# Lesson 5.2

Post-transcriptional regulation

Regulation of alternative splicing

Alternative exon inclusion/exclusion :   regulated process

- Exon-intron border sequence (splice-sites)

- Exon definition (intron definition in lower eukaryotes)

- SR and hnRNPs factors

- **_cis_-elements and _trans_-regulatory factors**

- **Tissue-specific Splicing Regulators**

- **Nucleosome positioning over exons**

- **Histone PTMs and readers-splicing factors association**

- **Transcriptional speed and pausing**

We can consider the <u>older competition model</u>, as exemplified by the case of a «skipping exon»


3' splice site

5' splice site
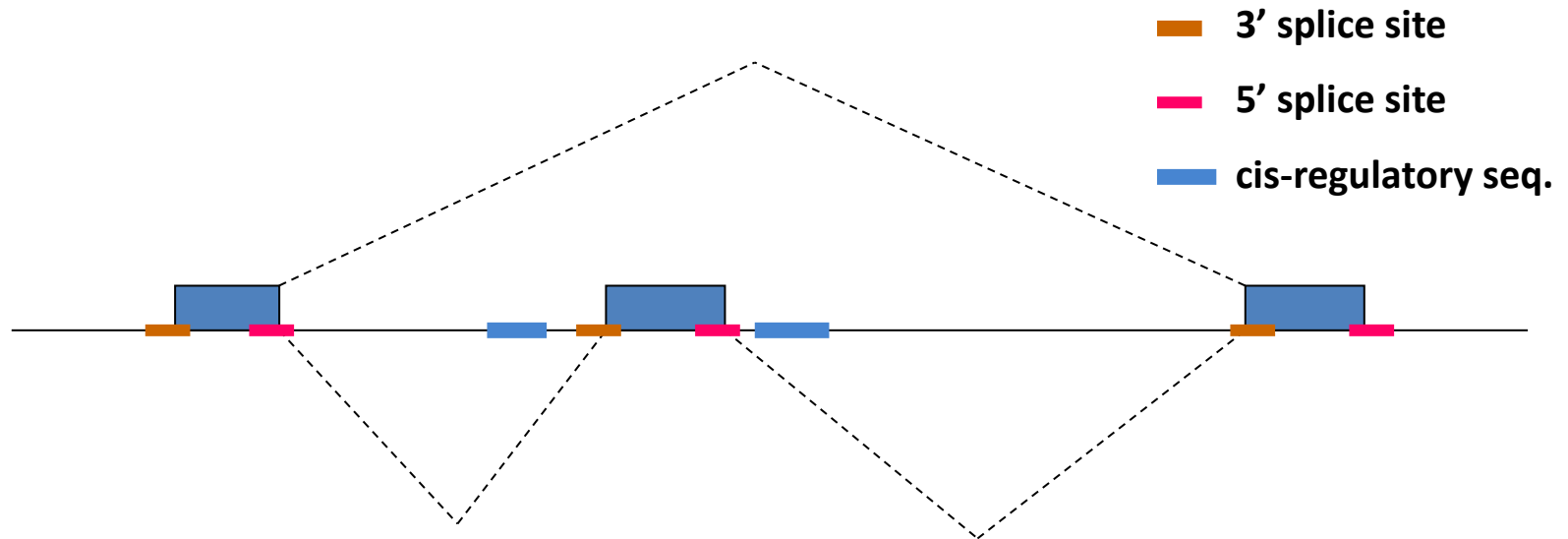
cis-regulatory seq.

In absence of other regulatory elements (cis-regulatory), only one result is seen: the <u>strongest splice sites will predominate</u>.

(*Actually, this model did not account for co-transcriptional splicing...discussed below*)
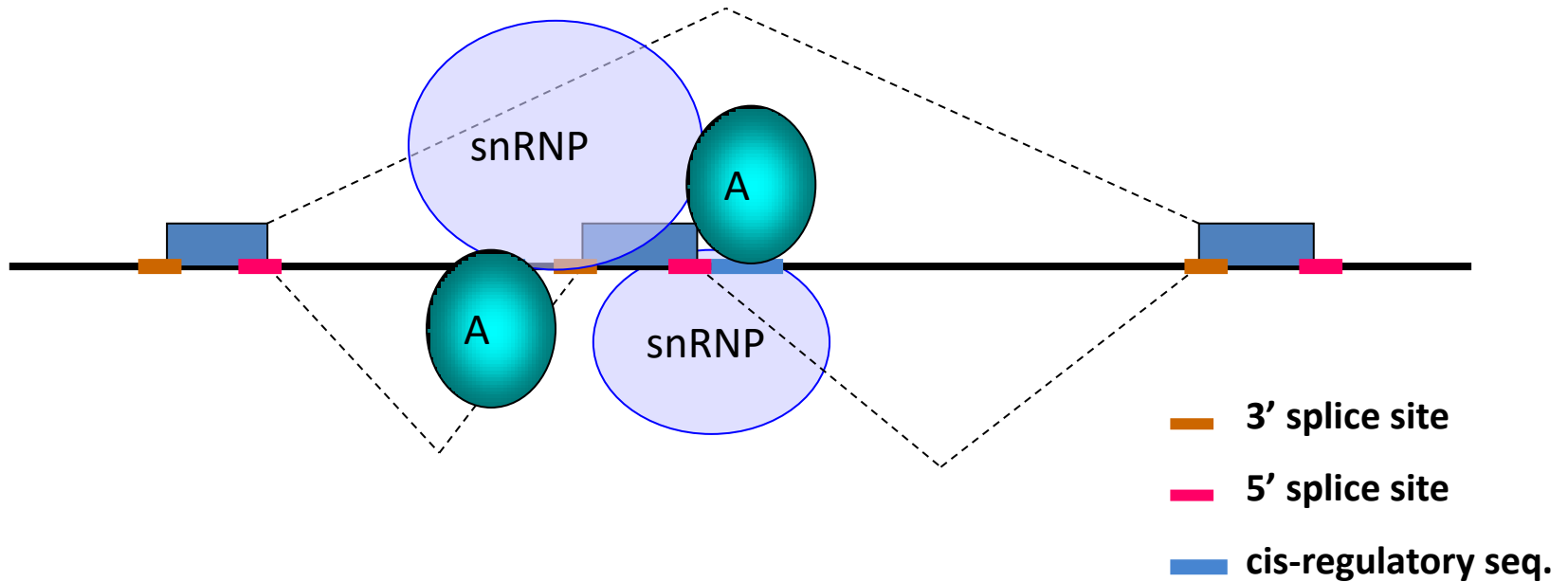
Mechanisms:

3' splice site

5' splice site

cis-regulatory seq.
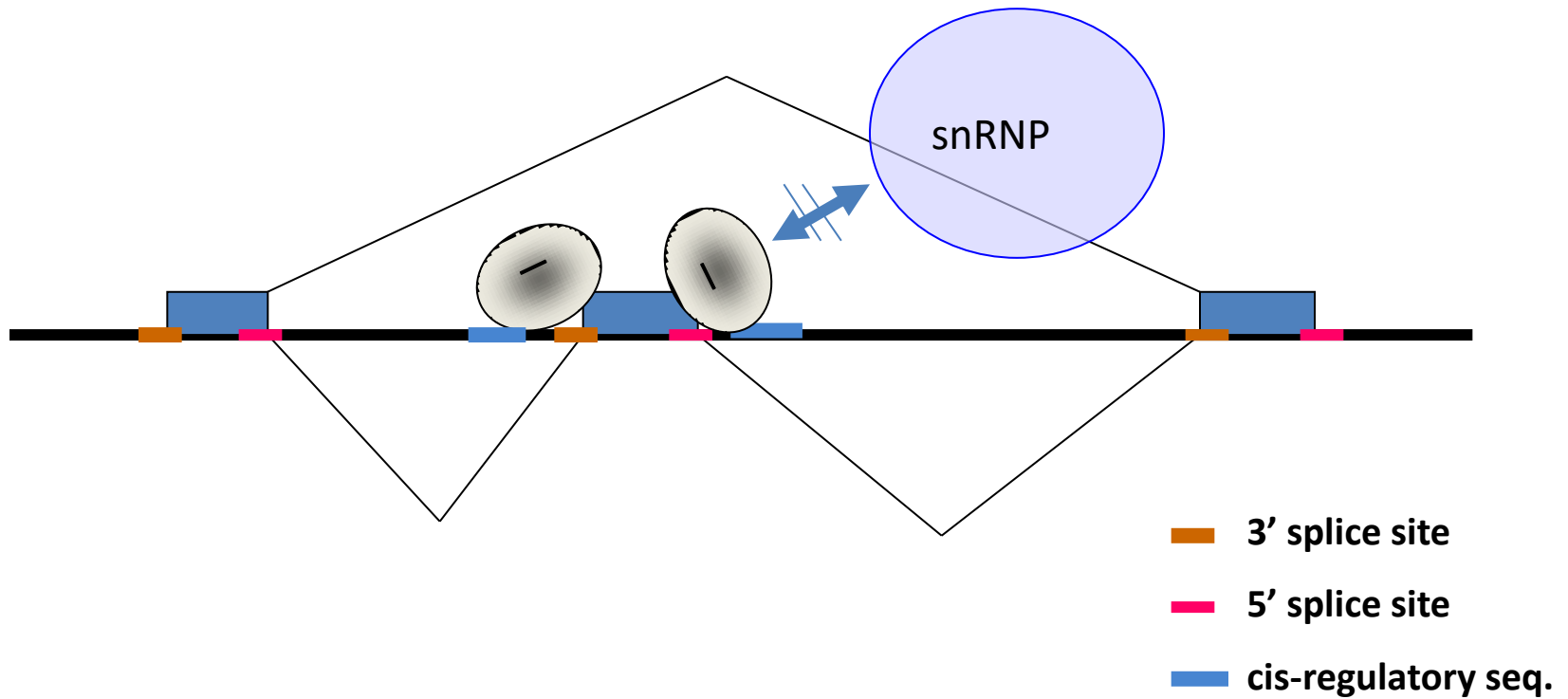
snRNP

1) Splicing sites are «strong», but they are occluded by inhibiting factor

Mechanisms:

3' splice site

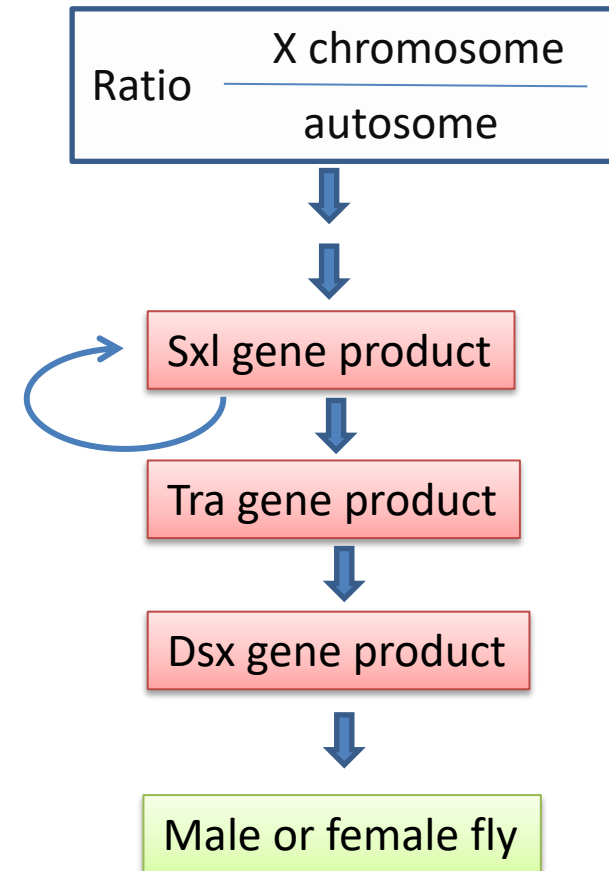5' splice site

cis-regulatory seq.
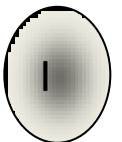
2) sites are «weak» but accessory *cis* sequences + factors will strengthen snRNP interaction.
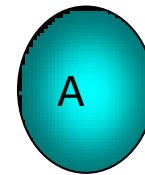
One of the first mechanisms studied illustrated exactly this simple situation.

Determination of sex in Drosophila gave the **first example** of exonic sequences enhancing a "poor" 3'-ss utilization as well as intronic sequences inhibiting 5'-ss

In Drosophila, the primary sex determinant is the X:A chromosome ratio.
This leads to a cascade of splicing regulatory signals, resulting in the production of two alternative splicing isoforms of the *dsx* transcription factor, repressing either female-specific or male-specific genes.

Ratio $\dfrac{\text{X chromosome}}{\text{autosome}}$

Sxl gene product

Tra gene product

Dsx gene product

Male or female fly

| GENE | trascritto primario di RNA MASCHILE X : A = 0,5 | trascritto primario di RNA FEMMINILE X : A = 1 |

sxl — Sex-lethal (Sxl)

sito di splicing 3' regolato
produzione di una proteina non funzionale

sito di splicing bloccato
proteina funzionale Sxl

tra — transformer (tra)

sito di splicing 3' regolato
produzione di una proteina non funzionale

sito di splicing bloccato
proteina funzionale Tra
Tra-2

dsx — doublesex (dsx)

il sito di splicing 3' regolato è inattivo

sito di splicing attivato

proteina Dsx N ———— C
400 aa
150 aa che sono specifici del maschio

proteina Dsx N ———— C
400 aa
30 aa che sono specifici della femmina

REPRIME I GENI DI DIFFERENZIAMENTO FEMMINILE

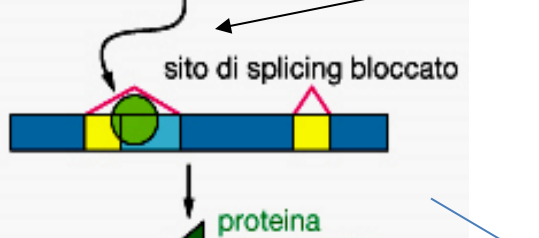REPRIME I GENI DI DIFFERENZIAMENTO MASCHILE

Male development

Female development

X:A=1 produces transient activation of an alternative promoter in the *sxl* gene, giving rise to a functional *sxl* protein
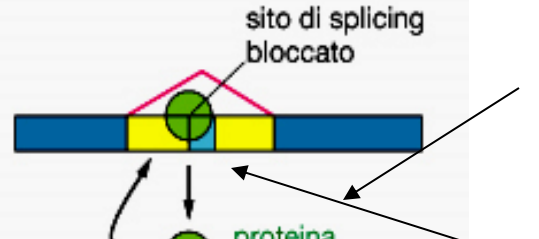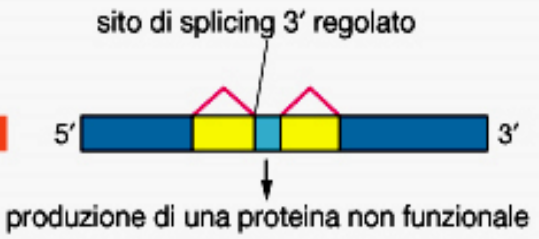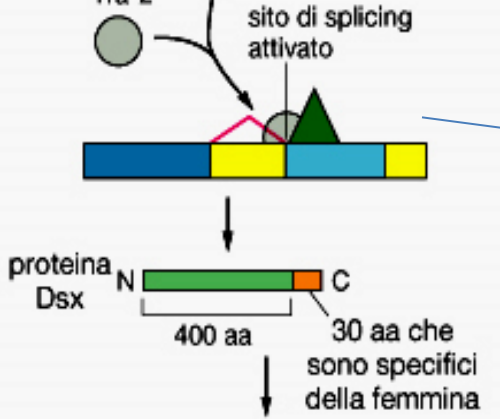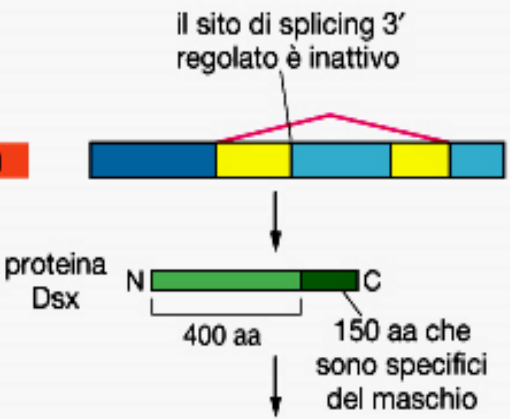
The *sxl* protein competes with U2AF for binding to the poly-pyrimidine tract

silencer

enhancer

....sorry for the Italian language, I did not find the slide in English

1st model - binding of a specific RBP to an intronic sequence overlapping poly-pyrimidine competes with U2AF65 binding and inhibits splicing


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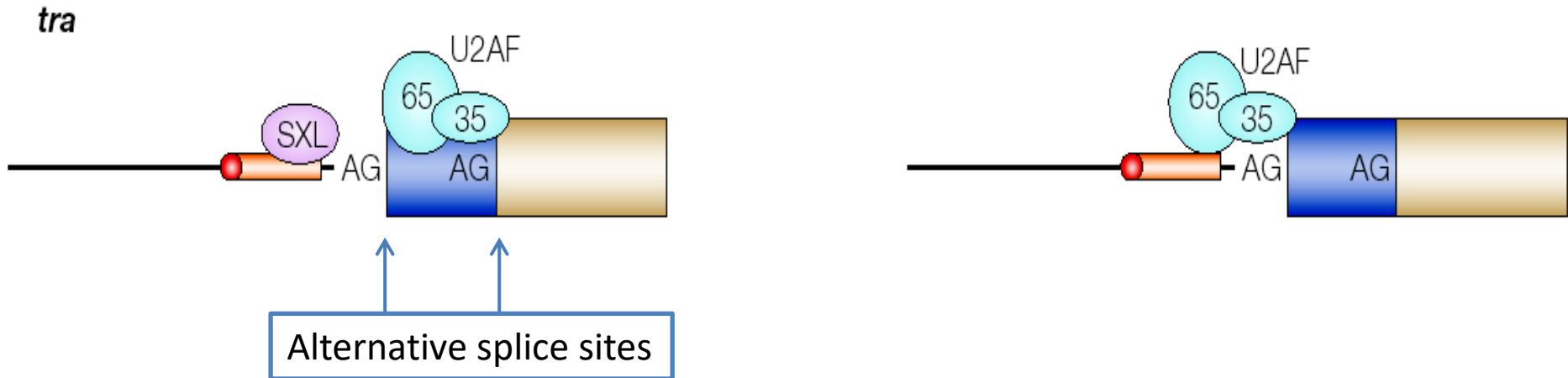
EXAMPLE: Repression of the non-sex-specific *tra* 3' splice site involves the interaction of SXL with an intron splicing silencer (ISS) embedded in the polypyrimidine tract and the prevention of U2AF binding. This leads to selection of the downstream female-specific 3' splice site.

