



Universidad de Buenos Aires

Facultad de Ciencias Exactas y Naturales

Departamento de Fisiología, Biología Molecular y Celular

Regulación del procesamiento en transcritos con múltiples eventos de *splicing* alternativo.

Juan Pablo Fededa

Tesis presentada para optar al título de

Doctor en el área de Ciencias Biológicas

Director de tesis y Consejero de Estudios: Alberto R. Kornblihtt

Laboratorio de Fisiología y Biología Molecular, IFIBYNE, UBA-CONICET

Buenos Aires, 2007

página del Jurado

Regulación del procesamiento en transcritos con múltiples eventos de *splicing* alternativo

El *splicing* alternativo juega un rol clave en la generación de la diversidad protéica. En esta tesis se estudió la coordinación en el procesamiento de dos exones alternativos distantes en un mismo gen. Se demostró que mutaciones que inhiben o estimulan la inclusión de un exón alternativo río arriba afectan profundamente el procesamiento de otro exón alternativo ubicado río abajo. Sin embargo, mutaciones similares en el exón alternativo río abajo tienen un efecto mínimo sobre el que se encuentra río arriba. Este efecto polar es específico de promotor y se encuentra aumentado cuando se inhibe la elongación transcripcional. Consistentemente, células de ratones mutantes con inclusión constitutiva o nula del exón alternativo EDA de la fibronectina muestran coordinación con otra región de *splicing* alternativo lejana, IIICS. Usando PCR específica de alelo, se demostró que la coordinación ocurre en *cis*, y es afectada, además, por la tasa de elongación transcripcional. Mediante estudios *in silico* se encontró que el 25% de los genes humanos contiene múltiples regiones de *splicing*, varios con combinaciones de isoformas que no se ajustan a distribuciones al azar, lo que sugiere la coordinación de las mismas. La coordinación en el *splicing* alternativo, acoplada a la transcripción de la RNA polimerasa II, supone un nuevo mecanismo de regulación de la expresión génica.

Regulation of processing in transcripts with multiple alternative splicing events

Alternative splicing plays a key role in generating protein diversity. In this thesis, the coordination in processing two distant, alternatively spliced exons in the same gene was studied. Mutations that either inhibit or stimulate inclusion of the upstream alternative exon deeply affect inclusion of the downstream one. However, similar mutations at the downstream alternative exon have little effect on the upstream one. This polar effect is promoter specific and is enhanced by inhibition of transcriptional elongation. Consistently, cells from mutant mice with either constitutive or null inclusion of a fibronectin alternative exon revealed coordination with a second alternative splicing region, located far downstream. Using allele-specific RT-PCR, the coordination was shown to occur in cis and to be affected by transcriptional elongation rates. Bioinformatics supports the generality of these findings, indicating that 25% of human genes contain multiple alternative splicing regions and identifying several genes with nonrandom distribution of mRNA isoforms at two alternative regions, suggesting a coordination between them. Alternative splicing coordination, coupled to RNA pol II transcription, sets a platform for a new mechanism involved in gene expression regulation.

A Maru, con todo mi amor

Come together.

John Lennon y Paul McCartney

Publicaciones

Trabajos originales

Fededa, J.P., Petrillo, E., Gelfand, M.S., Neverov, A.D., Kadener, S., Nogues, G., Pelisch, F., Baralle, F.E., Muro, A.F., and Kornblihtt, A.R. (2005). A polar mechanism coordinates different regions of alternative *splicing* within a single gene. *Mol Cell* 19, 393-404.

Trabajos en colaboración

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003). A slow RNA polymerase II affects alternative *splicing* in vivo. *Mol Cell* 12, 525-532.

Kadener, S., Fededa, J.P., Rosbash, M., and Kornblihtt, A.R. (2002). Regulation of alternative *splicing* by a transcriptional *enhancer* through RNA pol II elongation. *Proc Natl Acad Sci U S A* 99, 8185-8190.

Revisiones

Kornblihtt, A.R., de la Mata, M., Fededa, J.P., Munoz, M.J., and Nogues, G. (2004). Multiple links between transcription and *splicing*. *RNA* 10, 1489-1498.

