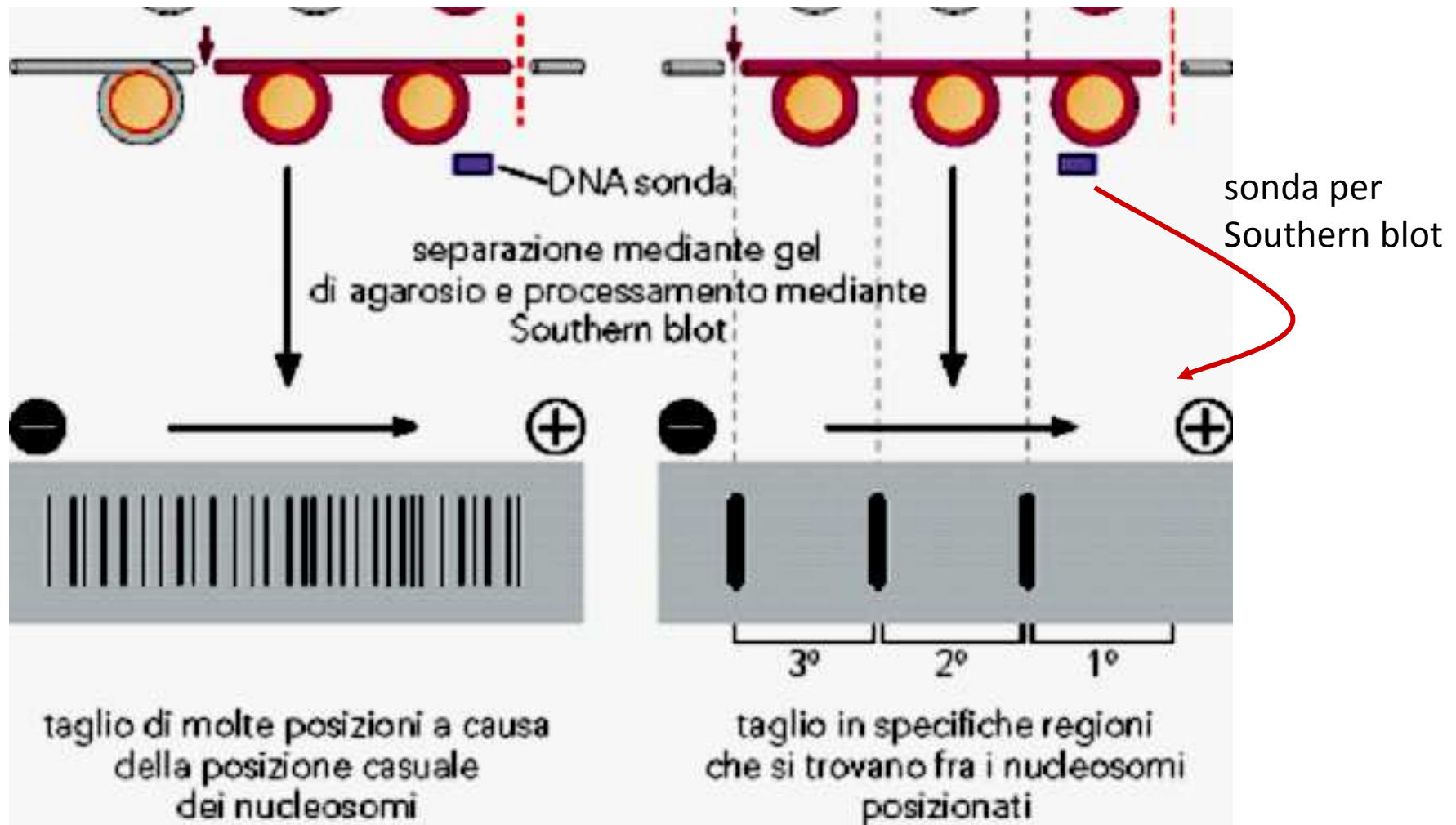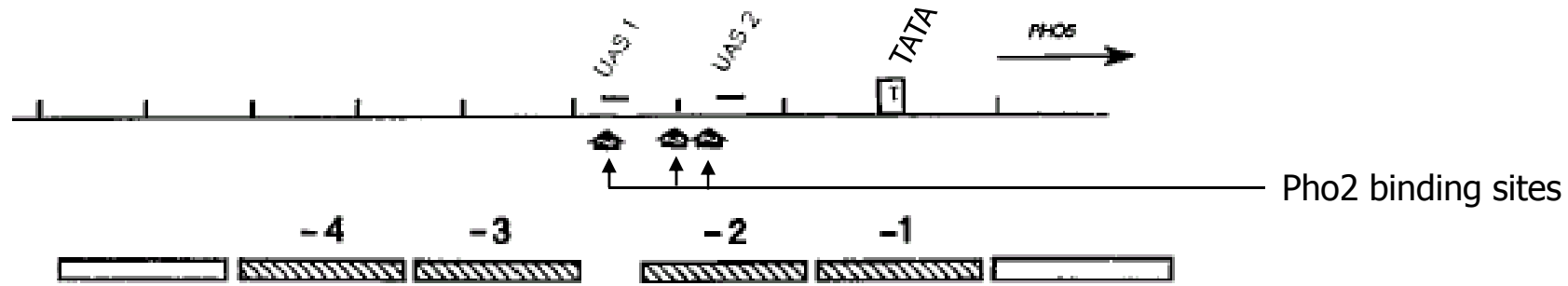Nucleosomes −1, −2, −3, and −4 are remodeled upon activation of the promoter by phosphate starvation conditions (Almer et al., 1986). The small circles mark UASp1 (open) and UASp2 (solid), which are Pho4-binding sites found by in vitro (Vogel et al., 1989) and in vivo (Venter et al., 1994) footprinting experiments. The positions are listed relative to the coding sequence (solid bar). T denotes the TATA box (Rudolph and Hinnen, 1987). The location of a ClaI site at −275 relative to the coding region is shown.
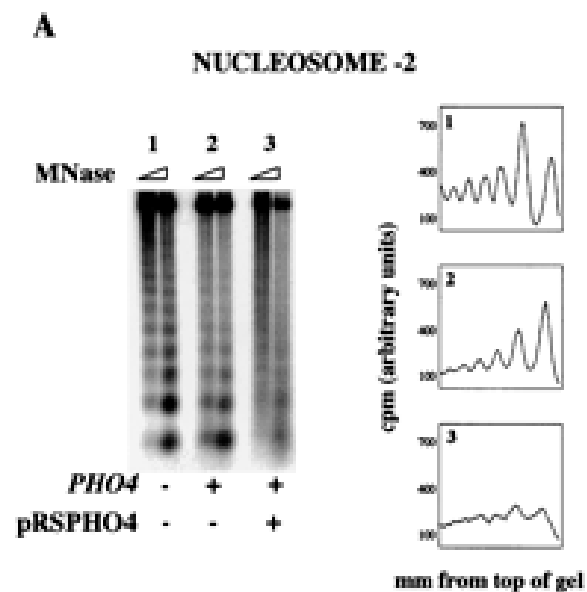