The balance between SR and hnRNP proteins may explain some cases of alternative splicing

2nd model - competition between SR proteins and hnRNP. (hnRNP A1 multimerizes)
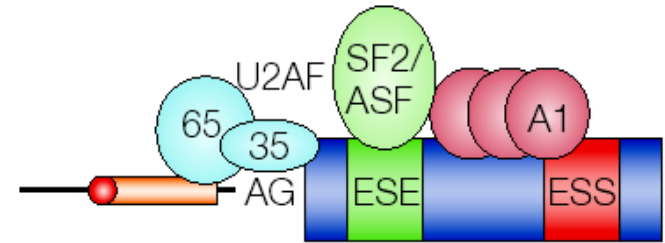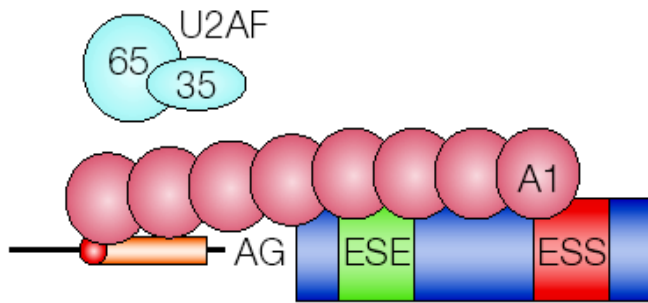
EXAMPLE: Inclusion of exon 3 of HIV1 *tat* pre-mRNA is determined by the nuclear ratio of specific heterogeneous nuclear ribonucleoprotein (hnRNP) and SR proteins. Propagative multimerization of hnRNPA1 from a high-affinity exon splicing silencer (ESS) is sterically blocked by the interaction of SF2/ASF with the upstream ESE. In this case, ESE function requires the RRM domains but not the RS domain of SF2/ASF.

3rd model - an RBP interacts with a small sequence in the pre-mRNA using an RRM domain and interacts with U1snRNP with another domain

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

A number of studies identified putative **motifs for positive and negative regulation** of splicing, classified on effect and position:

ESE = exonic splicing enhancer
ESS = exonic splicing silencer
ISE = intronic splicing enhancer
ISS = intronic splicing silencer

ESE (exonic splicing enhancers) were discovered and described first
        Note that the «code» is superimposed to coding sequences

ESE mutations were identified in various human diseases

When identified motifs are mapped to reference genes collection, it is observed that regulatory motifs spread a little distance from exons into introns, as predicted by mouse-human conservation analysis seen before.

Maps of potential regulatory sequences are obtained.

cis elements

Genome-wide identification of sequence elements that promote or repress splicing at adjacent «**ss**»

Approach 1 :   random select sequences that drive either exon inclusion or exon skipping in experimental model systems

Approach 2:   RNA immunoprecipitation using Abs against Splicing Factors (SR, hnRNP, other) and then NGS (RIP, CLIP, HITS-CLIP)

Note: Studies based on pure search of conserved motifs was not successfull: experimental + bioinformatics approaches needed

cis-regulatory element?
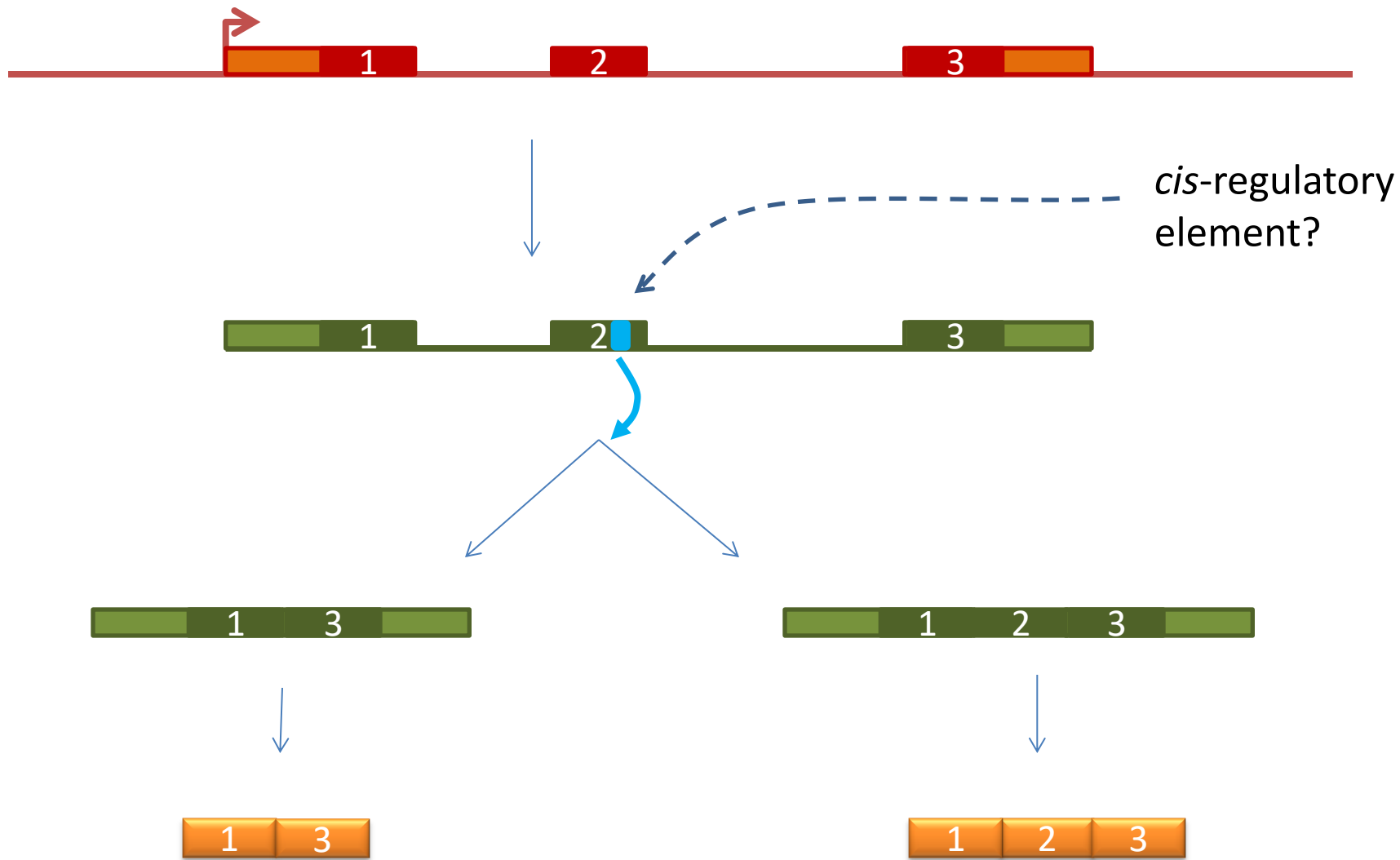
AS regulatory sequences characterized by de-novo functional assays
Example:

Systematic Identification and Analysis of Exonic Splicing Silencers

Zefeng Wang,[1] Michael E. Rolish,[1,2]
Gene Yeo,[1,3] Vivian Tung,[1]
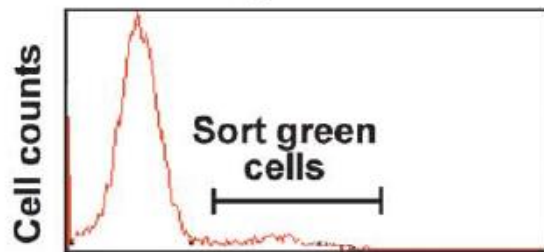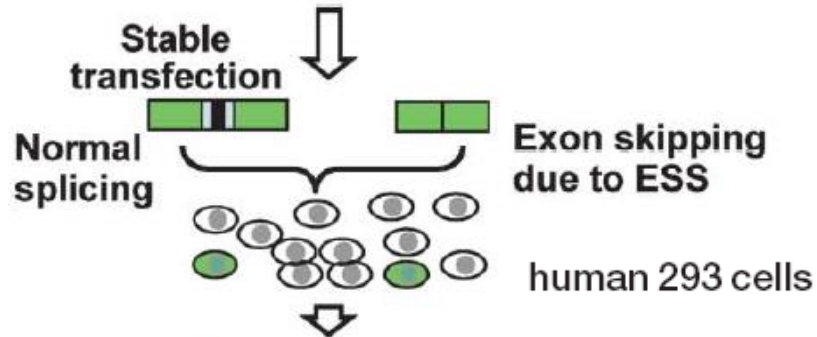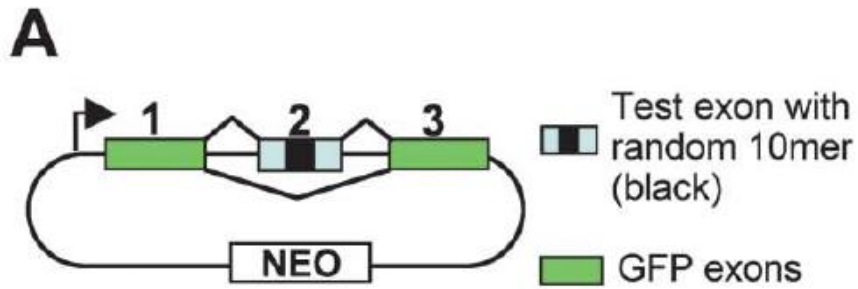Matthew Mawson,[1] and Christopher B. Burge[1,*]
[1]Department of Biology
[2]Department of Electrical Engineering
   and Computer Science
[3]Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Summary

Exonic splicing silencers (ESSs) are *cis*-regulatory elements that inhibit the use of adjacent splice sites, often contributing to alternative splicing (AS). To systematically identify ESSs, an in vivo splicing reporter system was developed to screen a library of random decanucleotides. The screen yielded 141 ESS decamers, 133 of which were unique. The silencer activity of over a dozen of these sequences was also confirmed in a heterologous exon/intron context and in a second cell type. Of the unique ESS decamers, most could be clustered into groups to yield seven putative ESS motifs, some resembling known motifs bound by hnRNPs H and A1. Potential roles of ESSs in constitutive splicing were explored using an algorithm, ExonScan, which simulates splicing based on known or putative splicing-related motifs. ExonScan and related bioinformatic analyses suggest that these ESS motifs play important roles in suppression of pseudoexons, in splice site definition, and in AS.

A

Test exon with random 10mer (black)

GFP exons

Stable transfection

Normal splicing

Exon skipping due to ESS

human 293 cells

Cell counts

Sort green cells

Log (GFP intensity)

Clone green cells, reconfirm by flow cytometry

PCR exon 2, sequence and analyze the ESS

random

......GATCCGNNNNNNNNNNTCTTCG......

exon 2 of the Chinese hamster dihydrofolate reductase (*DHFR*) gene, was used as the test exon.

D

Control   test 1   test2
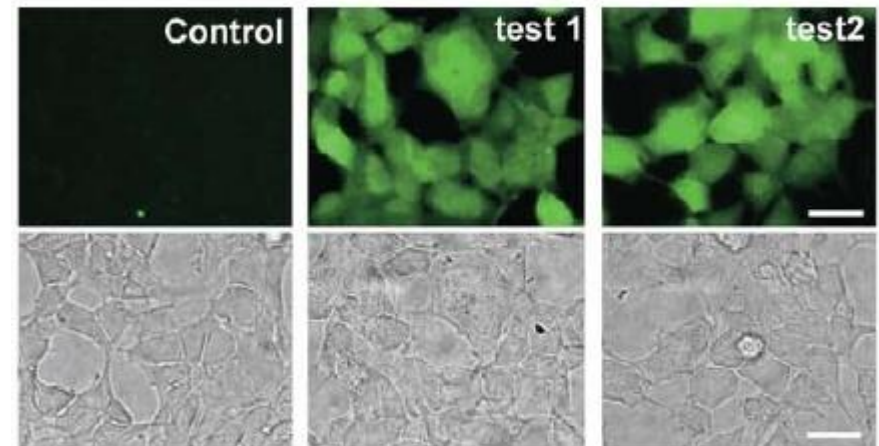
D) Microscopic images of transfected cells. Upper panel, GFP fluorescence. Lower panel, phase images. Scale bar, 50 μM.

Wang et al., 2004

Following selection, all positive cell clones are extracted → RNA
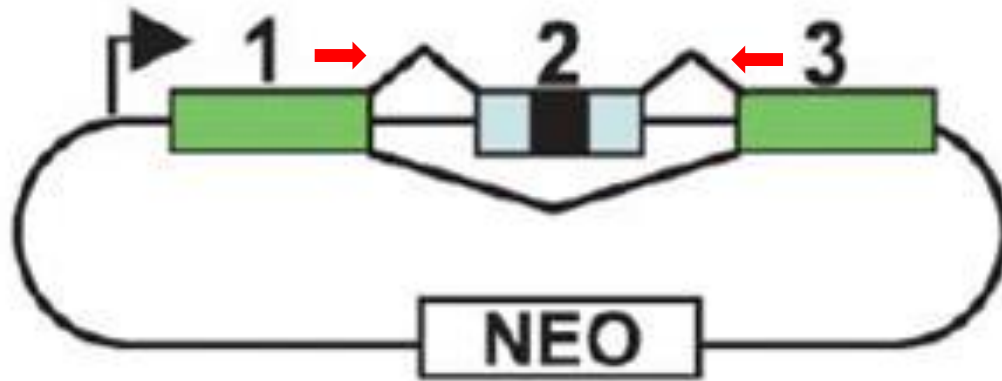
RT-PCR using primers pairs flanking the interrogated exon → cDNA library

The library is sequenced (clone-and-sequence or NGS) → a collection of sequences

Sequences are bioinformatically examined to find one or more «consensus» motifs
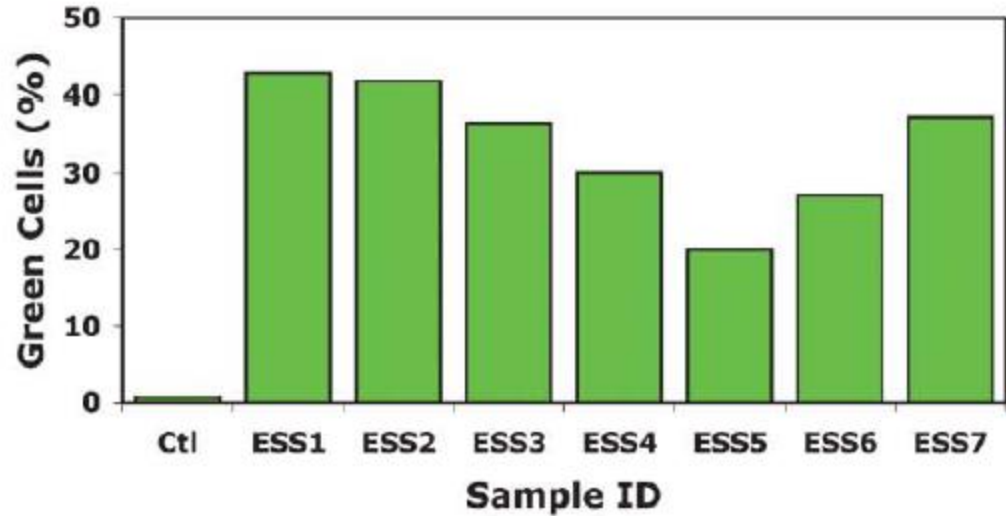Consensus motifs inserted in test exon and validated.

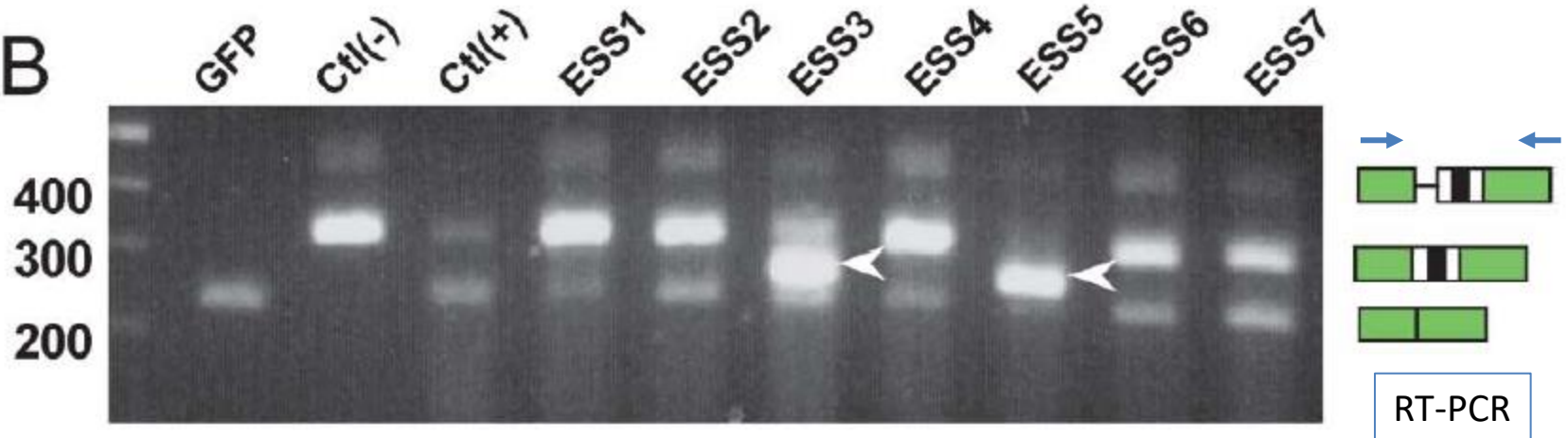Consensus motifs inserted in test exon and validated.

A consensus

ESS1	TTTGTTCCGT
ESS2	GGGTGGTTTA
ESS3	GTAGGTAGGT
ESS4	TTCGTTCTGC
ESS5	GGTAAGTAGG
ESS6	GGTTAGTTTA
ESS7	TTCGTAGGTA

B

RT-PCR

Regulatory sequences are found primarily close to the 5'-ss and 3'-ss i.e. around exons
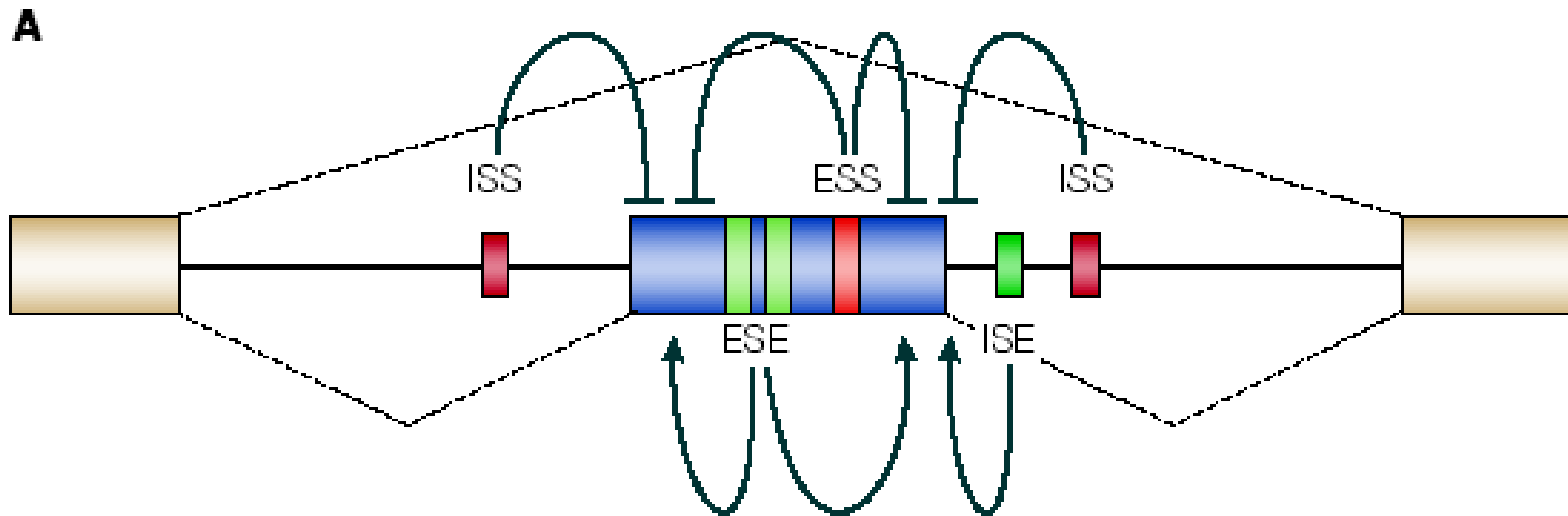

Figure 1 | **Elementary alternative splicing events and regulatory elements.**
A | In addition to the splice-site consensus sequences, a number of auxiliary elements can influence alternative splicing. These are categorized by their location and activity as exon splicing enhancers and silencers (ESEs 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.