Nogues, G., Kadener, S., Cramer, P., de la Mata, M., Fededa, J.P., Blaustein, M., Srebrow, A., and Kornblihtt, A.R. (2003). Control of alternative pre-mRNA *splicing* by RNA Pol II elongation: faster is not always better. *IUBMB Life* 55, 235-241.

Indice

Introducción.....	11
Splicing	11
Arquitectura exón – intrón.....	11
El Spliceosoma y la reacción de splicing	13
Explosión de intrones y la emergencia del splicing alternativo.....	15
Splicing Alternativo.....	20
Proteínas SR versus hnRNPs.....	24
Transcripción.....	27
Historia de un mRNA.....	30
Splicing cotranscripcional y acoplado a la transcripción.....	31
Combinatoria: Múltiples eventos, un transcrito.....	35
Objetivos.....	39
Resultados.....	42
Parte I- Coordinación en el gen de fibronectina.....	43
Minigenes con dos regiones de splicing alternativo para estudiar la coordinación.....	43
Coordinación entre eventos de splicing alternativo en un mismo transcrito.....	44
La coordinación en el procesamiento de los eventos de splicing alternativo tiene polaridad.....	45
La coordinación polar depende del nivel de inclusión y no del pegado de la proteína SF2 al ESE del exón proximal.....	48
Dependencia del ESE en el exón distal para la coordinación polar.....	50
El grado de inclusión de los exones río arriba de la región alternativa distal no es el causante del efecto de coordinación.....	51
Rol de la definición exónica de EDI en la coordinación.....	51
El promotor alfa-globina suprime la polaridad.....	54
La coordinación es recuperada por la inhibición de la elongación de la RNA polimerasa II....	54

El efecto de coordinación es exón específico.....	58
Splicing alternativo coordinado en el gen de FN murino endógeno.....	60
Elongación de la RNA pol II y los efectos en el gen endógeno de FN.....	62
Efectos alelo específicos.....	63
Parte II-Bioinformática.....	65
Parámetros de ocurrencia del splicing alternativo.....	65
Coordinación en el genoma humano.....	67
Splicing alternativo coordinado en gen PCBP2.....	70
Conclusiones.....	73
Parte I.....	73
Parte II.....	75
Discusión.....	76
Coordinación.....	76
La trampa transcripcional.....	81
Implicancias de la coordinación.....	84
Procedimientos Experimentales.....	86
Bioinformática.....	86
Plásmidos.....	86
Reporteros de splicing alternativo con promotor de α -gb humana.....	86
Reporteros de splicing alternativo bajo el control del promoter de FN humana.....	87
Construcción pSVED α 1W/mIIICS.....	88
Mutantes de delección en la figura r7.....	88
Construcciones de splicing alternativo del gen PCBP2.....	89
Vectores de expresión de proteínas.....	90
Transformación bacteriana, minipreps, maxipreps.....	90
Cuantificación de los plásmidos.....	91
Líneas celulares.....	91
Mantenimiento de las líneas celulares.....	91
Células Hep3B.....	92
MEFs.....	92
Transfecciones.....	92
Tratamientos.....	93
Ensayo de actividad de β -galactosidasa.....	93

Análisis del RNA.....	94
Preparación de RNA.....	94
RT-PCR.....	96
Transcripción reversa (RT).....	96
PCRs del cDNA de minigenes tándem con dos eventos de <i>splicing</i> alternativo.....	97
Evento Proximal EDI.....	97
Evento Distal EDI.....	98
Evento Distal EDII (humano).....	98
Evento Distal IIICS (ratón).....	99
Ensayos de <i>splicing</i> alternativo de genes endógenos.....	99
PCBP2 endógeno (humano) y minigen.....	99
Regiones alternativas endógenas de ratón.....	99
Observación y cuantificación de productos de PCR radiactiva.....	100
Referencias.....	101
Agradecimientos.....	120

Introducción

Splicing

Desde el descubrimiento del RNA, una variedad de experimentos demostró que, en promedio, las moléculas de RNA nuclear heterogéneo (hnRNAs) presentes en los núcleos de diferentes tipos de células eucariotas eran 5 veces más largas que las observadas en los RNAs mensajeros presentes en citoplasma (mRNAs). El hecho llamaba particularmente la atención dado que ambas clases de moléculas estaban vinculadas, no sólo por su naturaleza química inherente, sino también por compartir las mismas clases de estructuras en sus extremos: se detectaba una estructura capucha (*cap*) formada por un nucleótido de guanina unido mediante un enlace 5'-5' trifosfato a sus extremos 5', mientras que en los 3' se observaba una cola de nucleótidos de adenina (poliA).