FIG. 5. In vivo chromatin remodeling of episomal *PHO5*. Spheroplasts from the indicated strains were treated with micrococcal nuclease (MNase), and the DNA was purified and Southern blotted. (A) The blot was probed with probe A (Fig. 2). Data from each sample were quantified, and the distance from the top of the gel was graphed against the signal density. (B) The blot shown in panel A was stripped and reprobed with probe B (Fig. 2).
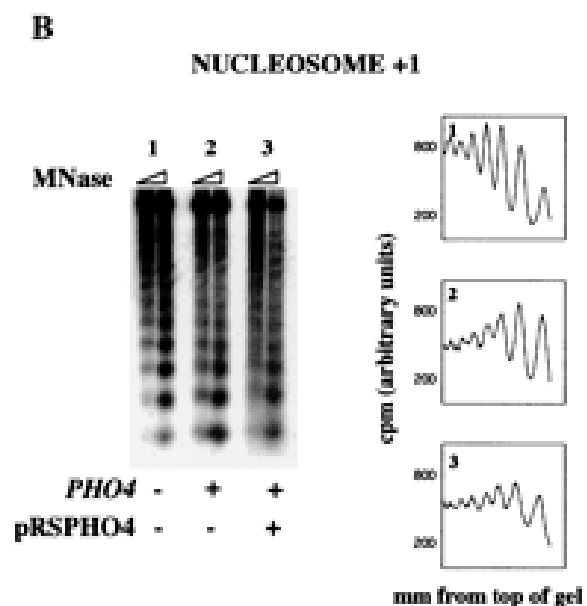
After the advent of genome-wide tiling microarrays or deep-sequencing methods, it has become possible to address the question:

are nucleosomes positioned in the genome?

Protocol   (as illustrated in previous slides)

1. isolate nuclei
2. MNase
3. purify DNA
4. go to tiling array or (better) deep-sequencing

5. draw the frequency map: if nucleosome are prevalently positioned at certain sequences, those will appear more frequently (give peaks).

Dynamic Regulation of Nucleosome Positioning in the Human Genome

Dustin E. Schones,[1,2] Kairong Cui,[1,2] Suresh Cuddapah,[1] Tae-Young Roh,[1] Artem Barski,[1]
Zhibin Wang,[1] Gang Wei,[1] and Keji Zhao[1,*]
[1]Laboratory of Molecular Immunology, The National Heart, Lung and Blood Institute, NIH, Bethesda, MD 20892, USA
[2]These authors contributed equally to this work.

SUMMARY

The positioning of nucleosomes with respect to DNA plays an important role in regulating transcription. However, nucleosome mapping has been performed for only limited genomic regions in humans. We have generated genome-wide maps of nucleosome positions in both resting and activated human CD4+ T cells by direct sequencing of nucleosome ends using the Solexa high-throughput sequencing technique. We find that nucleosome phasing relative to the transcription start sites is directly correlated to RNA polymerase II (Pol II) binding. Furthermore, the first nucleosome downstream of a start site exhibits differential positioning in active and silent genes. TCR signaling induces extensive nucleosome reorganization in promoters and enhancers to allow transcriptional activation or repression. Our results suggest that H2A.Z-containing and modified nucleosomes are preferentially lost from the −1 nucleosome position. Our data provide a comprehensive view of the nucleosome landscape and its dynamic regulation in the human genome.