Major SR proteins and hnRNP show almost ubiquitary expression, so that these proteins alone can hardly explain tissue-specific splicing

- Tissue-specific splicing
- Regulated splicing
- Epigenetic establishment of splicing patterns

1) **Tissue-specific splicing factors**
2) Signal transduction regulated factors
3) Chromatin effects on splicing choice

From: McManus & Graveley, COGD, 2011
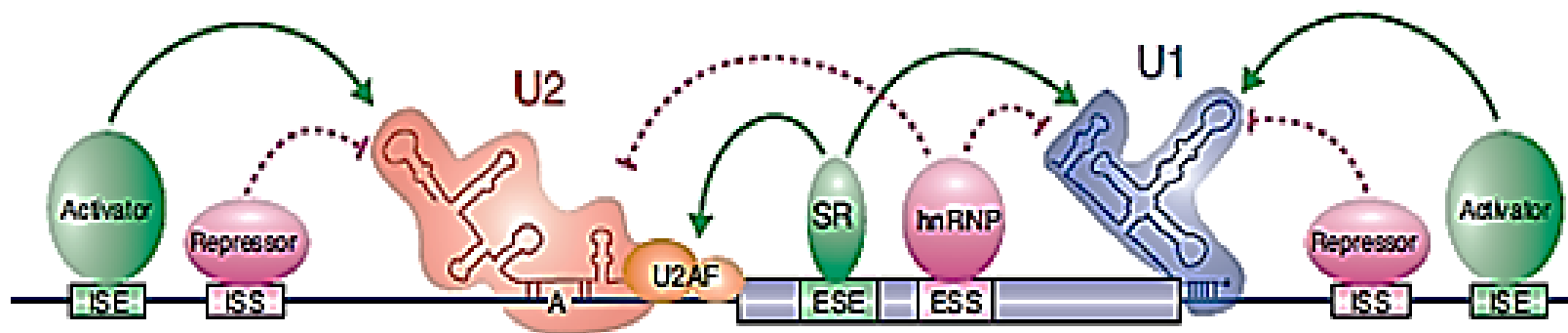
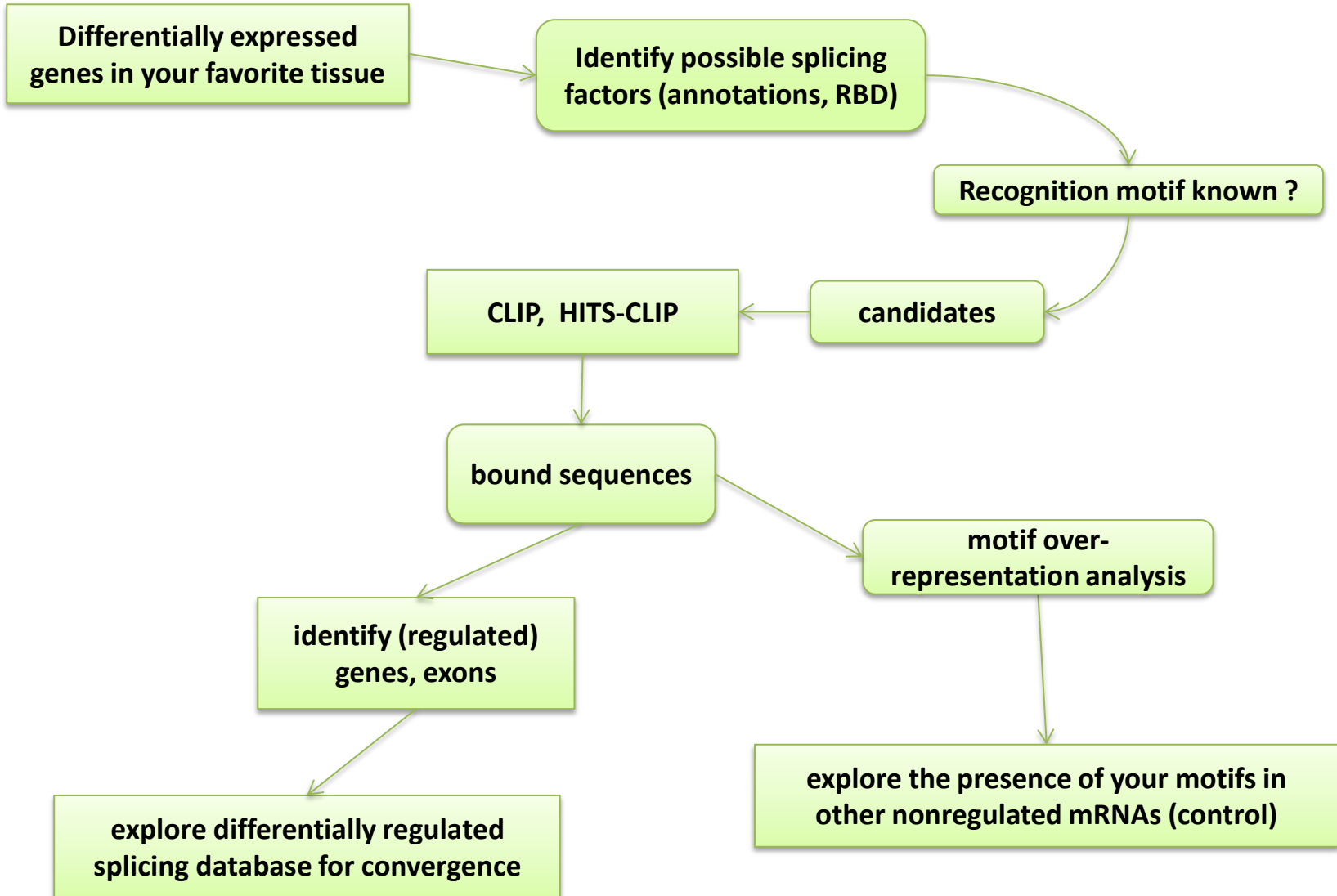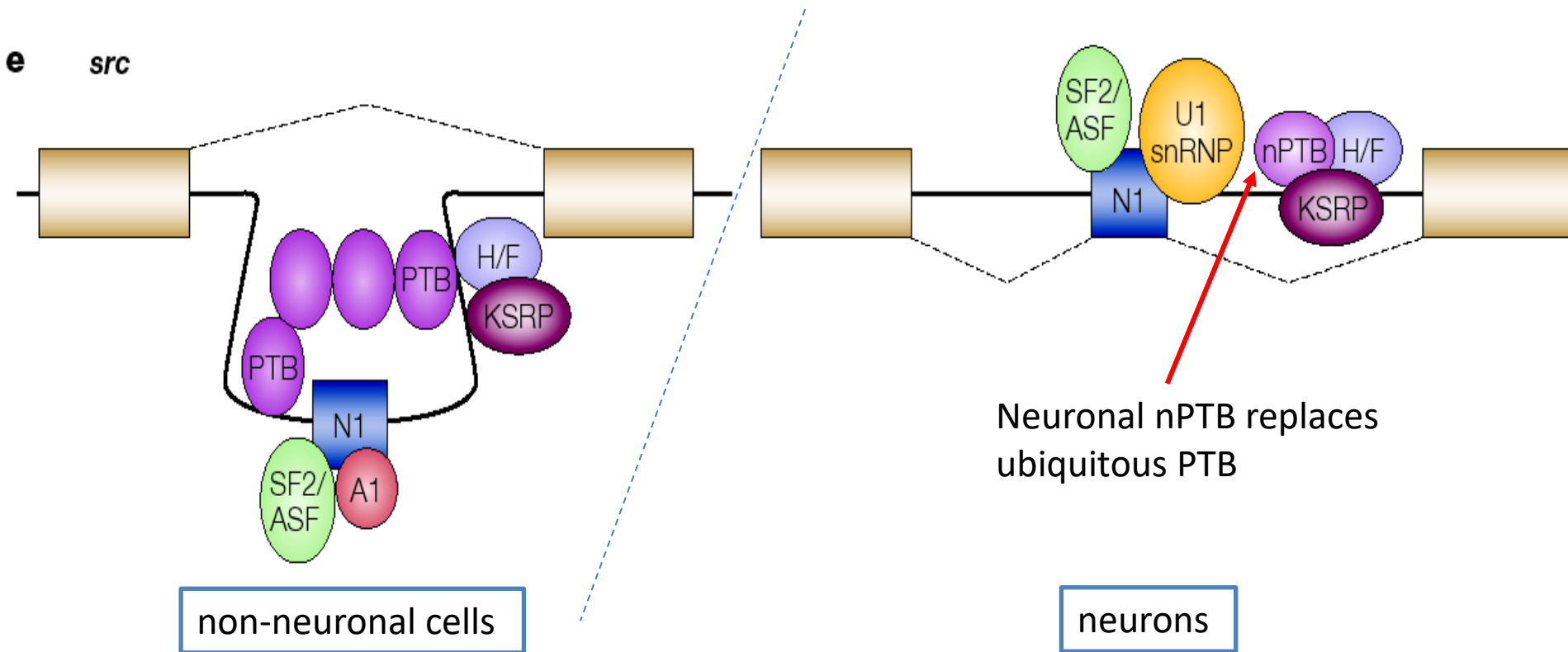# Appendix: The search for tissue-specific splicing factors

Table 2 | Tissue-specific alternative splicing factors

Name	Other names	Binding domain	Binding motif	Tissue expression	Target genes
nPTB	brPTB and PTBP2	RRM	CUCUCU	Neurons, myoblasts and testes	*BIN1, GLYRA2, ATP2B1, MEF2, NASP, SPAG9* and *SRC*
NOVA1	NA	KH	YCAY	Neurons of the hindbrain and spinal cord	*GABRG2, GLYRA2* and *NOVA1*
NOVA2	NA	KH	YCAY	Neurons of the cortex, hippocampus and dorsal spinal cord	*KCNJ, APLP2, GPHN, JNK2, NEO, GRIN1* and *PLCB4*
FOX1	A2BP1	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.

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

non-neuronal cells

neurons

Neuronal nPTB replaces ubiquitous PTB

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.

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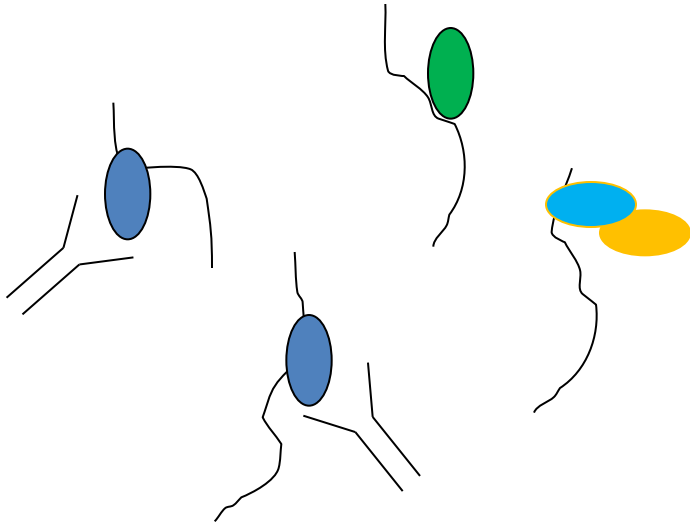
Nova: the first vertebrate tissue-specific
splicing factors (neurons)
(Nova1 – Nova2)

IMPT
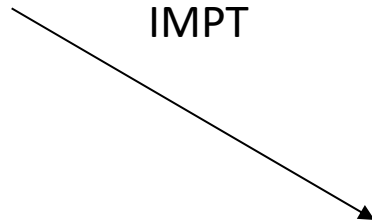
RNA identified on microarrays or NGS

48 targets identified in previous studies

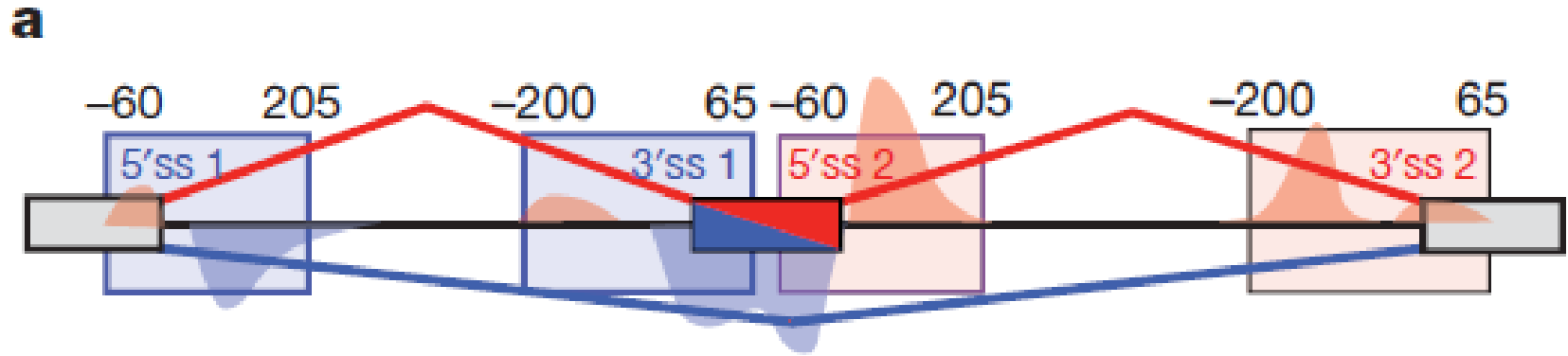# Clustering of "YCAY" Nova recognition sequences in 48 Nova-regulated exons

Figure 1 | Definition of the Nova–RNA binding map. a, A generic pre-mRNA 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)**.

The splicing regulatory effect of Nova 1-2 depends on the **position** of its cognate binding site relative to alternative exons.

Examples of predicted Nova-regulated exons: analysis in brain tissues from Nova1$^{-/-}$ / Nova2$^{-/-}$ double K.O. mice (dKO).
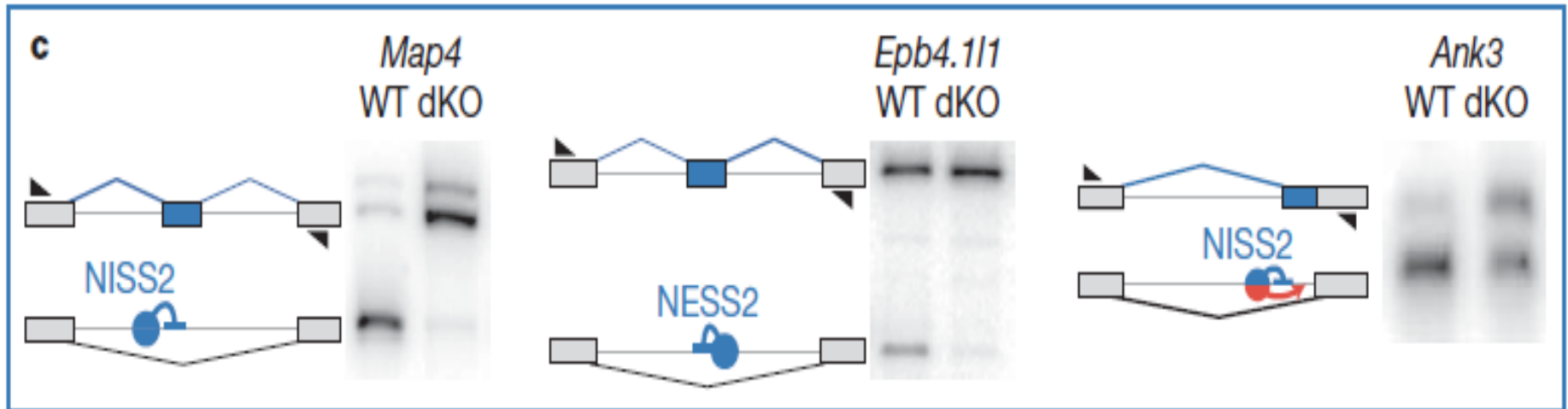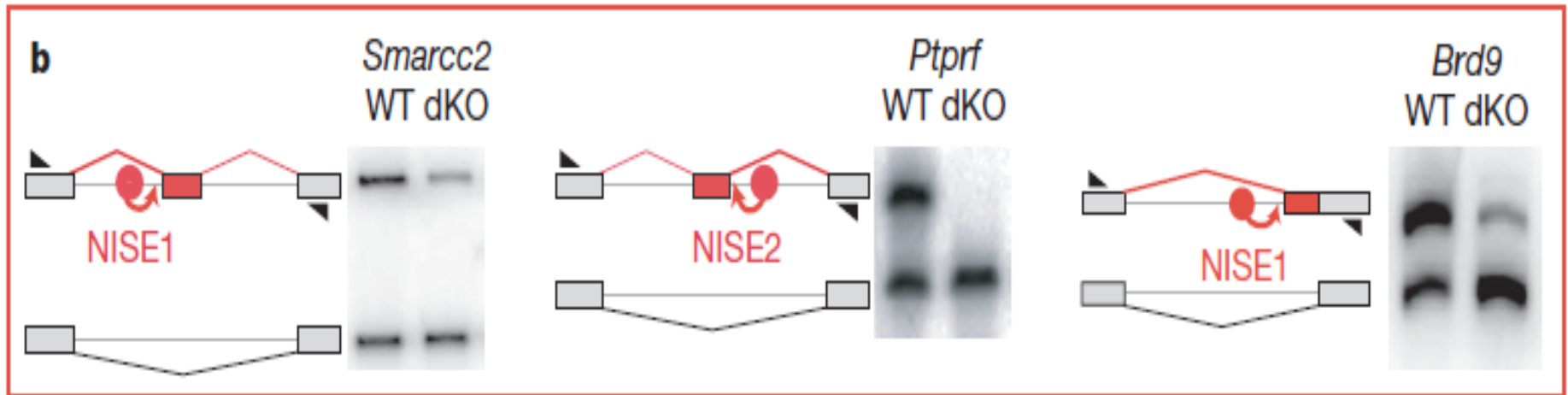
Indicate the primers used for RT-PCR analysis