El misterio fue resuelto en 1977, cuando los grupos de Sharp y Roberts-Brooker observaron bajo el microscopio electrónico híbridos del mRNA de la proteína estructural mayor del adenovirus II con su fragmento correspondiente de DNA viral simple cadena. Descubrieron que la secuencia del extremo 5' del mRNA hibridaba en cuatro sitios diferentes del DNA, separados entre sí por grandes bucles de diferentes longitudes (Fig. i1) y postularon que estos cuatro segmentos, "unidos" en el mRNA maduro, son el producto de múltiples eventos de corte y empalme (*splicing*) en la maduración de largos mRNAs precursores que incluirían las secuencias complementarias a los bucles de DNA observados (Berget et al., 1977; Chow et al., 1977a; Chow et al., 1977b). Más tarde, Walter Gilbert llamó "exones" a las secuencias presentes en el mRNA maduro o citosólico e "intrones" a las secuencias complementarias a los bucles de DNA, que estarían presentes sólo en los precursores y luego serían removidas.

A partir de entonces, los reportes de casos de *splicing* en genes eucariotas se sucedieron rápidamente: luego de ser transcritos, los precursores eucariotas sufren sistemáticamente una modificación de la información genética que consiste en la remoción de los intrones y el empalmado de los exones flanqueantes, convirtiéndose en mRNAs maduros que luego de franquear el núcleo están listos para ser traducidos a proteínas.

El *splicing* incluye una serie de reacciones bioquímicas que son catalizadas por el *spliceosoma*, un complejo macromolecular constituido por ribonucleoproteínas pequeñas (*snRNPs*), y la información que gobierna este proceso está impresa en porciones de las secuencias nucleotídicas de los precursores.

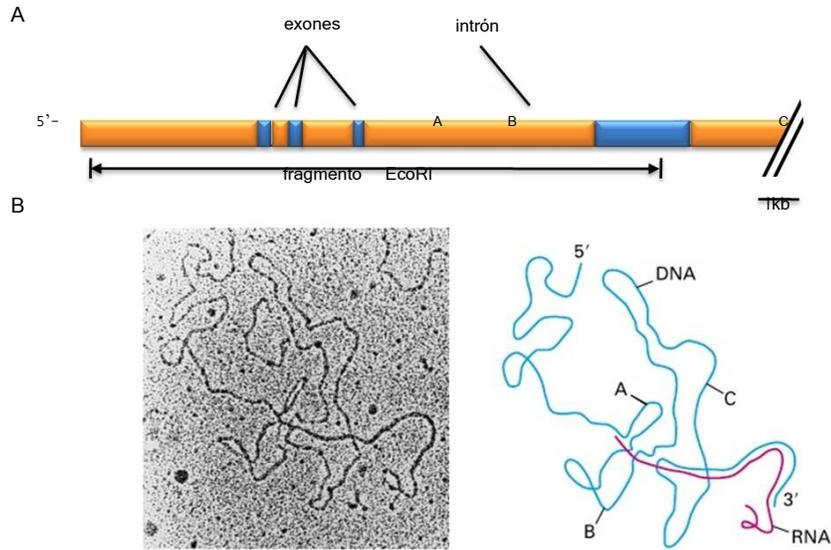


Fig.i1. (A) Estructura del gen de la proteína estructural mayor del adenovirus II y descripción del fragmento de DNA utilizado para la hibridación. (B) A, B y C *loops* de DNA viral (fragmento EcoRI) de una sola hebra sin hibridar (regiones complementarias a los intrones escindidos del mensajero maduro), que están unidos por regiones híbridas DNA-mRNA viral (tomado de Berget et al., 1977).

Arquitectura exón – intrón

La escisión de los intrones del pre-mRNA y la unión de los exones está dirigida por secuencias específicas en las juntas exón- intrón, llamadas sitios de *splicing*. El sitio 5' de *splicing* marca la junta exón-intrón en el extremo 5' del intrón. Incluye un dinucleótido GU en el extremo intrónico y abarca secuencias menos conservadas. En el otro extremo del intrón, la región del sitio 3' de *splicing* es más compleja y consta de tres elementos: un punto de ramificación, un trecho de polipirimidinas y un dinucleótido terminal AG en el extremo 3' del intrón (Fig. i2).