Mononucleosome isolated after Mnase then DNA mass-sequenced

Each mononucleosome has 147 or little more DNA

Illumina Solexa sequencing platforms read 25 nt from each end
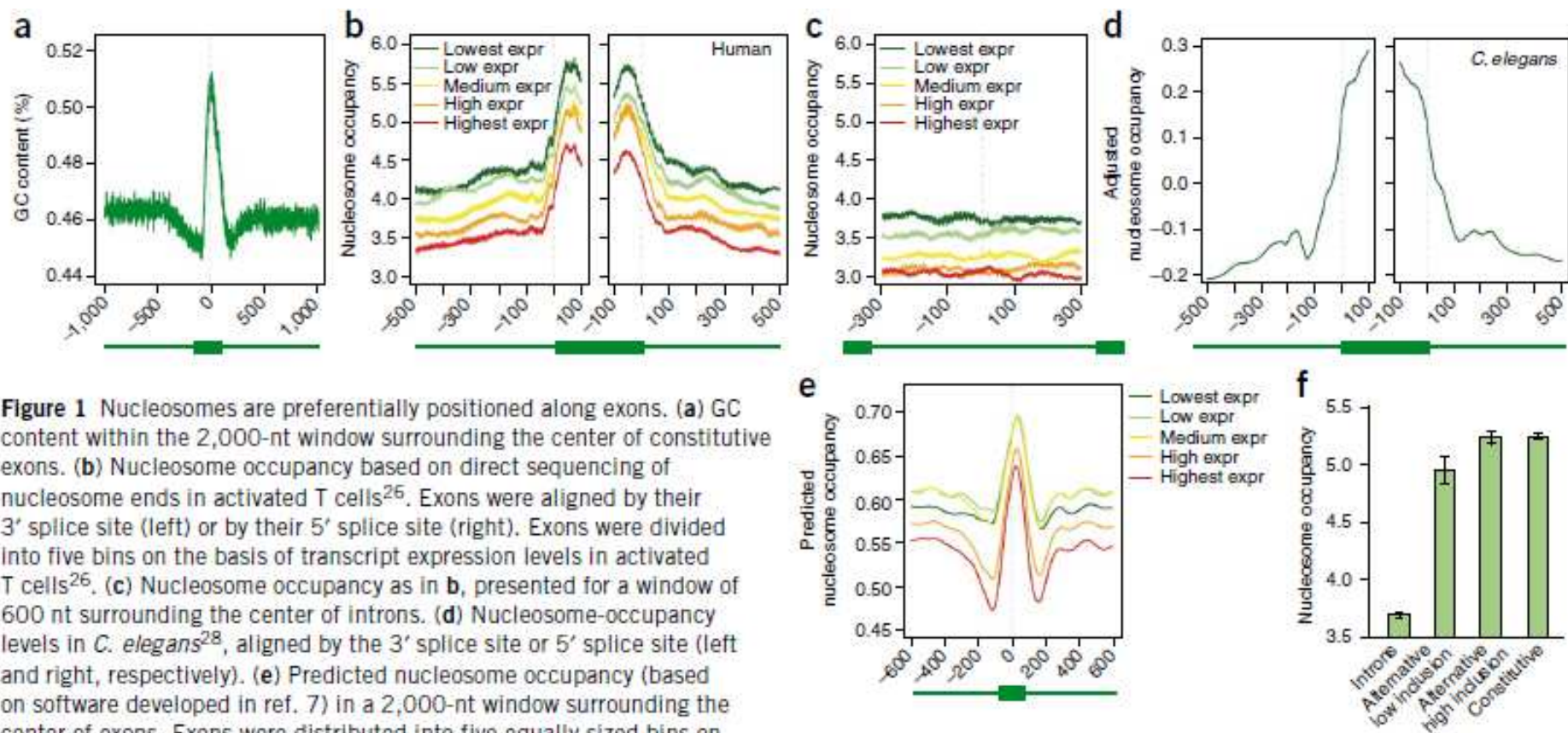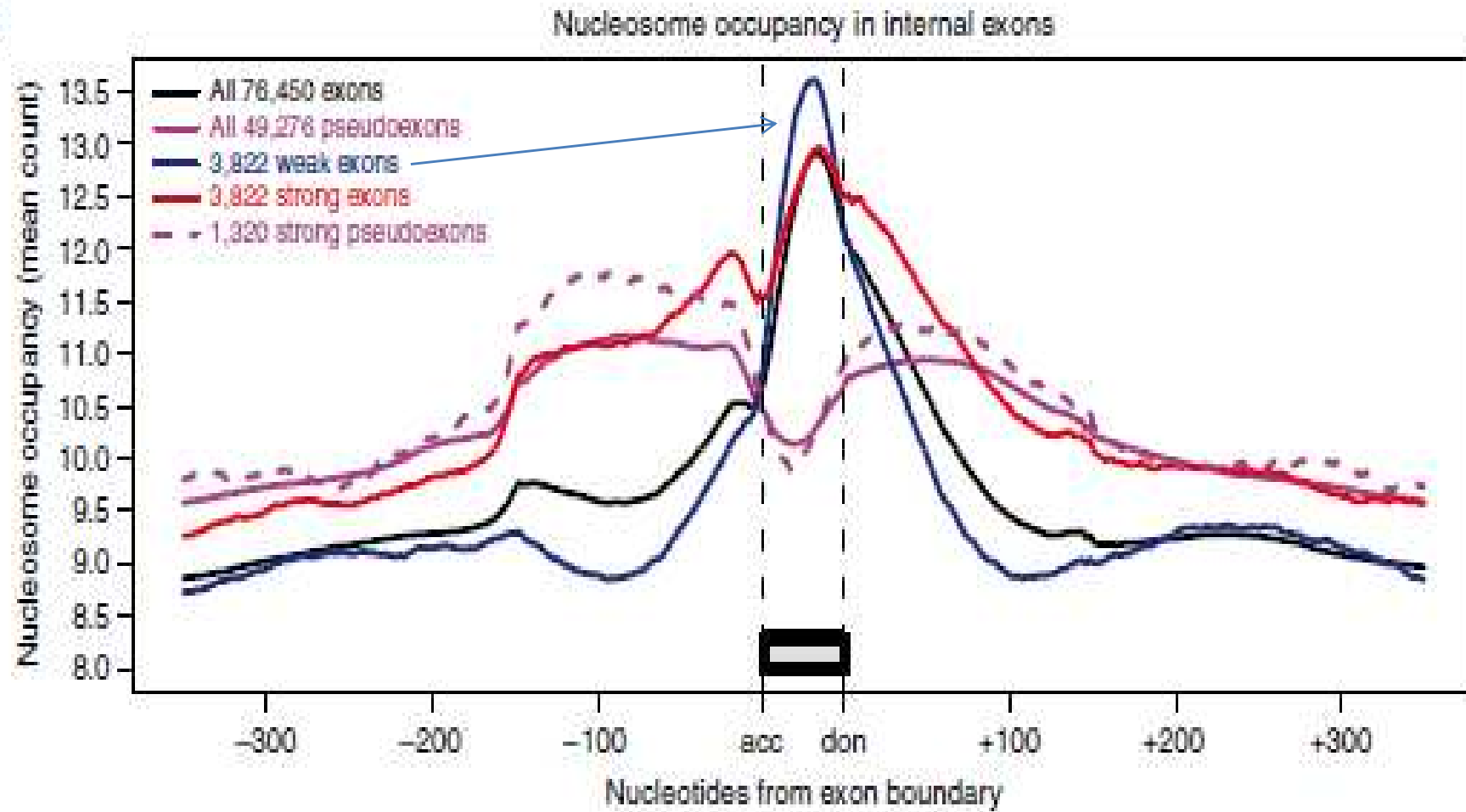
This allowed positioning