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.
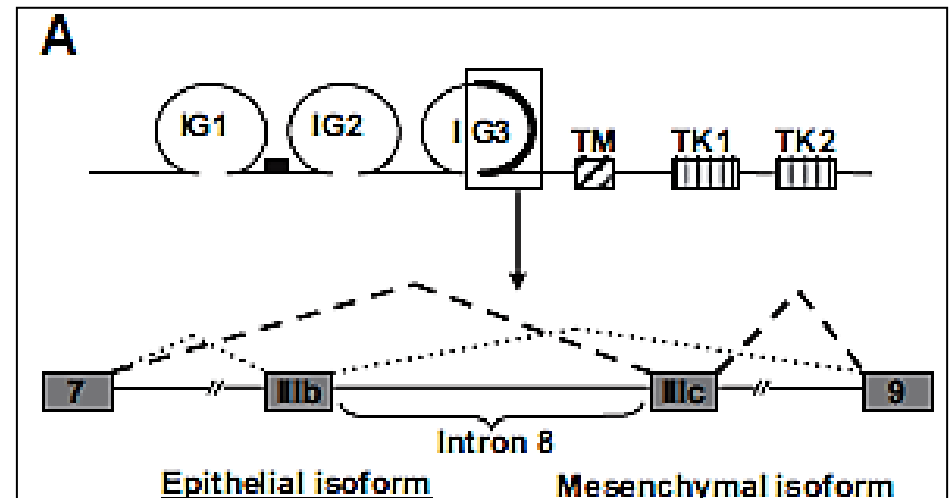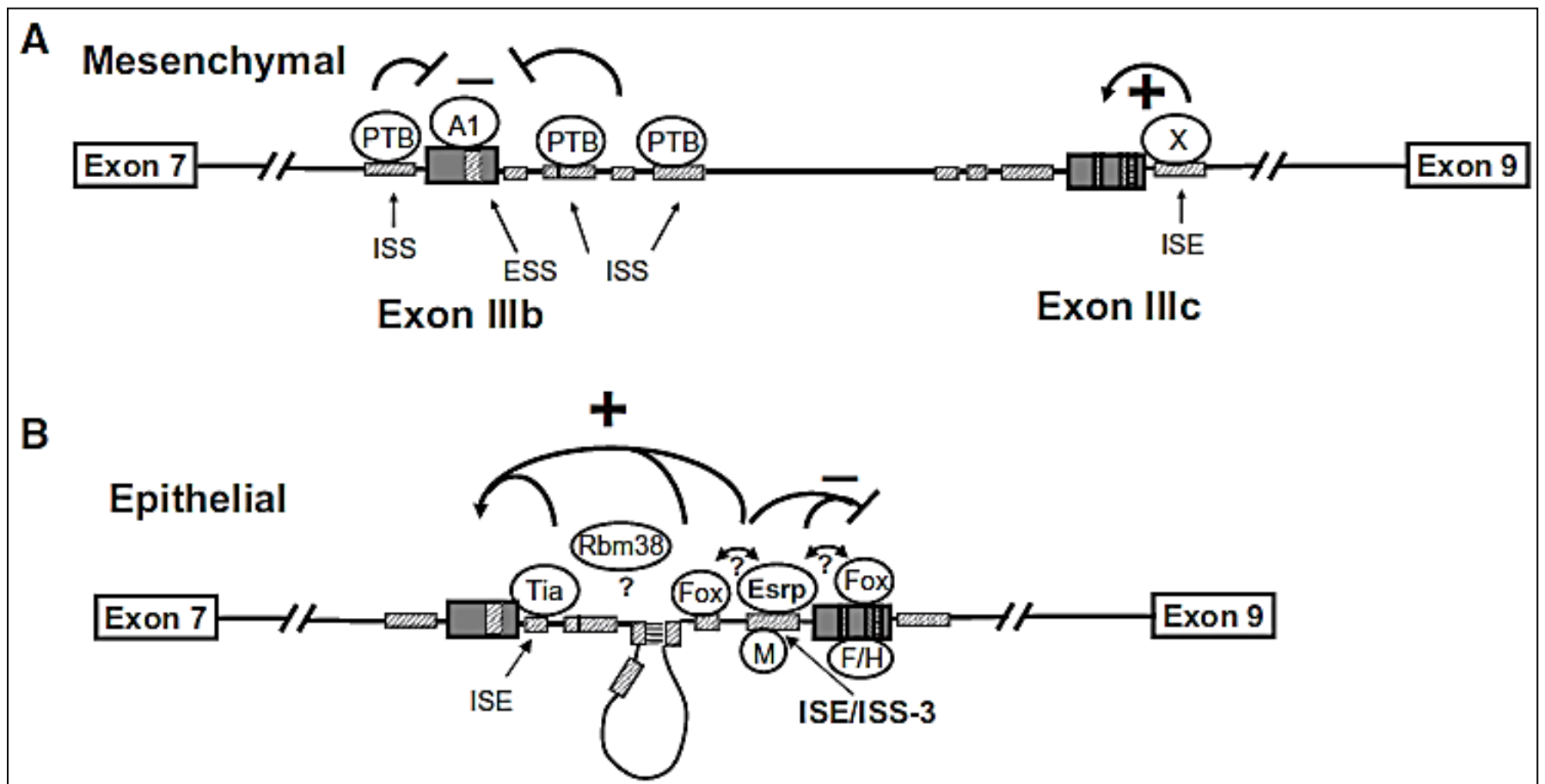
Figure 6. Model for the Mechanism of FGFR2 Cell-Type-Specific Splicing. (A) Combinatorial control by ubiquitous regulatory proteins favors exon IIIb silencing and exon IIIc inclusion in mesenchymal cells. Crosshatched boxes indicate known FGFR2 auxiliary ciselements. PTB, polypyrimidine tract binding protein; A1, hnRNPA1; X, Unknown factor(s) that bind an ISE downstream of exon IIIc. (B) In epithelial cells, the ESRPs collaborate with other regulatory proteins to activate exon IIIb splicing and silence exon IIIc splicing. Potential interactions between these proteins are indicated by double arrows and question marks. M, hnRNP M; Tia, Tia1 or TiaR; Fox, Fox family members; F/H, hnRNP F or H. A base-pairing interaction between two complementary sequences in the intron (ISE2 and ISAR) likely serves to position ESRP1 more closely to exon IIIb where it can act in conjunction with other factors (such as RBM38 and Tia1) to activate splicing (Muh et al., 2002). Note that proteins other than the ESRPs shown binding to FGFR2 transcripts only in mesenchymal or epithelial cells are expressed in both cell types and may be bound to the cognate elements in either cell type.

AS and nucleosomes

Third point, nucleosomes are best positioned on exons

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.
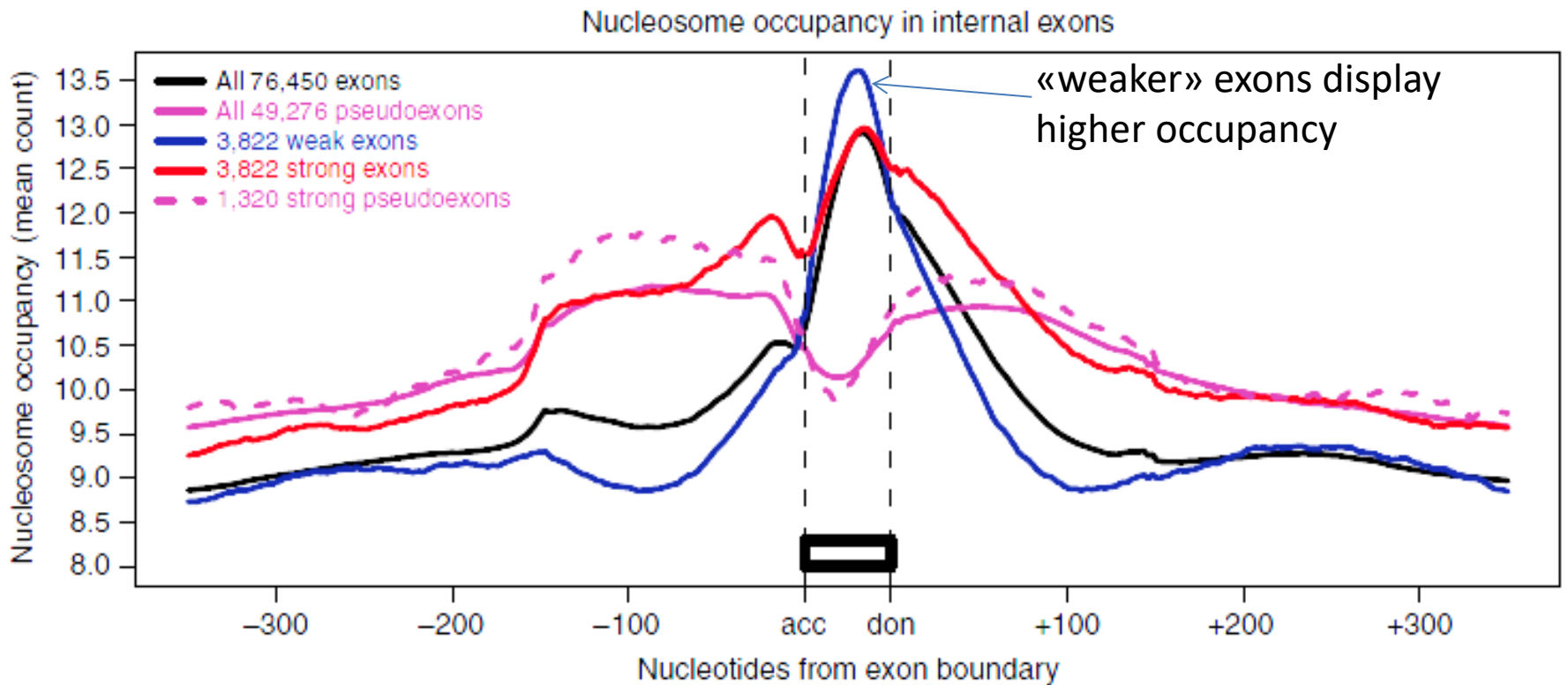
Figure 1 Observed and predicted nucleosome occupancy. (**a**) Nucleosome-occupancy profile across human internal constitutive exons in resting CD4+ T cells. We have computed the number of extended nucleosome reads overlapping each nucleotide. Upstream and downstream of an **idealized internal exon**, we plot the average number of nucleosome reads per nucleotide position, with negative positions relative to the acceptor (acc) site and positive positions relative the donor (don) site. Within the exon, reads have been mapped to 50 identically spaced intervals, irrespective of the length of the exon (see Online Methods). Strong exons are exons with a combined donor and acceptor score among the highest 5%; weak exons are the exons with a combined score among the lowset 5%; pseudoexons are intronic sequences bounded by splice sites; strong pseudoexons are exons with a combined score higher than the 90% percentile of real exons.
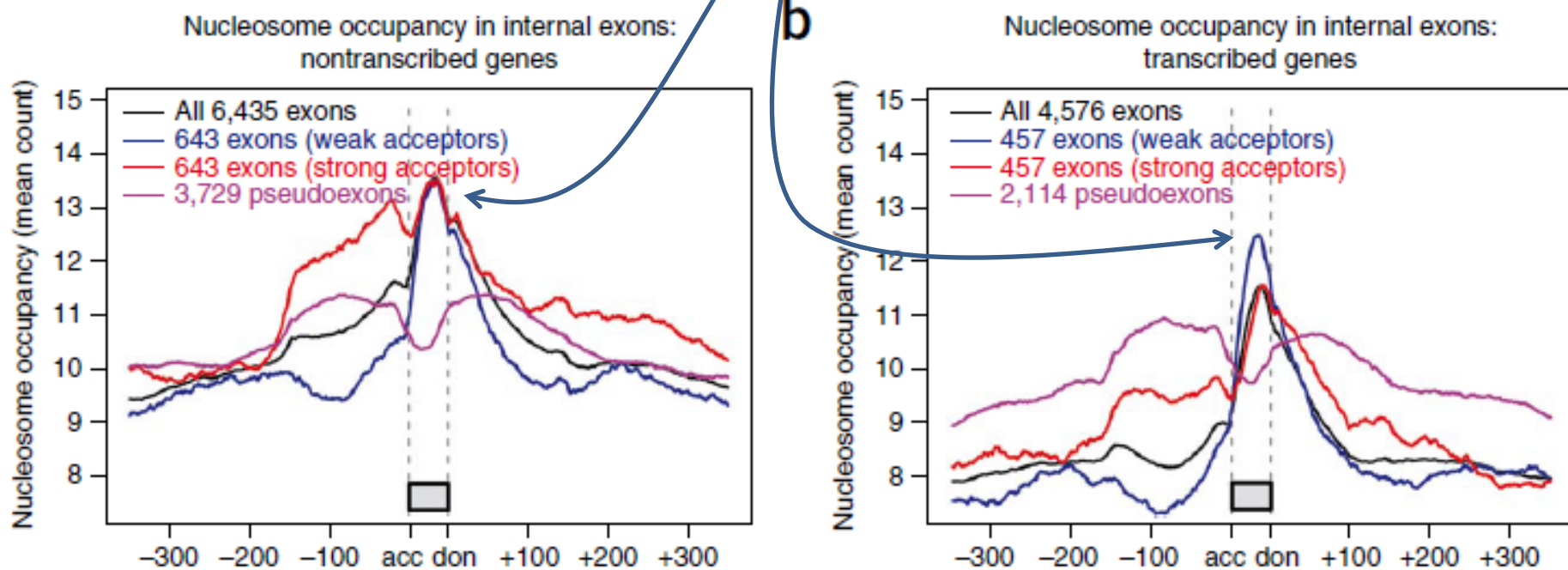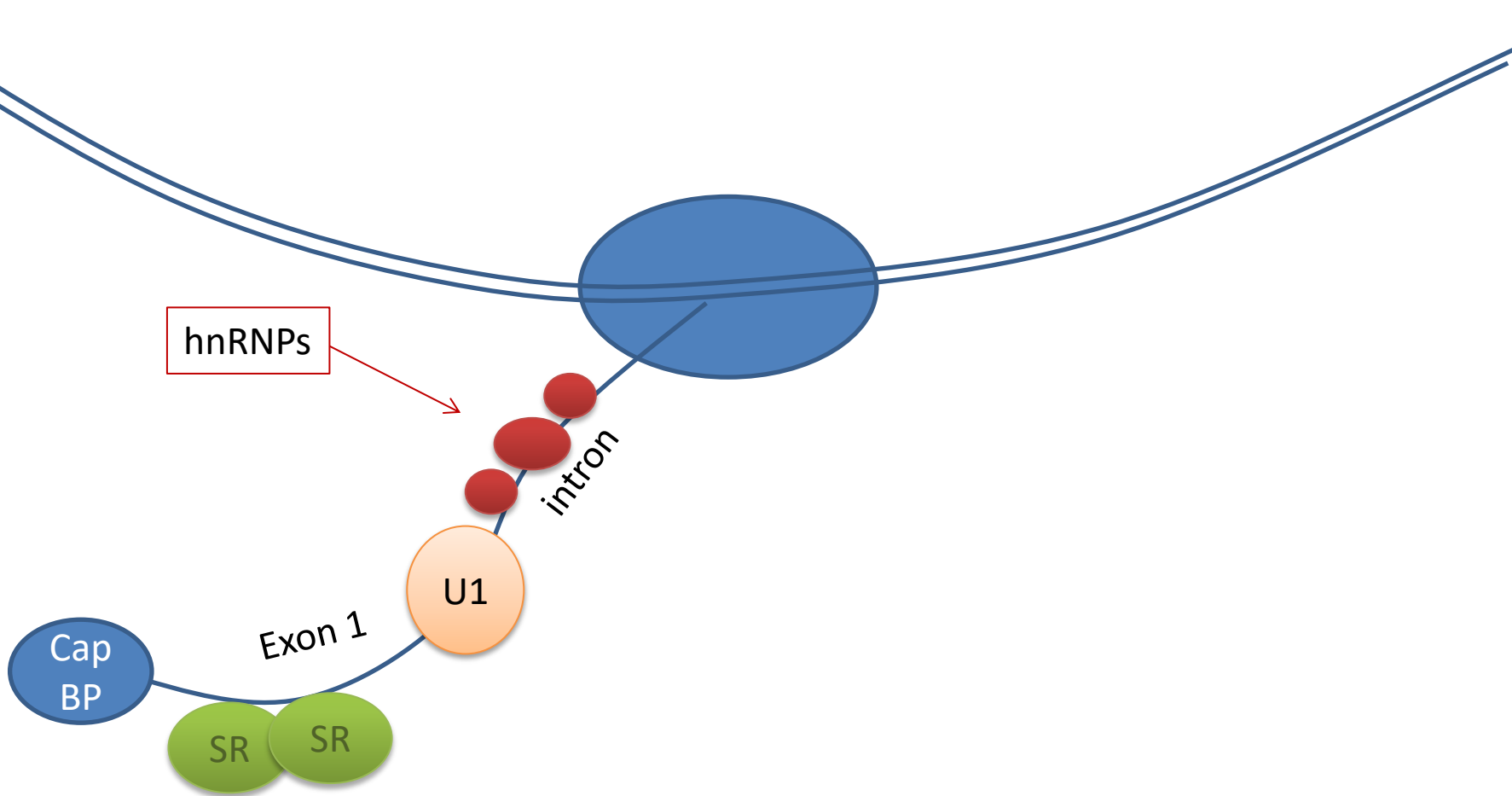
Tilgner et al., 2009