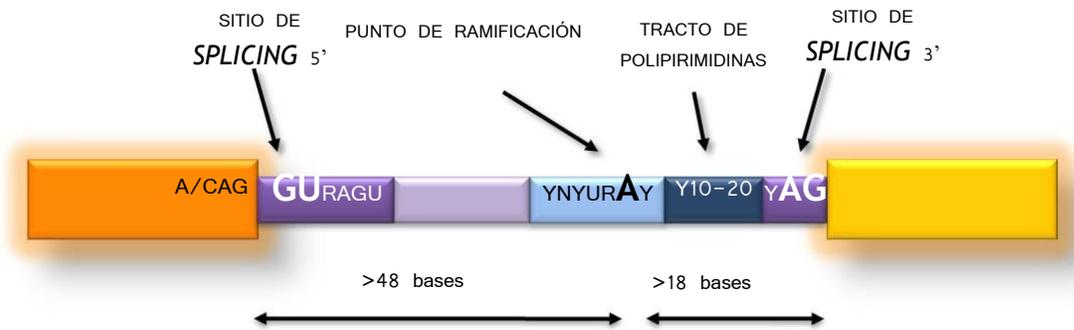


Fig.i2. Sitios de *splicing* consenso de un intrón vertebrado (Naranja y amarillo: exones, resto: intrón, Y=pirimidina, R=purina, N=cualquier nucleótido). Los dinucleótidos GU y AG en los extremos del intrón y la A en el punto de ramificación son invariables, las otras posiciones están menos conservadas. Se indican las distancias mínimas entre el sitio 5' y el punto de ramificación (*branch point*) y entre el sitio 3' de *splicing* y el sitio de ramificación (tomado de Smith and Valcarcel, 2000).

En la “definición intrónica” el reconocimiento del segmento a ser removido está regido por los sitios 5' y 3' de *splicing* en un mismo intrón. Esto es la regla en la gran mayoría de eucariotas inferiores como *S. cerevisiae*, en donde la mayoría de los intrones son cortos. Sin embargo, las secuencias consenso de *splicing* que se encuentran en el intrón muchas veces no son suficientes para activar el ensamblado del *spliceosoma* y desarrollar la reacción de *splicing*, siendo esto particularmente cierto en los vertebrados, donde los exones muchas veces son flanqueados por sitios de *splicing* “débiles” cuya secuencia se aleja del consenso y es necesaria una facilitación del proceso.

Además, los exones de vertebrados son fragmentos cortos (la mayoría tiene entre 50 y 300 nucleótidos de longitud) distribuidos entre intrones usualmente grandes, cuyo tamaño oscila entre 100 y cientos de miles de nucleótidos (Berget, 1995; Black, 1995). Dada su longitud, los intrones poseen gran número de “pseudoxones”, estructuras de 50 a 250 nucleótidos que están demarcadas por sitios de *splicing* canónicos cuya exclusión de los transcritos maduros debe ser controlada (Sun and Chasin, 2000).

El hecho es que en la mayoría de los exones internos de los vertebrados (Robberson et al., 1990) parece suceder la “definición exónica” (*exon definition*), que consiste en el reconocimiento de pares de sitios de *splicing* 5' y 3', ya no presentes en un mismo intrón, sino ubicados en las proximidades de los extremos de un exón. Estos sitios interactúan para estimular el *splicing* de los intrones que flanquean al exón (Hoffman and Grabowski, 1992) y el proceso de reclutamiento de la maquinaria de *splicing* a los

mismos es controlado por una serie de secuencias regulatorias, llamadas *splicing enhancers* si estimulan positivamente el ensamblado de *spliceosomas* o *splicing silencers* si lo inhiben.

Así, intrones y exones son reconocidos como una unidad por "definición intrónica" o "definición exónica", para intrones o exones cortos respectivamente.

Dada su posición particular, los exones en los extremos de los transcritos tienen mecanismos particulares de definición y promoción de *splicing*, los cuales están asociados a los procesos de *capping* en el extremo 5' y de corte y poliadenilación en el 3'.