Figure 1 Nucleosomes are preferentially positioned along exons. (**a**) GC content within the 2,000-nt window surrounding the center of constitutive exons. (**b**) Nucleosome occupancy based on direct sequencing of nucleosome ends in activated T cells[26]. Exons were aligned by their 3' splice site (left) or by their 5' splice site (right). Exons were divided into five bins on the basis of transcript expression levels in activated T cells[26]. (**c**) Nucleosome occupancy as in **b**, presented for a window of 600 nt surrounding the center of introns. (**d**) Nucleosome-occupancy levels in *C. elegans*[28], aligned by the 3' splice site or 5' splice site (left and right, respectively). (**e**) Predicted nucleosome occupancy (based on software developed in ref. 7) in a 2,000-nt window surrounding the center of exons. Exons were distributed into five equally sized bins on the basis of expression levels as in **b**. (**f**) Mean nucleosome occupancy in introns, alternatively spliced exons included in less than 50% of transcripts, alternatively spliced exons included in at least 50% of transcripts and constitutively spliced exons. Error bars represent the s.e.m.

a

Nucleosome occupancy in internal exons

- All 76,450 exons
- All 49,276 pseudoexons
- 3,922 weak exons
- 3,922 strong exons
- 1,320 strong pseudoexons

Nucleosome occupancy (mean count)

13.5, 13.0, 12.5, 12.0, 11.5, 11.0, 10.5, 10.0, 9.5, 9.0, 8.5, 8.0

−300, −200, −100, acc, don, +100, +200, +300

Nucleotides from exon boundary

Tilgner et al., 2009

Review

Epigenetics in Alternative Pre-mRNA Splicing

Reini F. Luco,[1] Mariano Allo,[2] Ignacio E. Schor,[2] Alberto R. Kornblihtt,[2] and Tom Misteli[1,*]
[1]National Cancer Institute, National Institutes of Health, Bethesda, MD 20892 USA
[2]Departamento de Fisiologia Biologia Molecular, LFBM and IFIBYNE-CONICET, Facultad de Ciencias Exactas Naturales, Universidad de Buenos Aires, Buenos Aires, Argentina
*Correspondence: mistelit@mail.nih.gov
DOI 10.1016/j.cell.2010.11.056

Alternative splicing plays critical roles in differentiation, development, and disease and is a major source for protein diversity in higher eukaryotes. Analysis of alternative splicing regulation has traditionally focused on RNA sequence elements and their associated splicing factors, but recent provocative studies point to a key function of chromatin structure and histone modifications in alternative splicing regulation. These insights suggest that epigenetic regulation determines not only what parts of the genome are expressed but also how they are spliced.