Figure 2 Nucleosome occupancy and expression of genes and exons. (**a**) Nucleosome-occupancy profile across internal acceptor sites from genes that are not expressed in resting CD4+ T cells. Gene expression has been determined using the Affymetrix platform. We plot the average number of nucleosome reads per position in all exons considered together (black), only in exons with strong (red) and weak (blue) acceptor sites, and in intronic pseudoexons. (**b**) Nucleosome-occupancy profile across internal acceptor sites from genes expressed in resting CD4+ T cells, shown as in **a.**
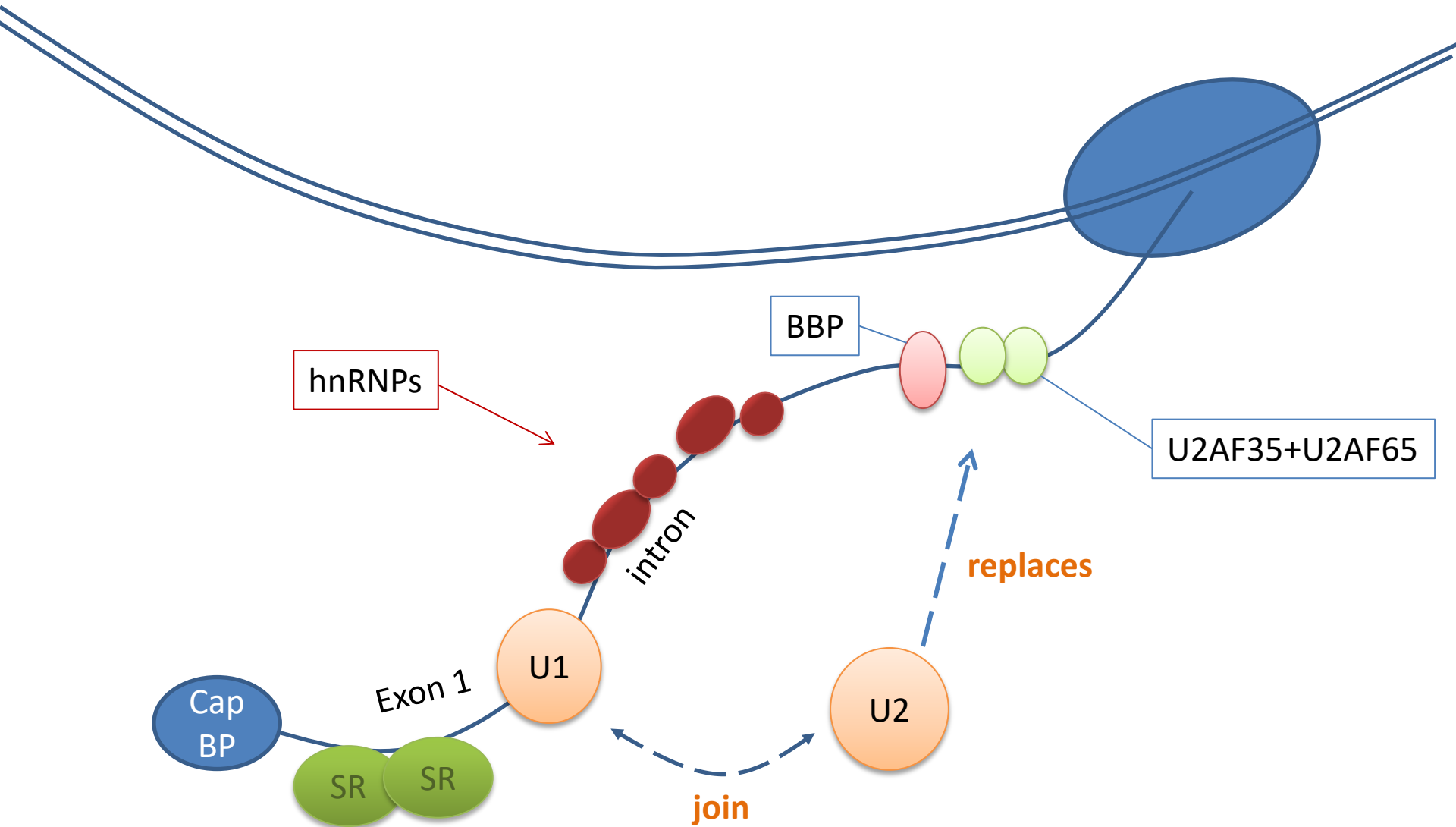
Tilgner et al., 2009

Splicing regulation through trenascriptional regultion

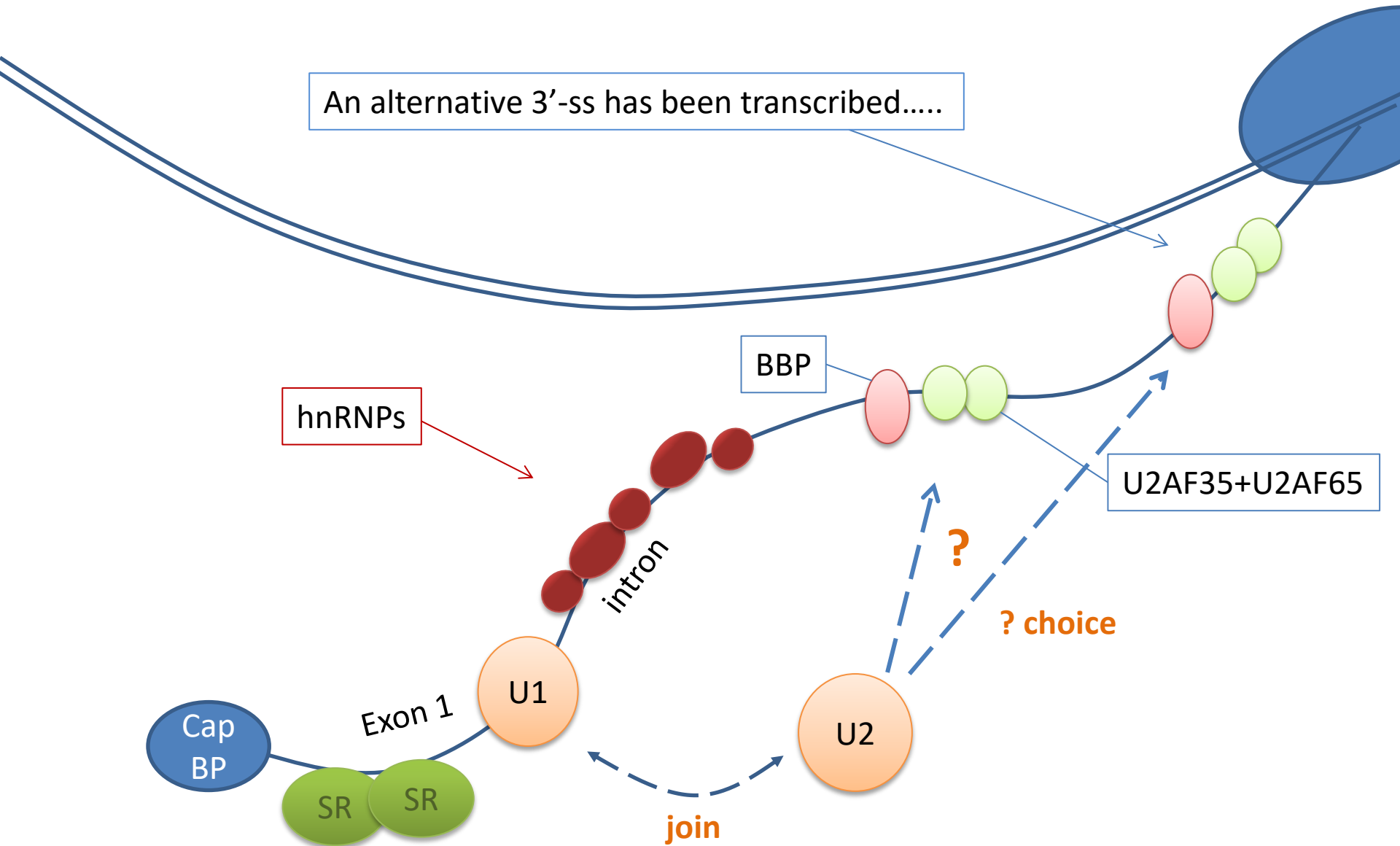Chromatin status influence – epigenetic mechanisms

The kinetic competition model

Exon definition

hnRNPs

intron

U1

Exon 1

Cap
BP

SR SR

Michele De Bortoli, UniTo, 2015

Intron definition

hnRNPs

BBP

U2AF35+U2AF65

intron

Exon 1

Cap BP

SR   SR

U1

U2

join

replaces

Michele De Bortoli, UniTo, 2015

Alternative Splicing

An alternative 3'-ss has been transcribed…..

hnRNPs

BBP

U2AF35+U2AF65

intron

Exon 1

Cap BP

SR   SR

U1

U2

?

? choice

join

Michele De Bortoli, UniTo, 2015

The « first arrived, first served » model

Is compatible with the fact that, in general, splicing is well ordered and that Alternative Splicing normally concerns a choice between two consecutive splicing sites, only very rarely separated by multiple exons/introns

This model is also compatible with the observed cases in which specific intervention of a ESE-ISE/SR or ISS-ESS/hnRNP takes place (see above).

In special cases, two or more unspliced sites can be present in the mRNA, then true competition is plausible

How un-spliced sites may arise ?

Of course, <u>inhibitory RBPs</u> as seen before

The <u>speed of the RNA Polymerase</u>
- elongation factors,
- RNA 2ary structure,
- nucleosome stability

Also, remember that RNA Pol II carries splicing factors loaded on CTD

Effects of promoters and enhancers

Inclusion of weak exons may depend on the ability of **promoters** to regulate :

✓ the "speed" of RNA Polymerase II (the rate at which Pol II synthesizes RNA is variable and discontinuous, and may depend on elongation factors that are «charged» on RNA Pol II by different promoters/enhancers)

✓ the «loading» of RNA Polymerase II CTD with specific Splicing Factors (interaction of several splicing factors with Pol II CTD identified by co-IP)

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.

The Speed Bump model was born in Kornblihtt lab more than a decade ago.

Creating a splicing reporter with a weak exon
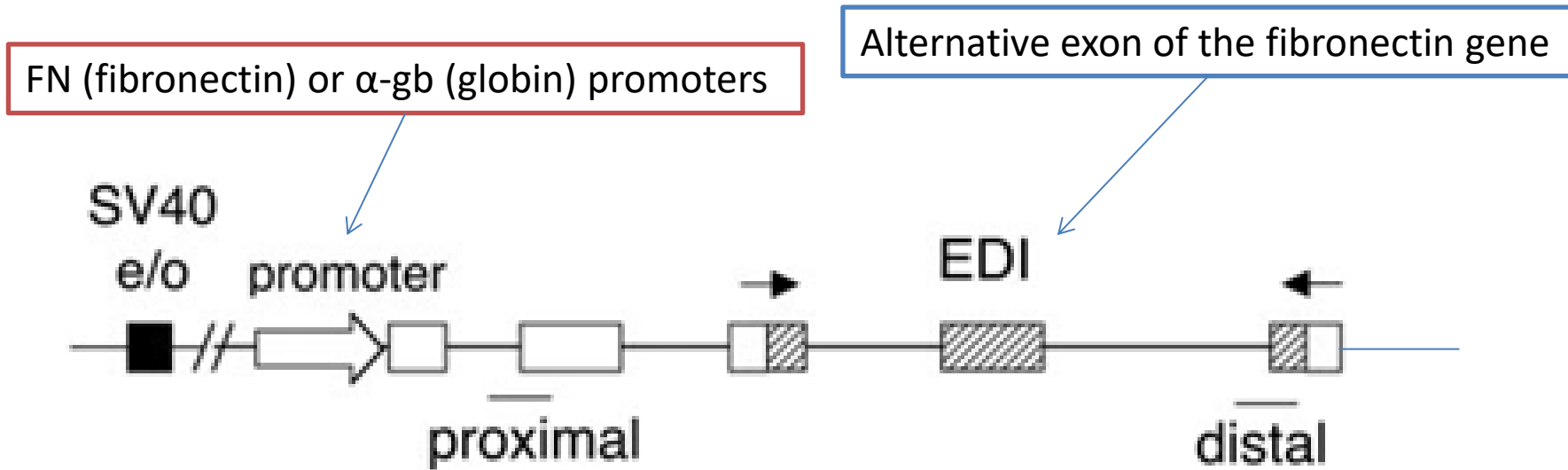
A) Scheme of the minigenes transfected to assess alternative splicing. Open exons, human-gb; dashed exons, human FN; black box, SV40 e/o; arrows, primers used to amplify the mRNA splicing variants by RT-PCR, and lines, proximal and distal probes used for RPA.
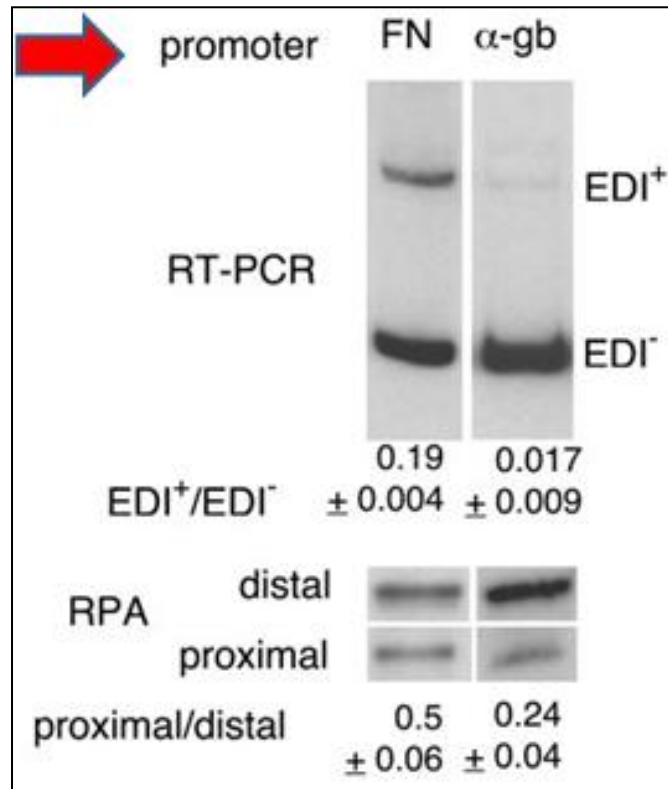
Hep3B cells were transfected with 600 ng of pSVEDAFN (FN promoter) or pSVEDATot (-gb promoter) plus 400 ng of pCMVgal. RNA splicing variants were detected by radioactive RT-PCR and analyzed in 6% native polyacrylamide gels. Ratios between radioactivity in EDI bands and radioactivity in EDI bands are 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.
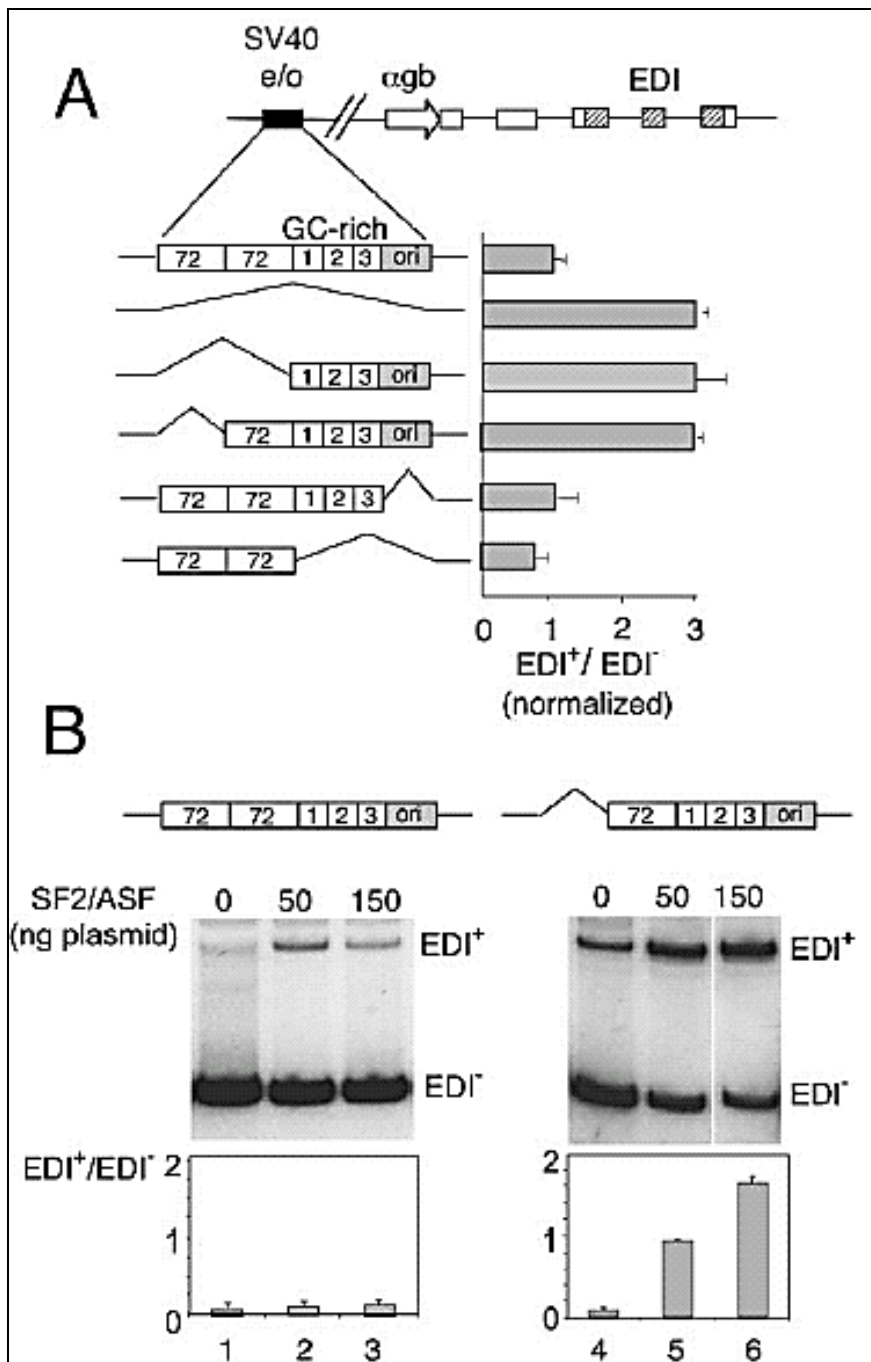
Fig. 3. (*A*) Deletion analysis of the SV40 eo with respect to alternative splicing of the EDI exon. Horizontal bars indicate normalized EDIEDI ratios of Hep3B cells transfected with a series of α–gb promoter constructs carrying different internal deletions of the SV40 eo. Results correspond to the mean SD of at least three independent transfection experiments.

(*B*) Deletion of only one 72-bp repeat confers responsiveness to SF2ASF to the α-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.
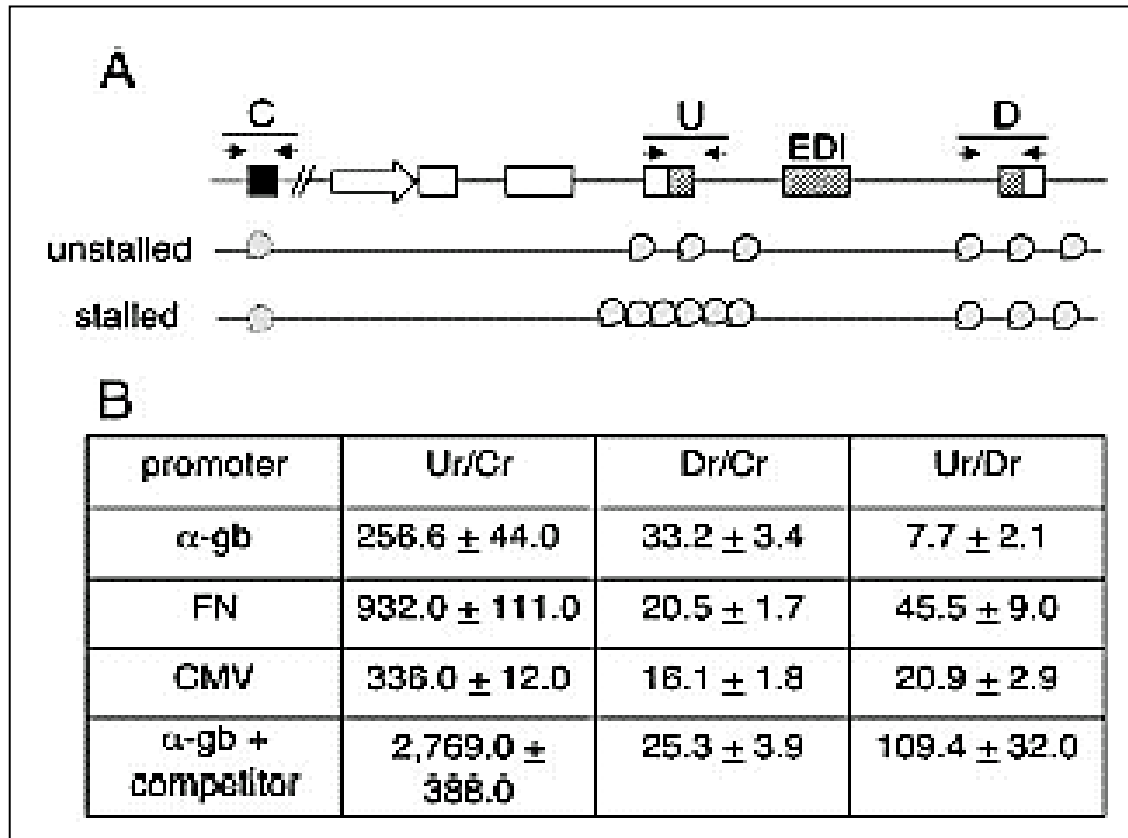
Fig. 5. ChIP with an Ab to RNA pol II. (*A*) Scheme of the minigenes transfected to assess pol II densities. Arrows indicate the pairs of primers used in real time PCRs to quantitatively amplify DNA that is bound to the immunoprecipitated pol II, at two regions mapping U, D of the EDI alternative exon, and at a third C region outside of the transcription unit. (*B*) Cells were transfected with -gb, FN, or CMV promoter constructs and, where indicated, co-transfected with a 10-fold molar excess of a <u>competitor plasmid carrying the SV40 e/o</u>. After 48 h, cells were fixed with formaldehyde and treated for ChIP and real time PCR analysis as described in *Experimental Procedures*. UrUimUin; CrCimCin; DrDimDin 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 SD of three real time PCR determinations.