Las secuencias *enhancer* presentes en los exones (*ESEs=Exonic Splicing Enhancers*) promueven el reconocimiento de los mismos y por ende el *splicing* de los intrones lindantes. El análisis mutacional y la bioinformática permitieron definir, además, elementos supresores en intrones y exones (*ISSs or ESSs= Intronic or Exonic Splicing silencers*) e inclusive *enhancers* intrónicos (*ISEs= intronic splicing enhancers*).

Los *ESEs* funcionan casi exclusivamente como plataforma de unión de proteínas con repeticiones de arginina-serina (SR), las cuales regulan el *splicing* mediante el reclutamiento del resto de los componentes del *spliceosoma* y la promoción de la interacción entre los mismos (Fig.i3) (Black, 2003).

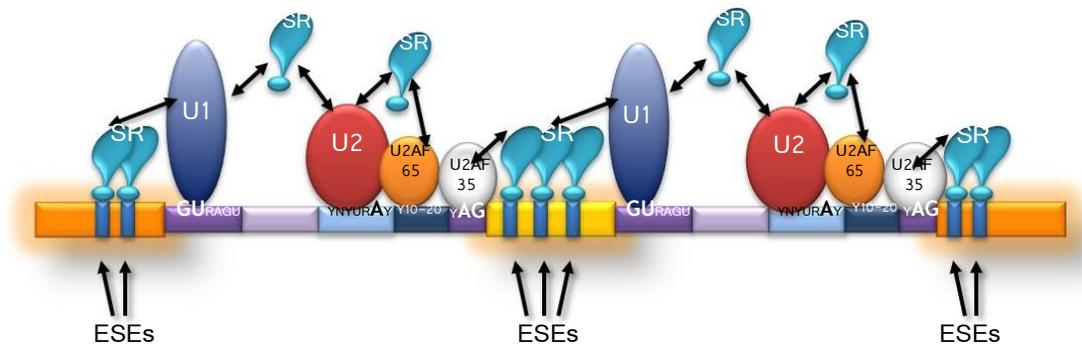


Fig.i3. Los sitios de *splicing* 5'y 3' son reconocidos por su proximidad a los exones. Éstos contienen *splicing enhancers* (*ESEs=exonic splicing enhancers*) que son sitios de unión para proteínas SR. Las mismas actúan generando complejos-puente de reconocimiento y "definición exónica", reclutando a la ribonucleoproteína U1 al sitio 5'de *splicing* río abajo y a la proteína U2AF al sitio 3' de *splicing* río arriba. A su vez funcionan en la generación de la identidad intrónica, facilitando las interacciones entre U1 y U2 (adaptado de Maniatis and Tasic, 2002 y de Smith and Valcarcel, 2000).

El *spliceosoma* y la reacción de *splicing*

El *spliceosoma* y su relación con el *splicing* fueron originalmente postulados por Lerner y Steitz, al caracterizar los inmunoprecipitados de un extracto nuclear con anticuerpos de pacientes que desarrollaron la enfermedad autoinmune *lupus erythematosus*.

Al analizar los inmunoprecipitados, descubrieron que están formados por 7 polipéptidos principales de 12 a 35 kDa y 5 RNAs nucleares pequeños (snRNAs): U1, U2, U4, U5 y U6, entre los cuales identifican que las secuencias cercanas al extremo 5' del más abundante, U1, son complementarias a las secuencias 5' de los sitios de *splicing* (Lerner and Steitz, 1979).

El *spliceosoma* es el complejo en el cual se combinan las ribonucleoproteínas nucleares pequeñas (*snRNPs=snRNAs* + proteínas asociadas), la molécula de precursor o pre-mRNA y otras proteínas no-*snRNPs*, para catalizar la reacción de *cis-splicing* en el núcleo. Análisis recientes en el área de la proteómica, muestran que estas partículas consisten en 5 *snRNAs* y alrededor de 150 proteínas asociadas (Zhou et al., 2002), convirtiéndose en la maquinaria celular más compleja caracterizada hasta el momento.

La reacción de *splicing* está dividida en varios pasos de ensamblado y catálisis, y diferentes subgrupos o complejos de proteínas del *spliceosoma* están asociados específicamente a cada uno de estos. La identificación de estos pasos fue llevada a cabo mediante ensayos de *splicing in vitro* sobre precursores radioactivos presintetizados y expuestos a extractos nucleares de células HeLa en presencia de ATP (Padgett et al., 1986). Esto, junto con las técnicas de fraccionamiento de extractos celulares (Brody and Abelson, 1985) e inmunodepleción específica de snRNPs, permitió dilucidar el funcionamiento del *spliceosoma*.

En el pre-ensamblado, el snRNP U1 se une al sitio 5' de *splicing* por complementariedad de bases y cada elemento del sitio 3' de *splicing* se une a una proteína específica: el punto de ramificación (*branch point*) a SF1, el trecho de polipirimidinas a U2AF65 (factor auxiliar de U2) y el dinucleótido AG del extremo 3' del intrón a U2AF35, formando el complejo E o CC (*Early=temprano, Commitment Complex* = complejo de compromiso) (Jamison and Garcia-Blanco, 1992; Seraphin and Rosbash, 1989). Este complejo parece contener, además de algunas proteínas SRs, al snRNP U2 aún no unido al punto de ramificación, no es dependiente de ATP y "compromete" al pre-mRNA a sufrir *splicing*, siendo considerado el punto limitante en toda la reacción (Fig.i4)

Luego, en el ensamblado del *spliceosoma*, la proteína U2AF65 previamente unida al trecho de polipirimidinas, asiste al snRNP U2 a unirse al punto de ramificación por apareamiento de pares de bases RNA/RNA, y se supone que éste activa la adenosina que va a formar la estructura de lazo o *lariat*, de

manera que sea más “reactiva” para atacar nucleofílicamente al sitio 5' de *splicing* (Krainer and Maniatis, 1985; Newby and Greenbaum, 2002). En la formación de este complejo, llamado A, se hidroliza ATP (Grabowski and Sharp, 1986; Konarska and Sharp, 1986; Query et al., 1994).

El tri-snrNP U4-U5-U6 se une a la reacción, aparentemente mediante la interacción entre U2 y U6, formando el complejo B de *splicing* temprano (Staley and Guthrie, 1998) en donde los componentes proteicos y el *loop* invariante de U5 interaccionan con los sitios 5' y 3' de *splicing* (Chiara et al., 1997; Teigelkamp et al., 1995).

Este último es cortado y de forma concertada se une el extremo 5' libre del intrón, con el hidroxilo 2' de la adenosina del punto de ramificación.

El 3'-hidroxilo del intermediario exónico 5' ataca al fosfato en el extremo 3' del intrón en el intermediario que contiene al *lariat* (Fig.i5).