This study demonstrated that:

1. Promoters/enhancers can influence the rate of inclusion of weak exons

2. They do so in part by influencing the speed of RNA Pol II

3. When a weak exon is included, RNA PolII «slows down» in proximity of that exon (*remember that weak exons have more strongly positioned nucleosomes !*)

… and RNA Pol II CTD carries splicing factors.
All this led to:

the speed-bump model

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

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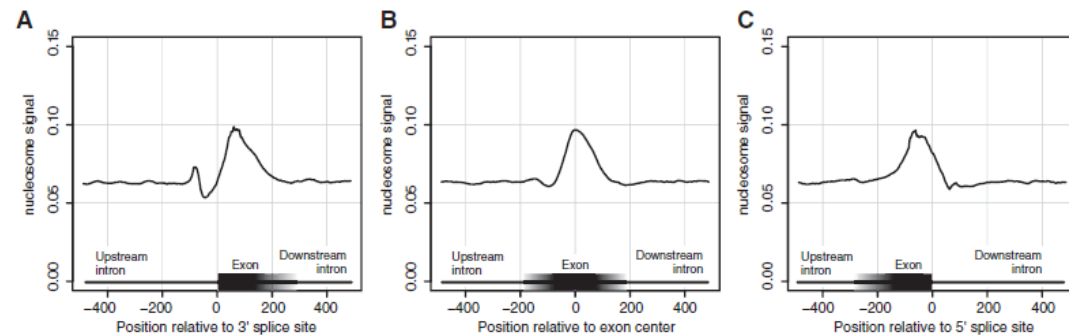
nucleosomes are positioned at exons
(data from MNase-Seq experiments)

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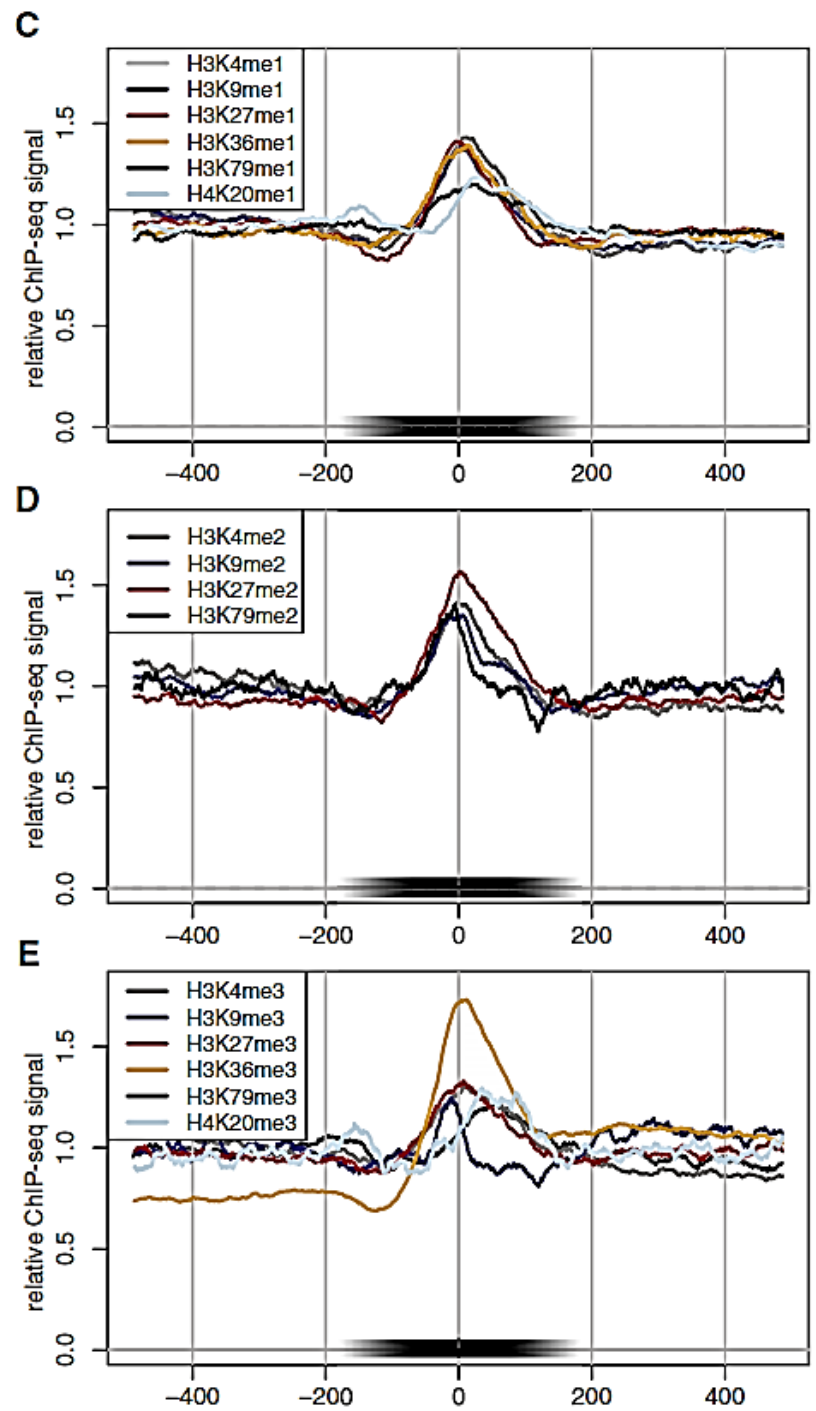

Exon/intron ratio

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**).
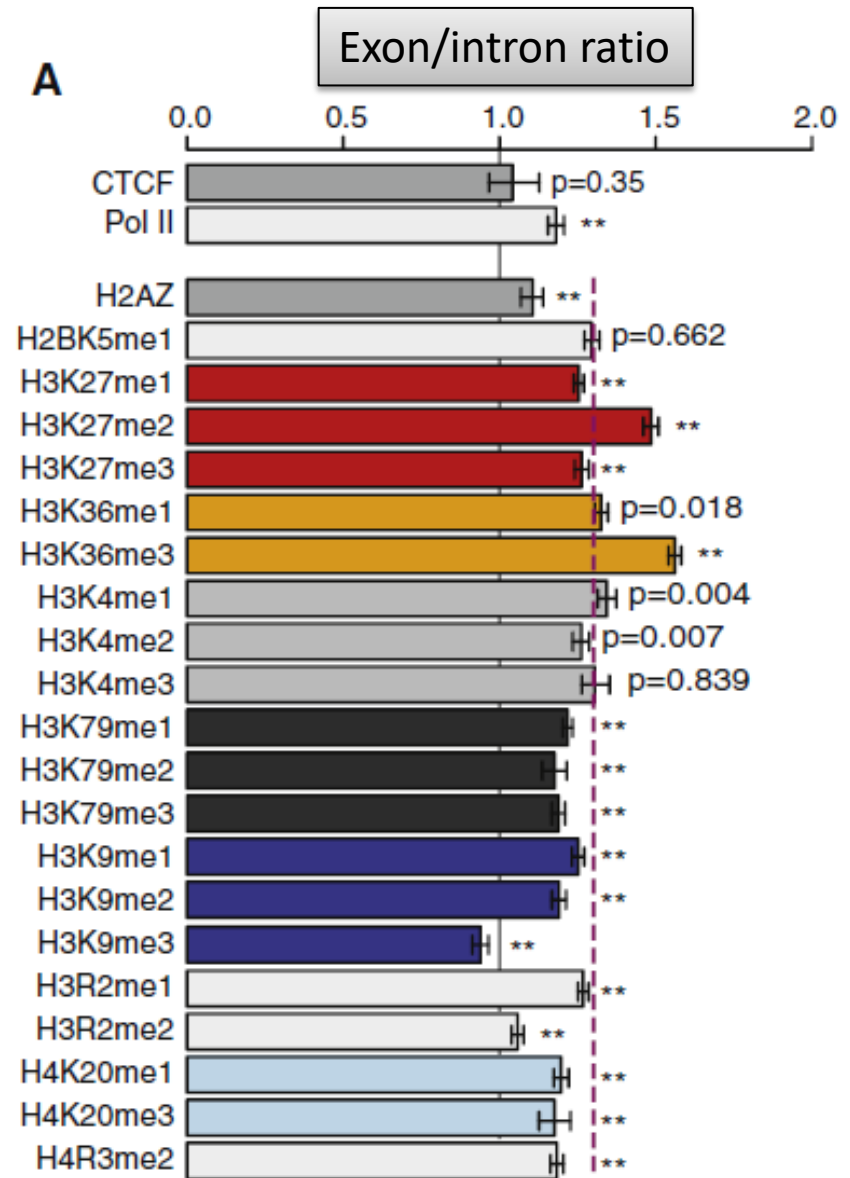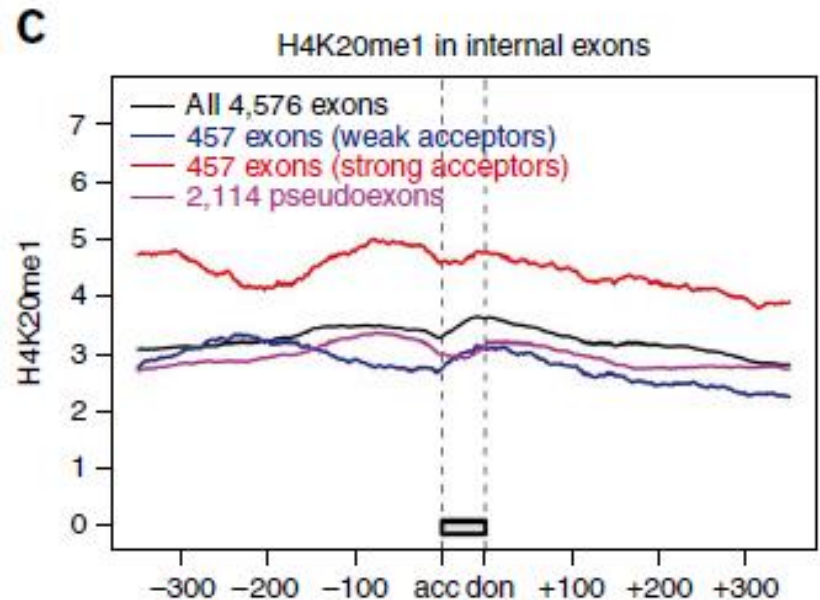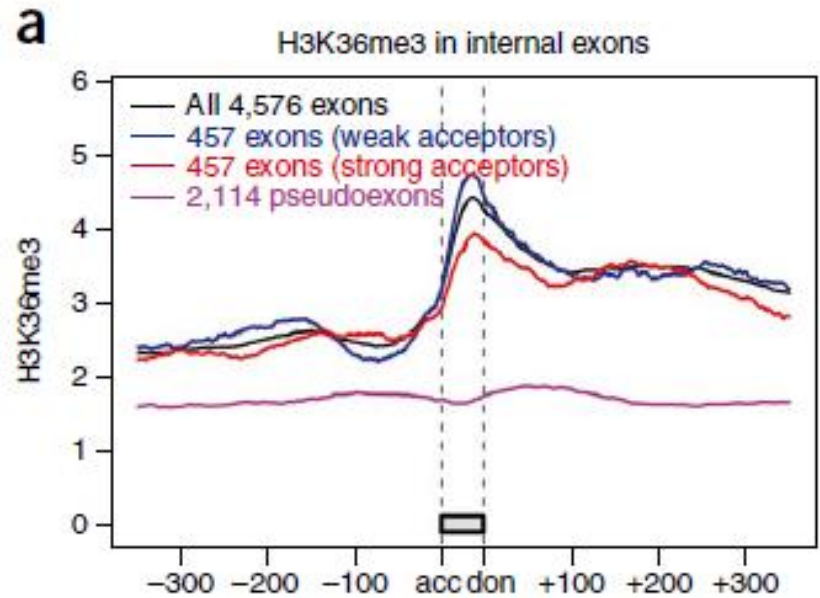
What is Bonferroni correction ?
Wiki methods

Figure 2. Exon-Biased Distribution of Specific Histone H3 Methylation Marks.
Profiles centered on exons for:
(C) monomethyl histone marks,
(D) dimethyl histone marks,
(E) Trimethylated histone marks.
(C)–(F) are normalized to average library ChIP signal across the displayed region.

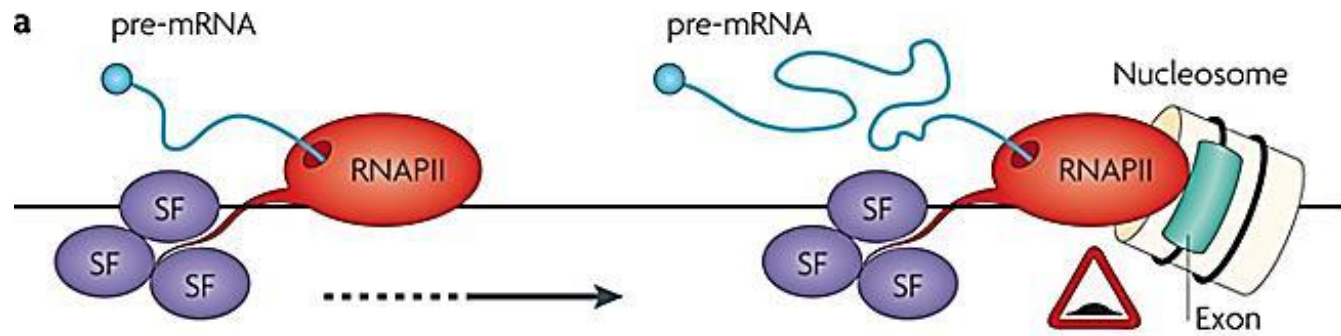H3K36me3 is one of the most evident at exons

H3K36me3 is more evident at weak exons

a   H3K36me3 in internal exons
— All 4,576 exons
— 457 exons (weak acceptors)
— 457 exons (strong acceptors)
— 2,114 pseudoexons

Data from:
Barski et al. (2007) Cell 129, 823–837.

c   H4K20me1 in internal exons
— All 4,576 exons
— 457 exons (weak acceptors)
— 457 exons (strong acceptors)
— 2,114 pseudoexons

the speed-bump model

a

pre-mRNA
RNAPII
SF
SF SF

pre-mRNA
Nucleosome
RNAPII
SF
SF SF
Exon

b

pre-mRNA
Cap
3' SS
SF Exon
RNAPII
SF SF
Nucleosome
Exon

pre-mRNA
Cap
5' SS   3' SS
Exon   SF
SF SF
RNAPII

Nature Reviews | Genetics
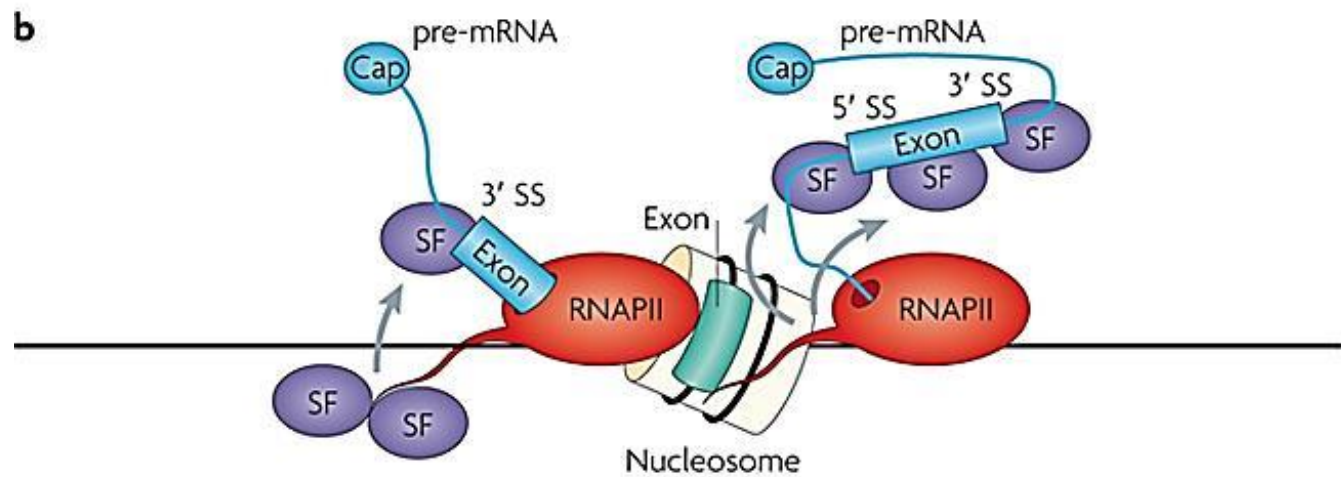
Nucleosome occupancy marks exons and is coupled to transcription.
a | RNA polymerase II (RNAPII), associated with different splicing factors (SFs), travels along the gene and transcribes it. When RNAPII reaches an area with high nucleosome occupancy and encounters specific histone modifications that mark an exon, it is slowed down.
b | This panel shows RNAPII and the nucleosome at the point at which their coupling marks the exon boundaries for the splicing machinery. RNAPII transcribes the exon and SFs detach from the carboxy-terminal domain of RNAPII and bind to the 3' splice site (3' SS) region of the precursor mRNA (pre-mRNA). During transcription elongation, additional SFs bind intronic and exonic splicing regulatory elements and the 5' SS.

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.

Differential splicing of FGFR depends on the PTB regulator, which binds to a splicing silencer around exon IIIb and represses its inclusion

FGFR2 pre-mRNA tissue-specific exon IIIb / IIIc alternative splicing was studied in **PNT2** (prostate normal epithelium) and in **hMSC** (human mesenchymal 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:
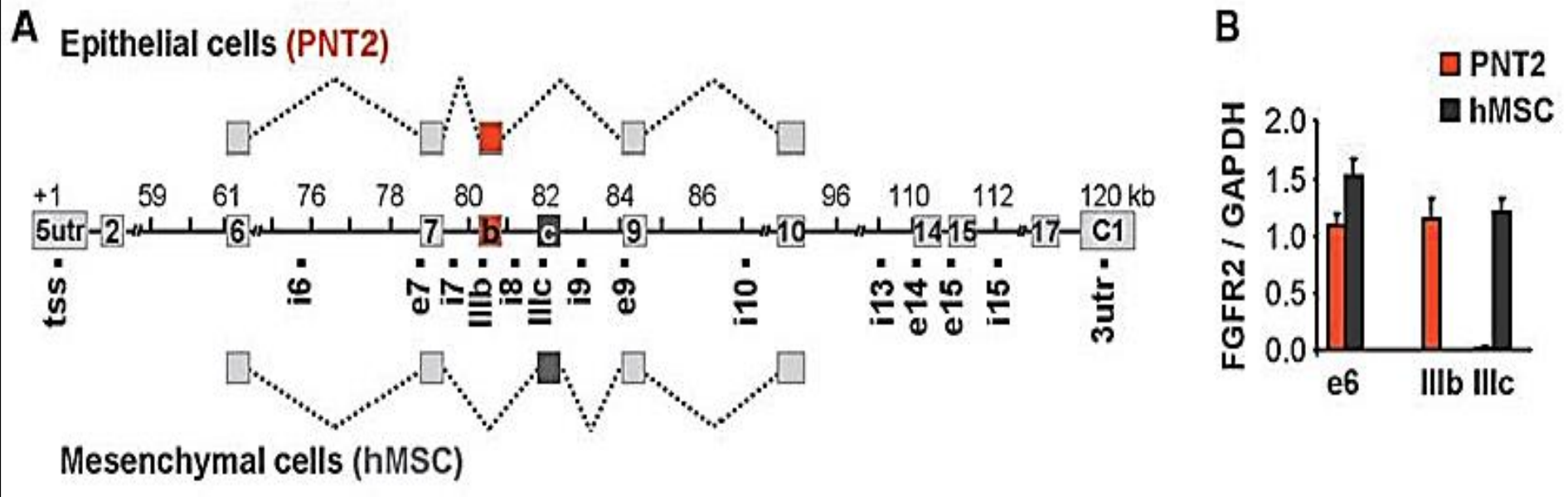
Fig. 1. Splicing-specific histone modifications. (A) Schematic representation of the human FGFR2 gene. Exon IIIb (red) is included in PNT2 epithelial cells, exon IIIc (black) is included in hMSCs. Square dots indicate oligonucleotide pairs used in analysis. (B) Levels of FGFR2 **exon inclusion** relative to GAPDH in PNT2 (red) or hMSCs (black) determined by quantitative polymerase chain reaction (PCR).

A Epithelial cells (PNT2)
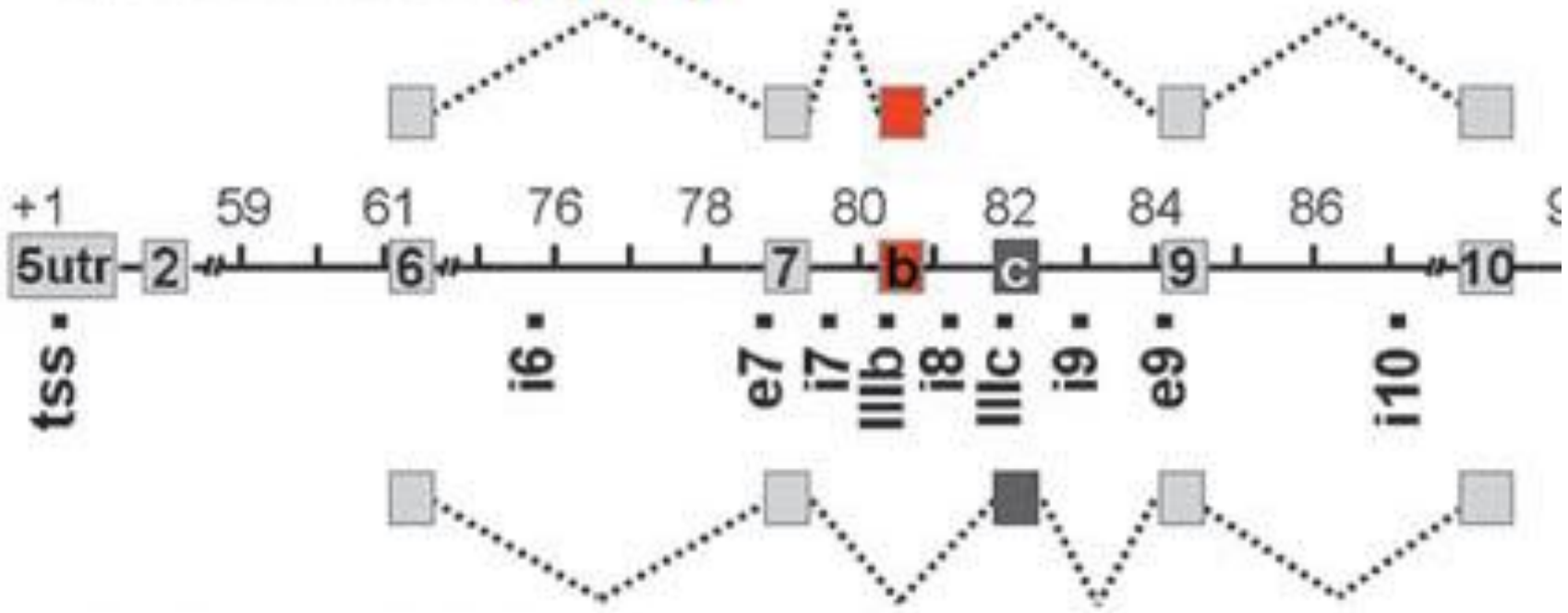
Mesenchymal cells (hMSC)

To analyze   e6
IIIb
IIIc

?

Why ratio?
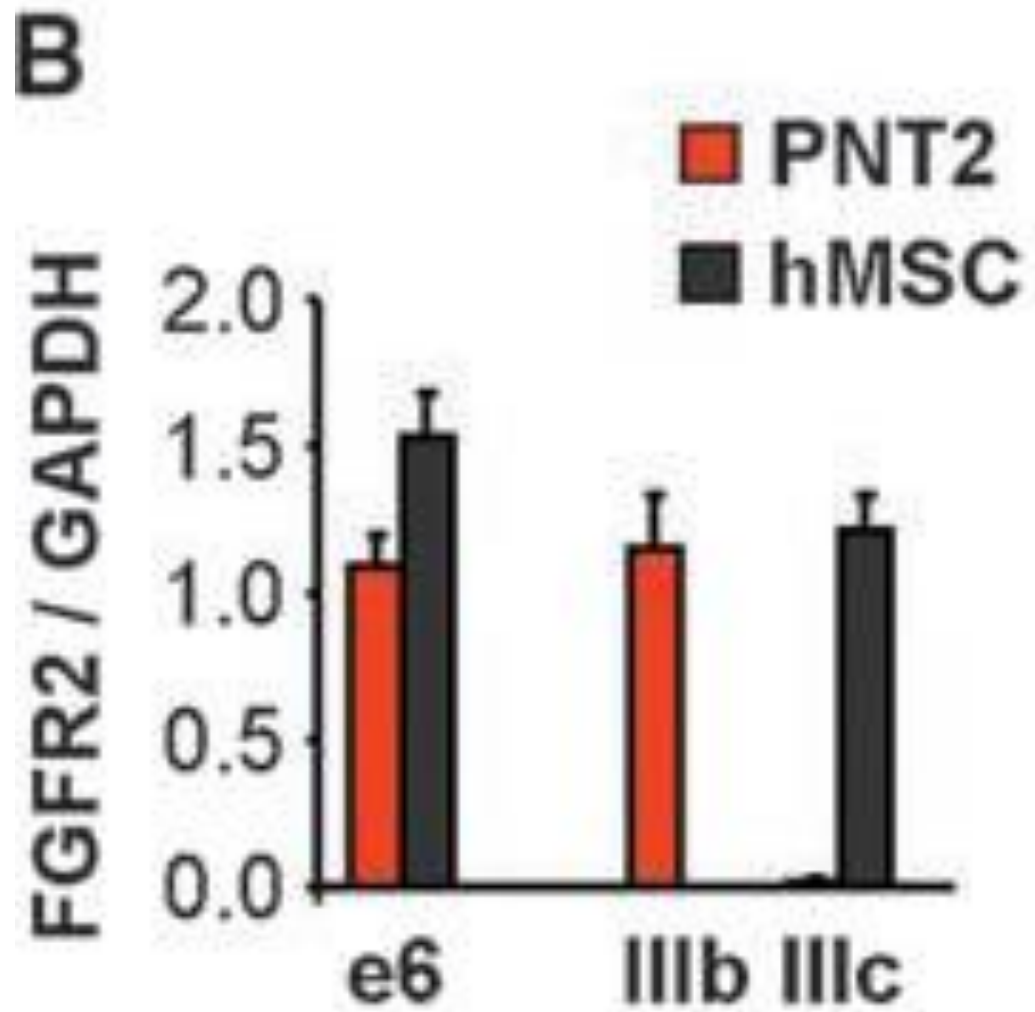
Is there any special histone PTMs at these exons ?

- Chromatin Immunoprecipitation using Antibodies against PTMs

- PCR analysis of single sites as in Figure 1
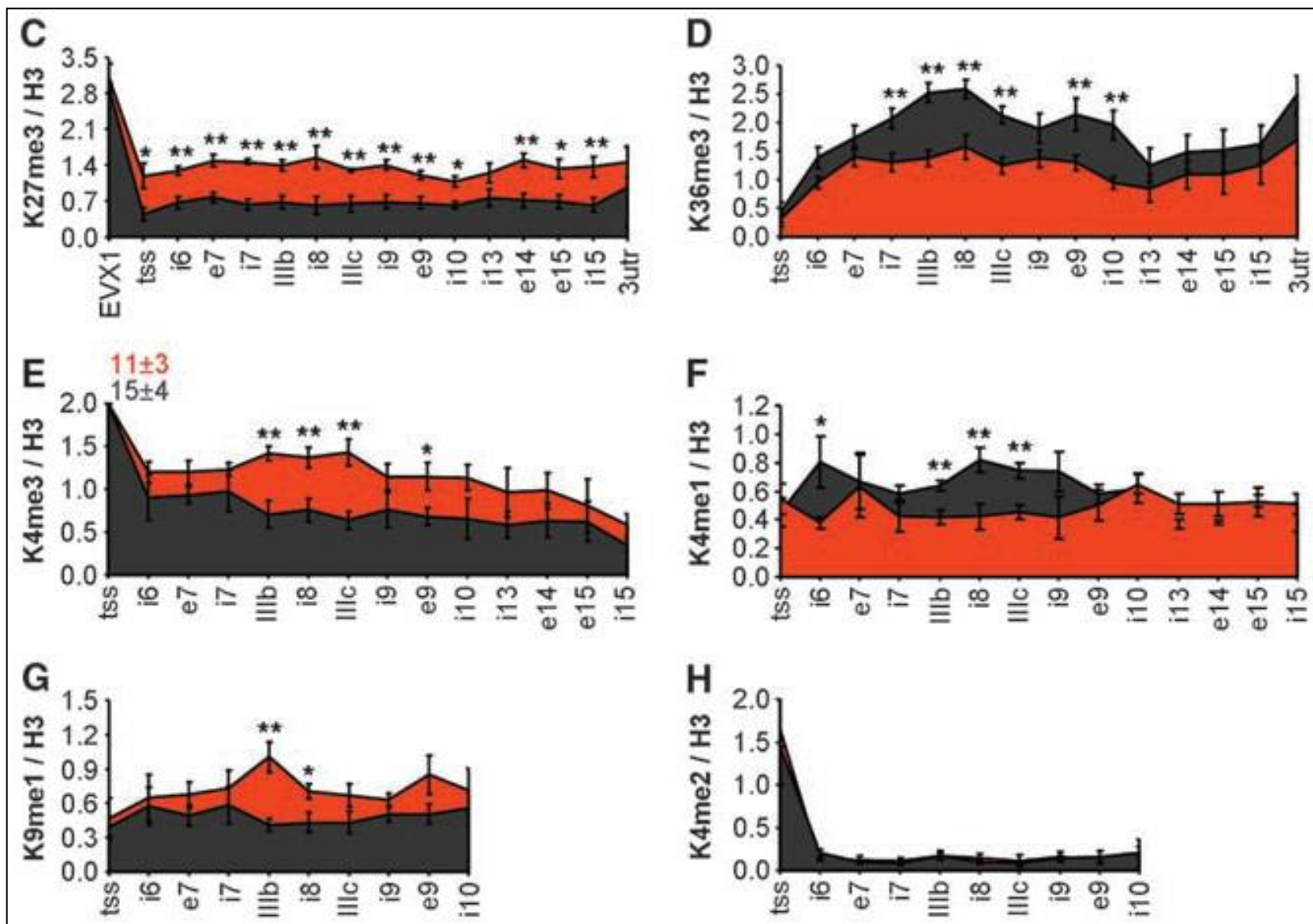
Fig. 1 - (C to H) Mapping of H3-K27me3 (C), H3-K36me3 (D), H3-K4me3 (E), H3-K4me1 (F), H3-K9me1 (G), and H3-K4me2 (H) in FGFR2 in **PNT2 (red)** and **hMSC (black)** cells by **quantitative ChIP.** The percentage of input was normalized to unmodified H3. Values represent means ± SEM from four to six independent experiments. *P <0.05, **P < 0.01, Student's t test.

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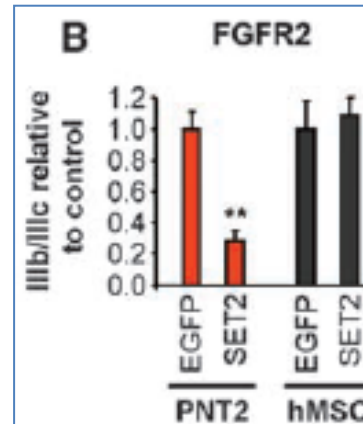
SET2 overexpression

H3K36(me3) is recognized by the
bromodomain protein **MRG15 (reader)**
When MRG15 is overexpressed in epithelial
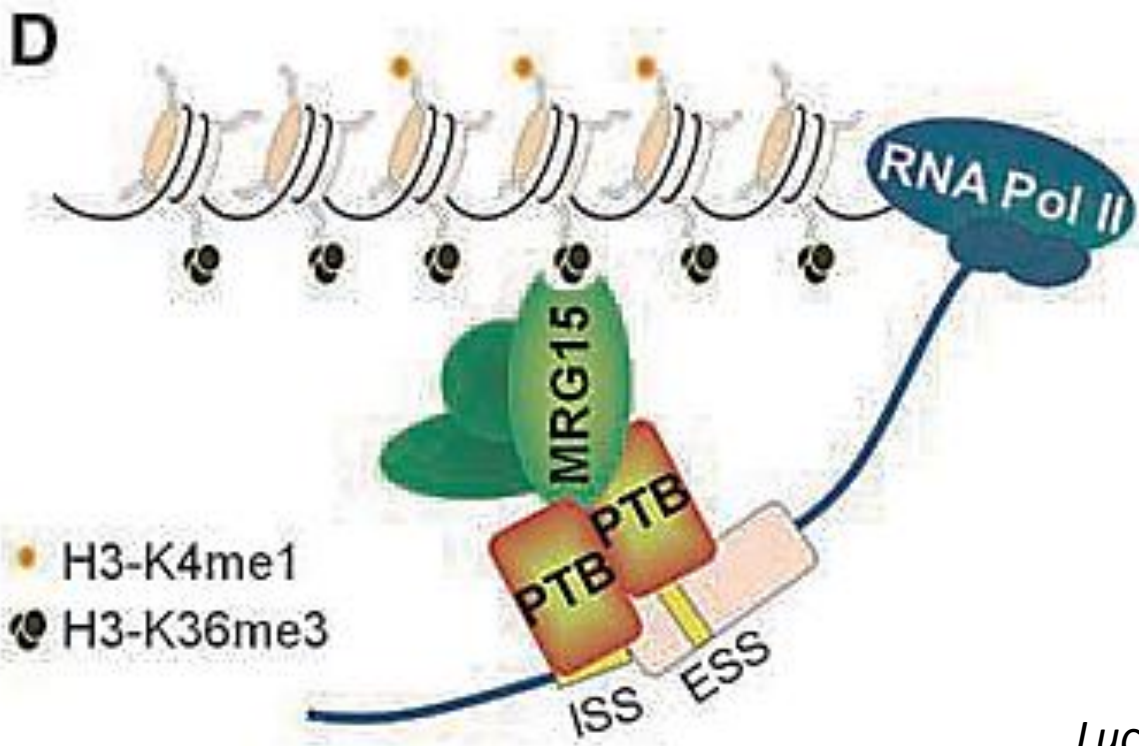cells, IIIb/IIIc raio falls by 75%:
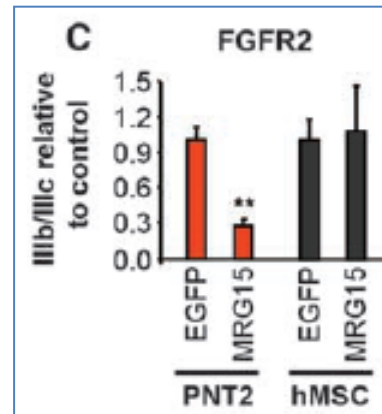
MRG15 overexpression

Exon IIIb has weak site for **PTB**.
MRG15 co-immunoprecipitates with
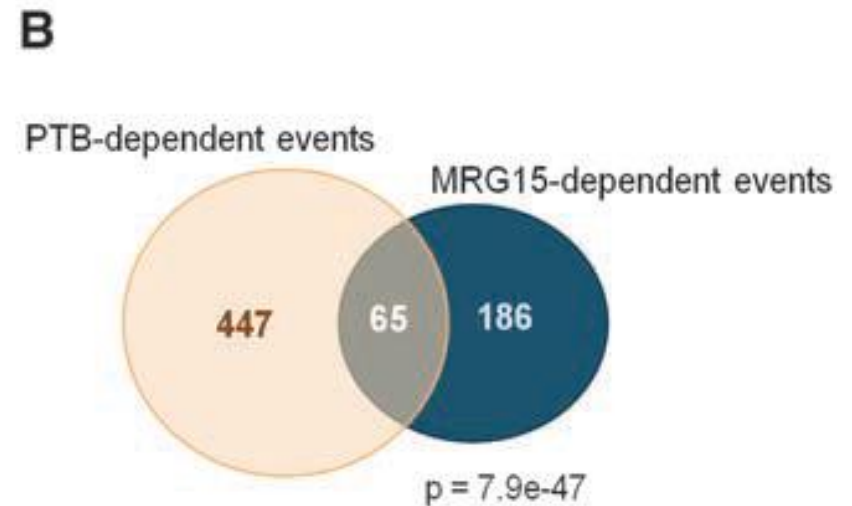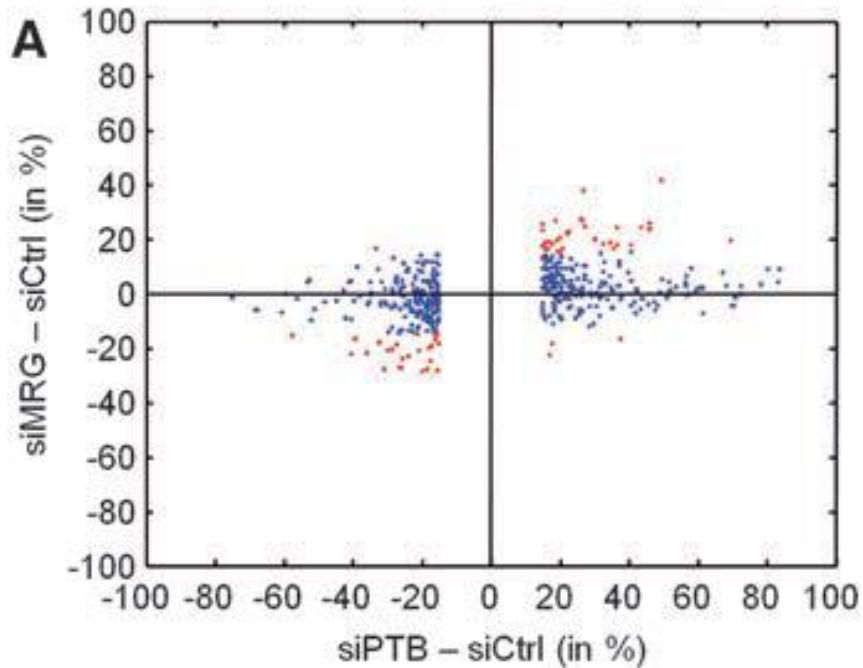the RNA binding protein PTB   →
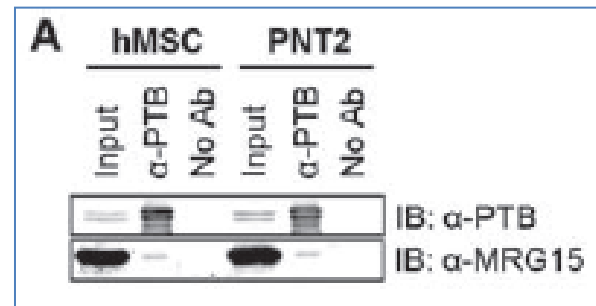
Co-IP

Luco et al, 2011

Are PTB and MRG15 effects limited to FGFR2 exons ?

siRNA-mediated down-regulation of either PTB or MRG2

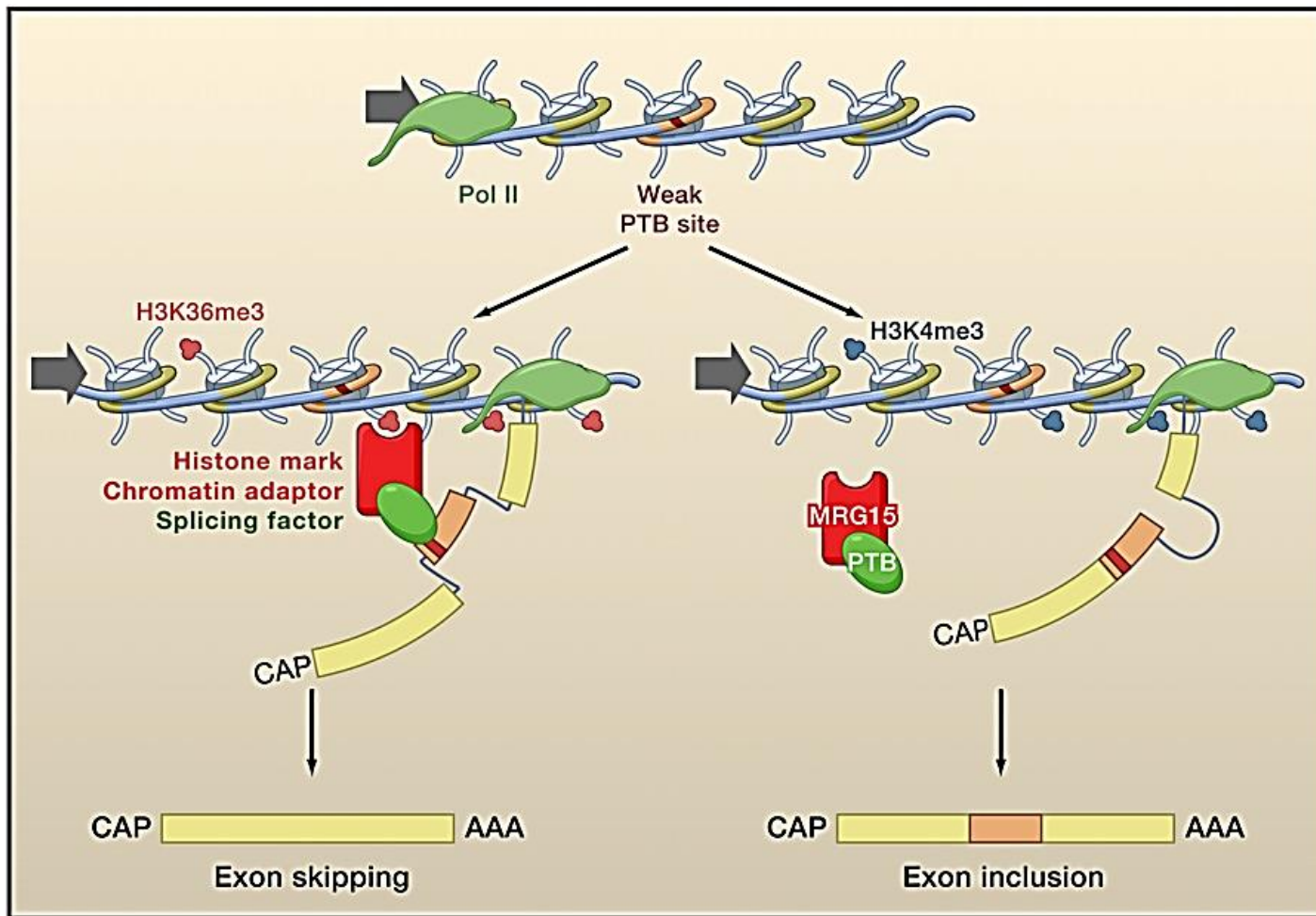RNA-seq → splicing read mapped and quantitated for each AS event

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.
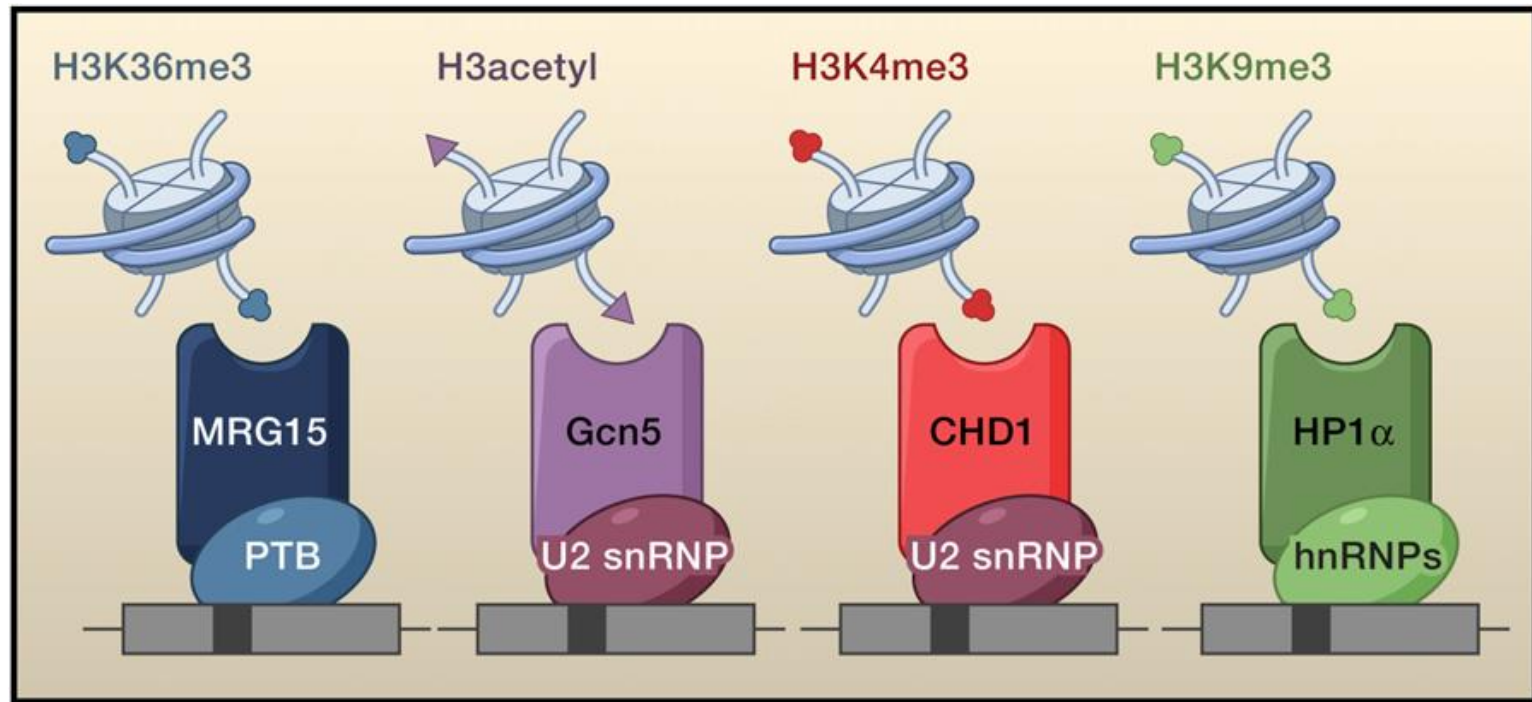

Figure 4. Chromatin-Adaptor Complexes
Several histone modification-binding chromatin proteins interact with splicing factors (Luco et al., 2010; Sims et al., 2007; Gunderson and Johnson, 2009; Piacentini et al., 2009; Loomis et al., 2009).

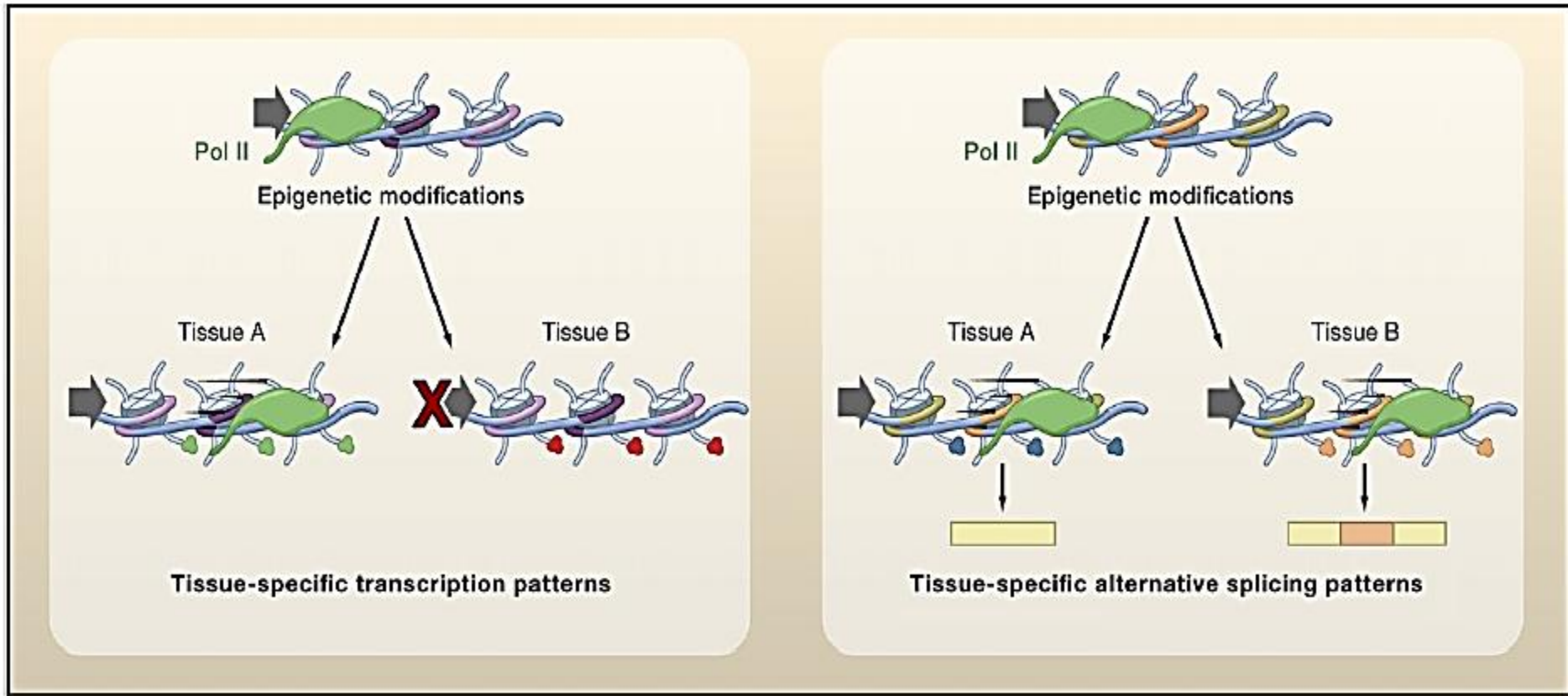A model of epigenetic "memory" of alternative splicing in the cells

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).

The speed-bump model

After years, this model is now largely accepted, also because it can explain epigenetic effects on the choice of Alternative Splicing patterns

First mechanism
Histone modification can «stabilize» or «destabilize» nucleosomes sat on specific exons, thus modifying the rate of PolII at that specific exon.

Second mechanism
Specific histone PTMs enriched at one exonic nucleosome can bind «readers» associated to Splicing factors, thus increasing the local availability of that Splicing factor, when the exon is transcribed.

Exon circularization may represent a case of alternative splicing

From Salzman 2016   Trends in Genetics
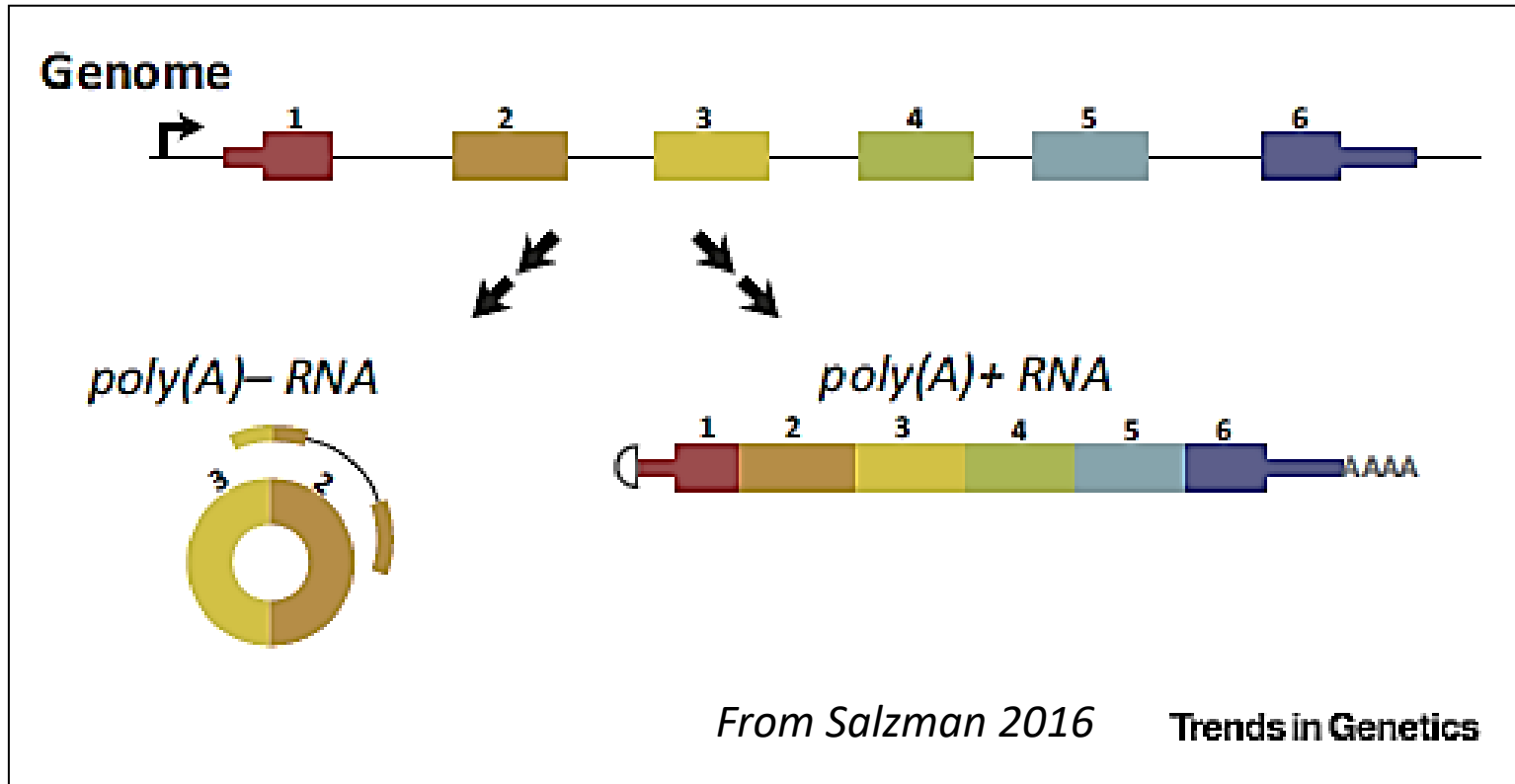
The sequence of circular RNA comprises exons from protein and noncoding loci, spliced at canonical splice sites. Bi-colored reads represent paired-end RNA-Seq reads used to discover circular RNA that are enriched in polyA-depleted RNAsamples.

circRNAs

a

Spliced read

Linear splicing

«back-splicing reads»

Circularization

circRNA
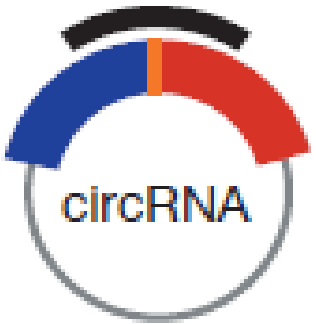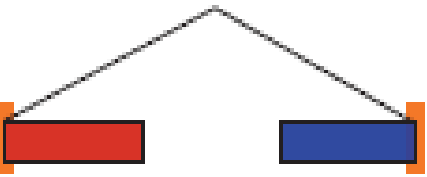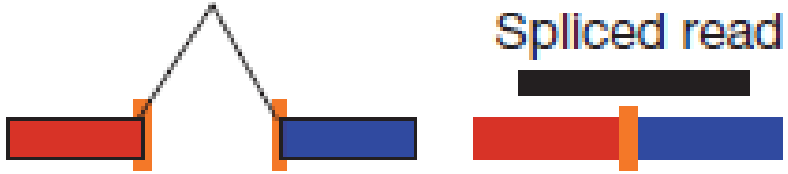
From Memczak et al. 2013

Possible functions

- Endogenous competing RNAs (ceRNA, see last lesson)

- Activators/repressors of transcription

- Alternative protein-encoding RNAs

From Chen 2016, NRMCB 17:205