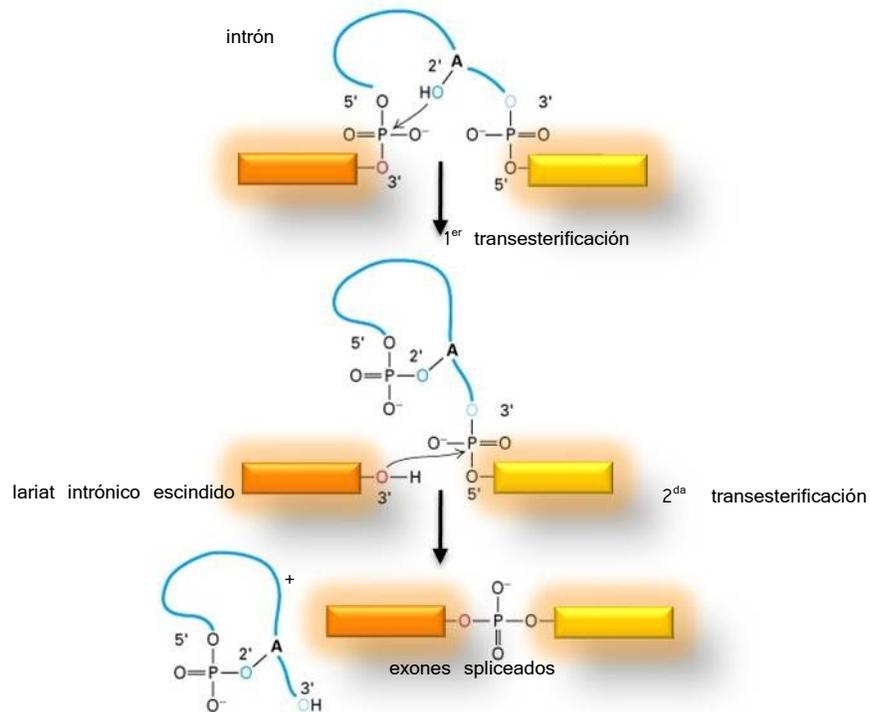


Fig.i5. El *splicing* ocurre mediante dos reacciones de transesterificación. Las flechas indican la reacción del oxígeno de los hidroxilos con los átomos de fósforo (tomado de Lodish, 2004). ○: oxígeno 3' del exón 1, ○: oxígeno 2' de la A del punto de ramificación, ○: oxígeno 3' del intrón.

El sitio 3' de *splicing* es entonces cortado y la ligación de los exones se consigue mediante un segundo paso de transesterificación, consumiendo una tercer molécula de ATP.

Finalmente, el *lariat* se suelta unido a snRNPs.

La similitud en la formación de la ramificación entre los intrones autocatalíticos tipo II presentes en transcriptos mitocondriales de *S. cerevisiae* y los que sufren *splicing* mediado por *spliceosoma*, sugiere que el *splicing* de estos últimos es también una reacción catalizada por moléculas de RNA.

Experimentos recientes en donde la reacción de ramificación sobre un oligo que emula el sitio de ramificación es catalizada por un complejo formado sólo por los snRNAs U2 y U6 parecen indicar que la especulación es correcta (Valadkhan and Manley, 2001, 2003) (Fig.i6).

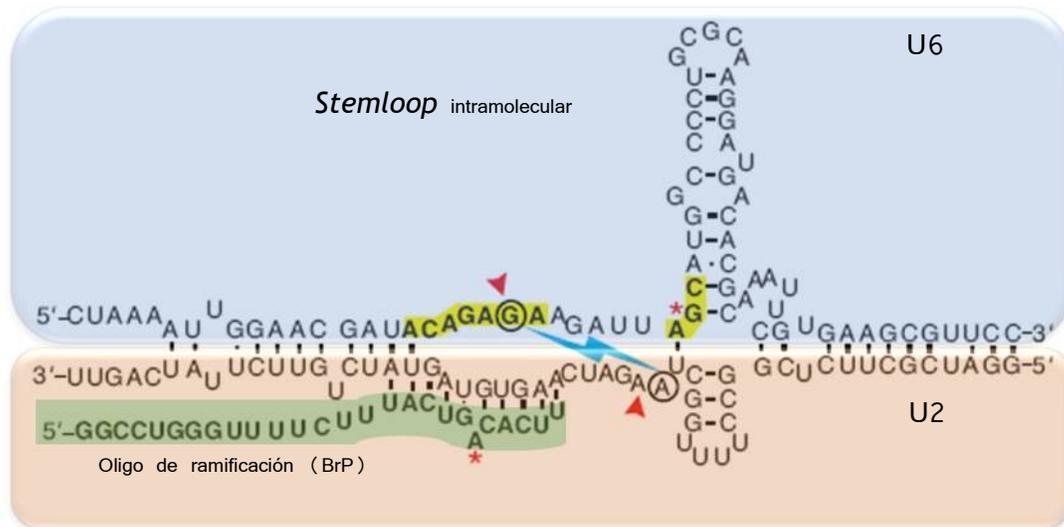


Fig.i6. Sitio activo del *spliceosoma*. Esquema del complejo U2/U6 ensamblado *in vitro* por Valadkhan et. al 2003, libre de proteínas. Azul: U6 snRNA, rojo: U2 snRNA, verde: oligo que emula el punto de ramificación. Las cajas amarillas marcan las regiones conservadas de U6 snRNA. Los apareamientos de bases previamente establecidos están indicados. Los residuos marcados con un rayo, pueden ser *cross-linkados* por luz UV. Las flechas marcan los residuos envueltos en interacciones probadas genéticamente. Los asteriscos denotan residuos involucrados en uniones covalentes entre el oligo BrP y U6 en la reacción de *splicing* catalizada por U2 y U6.

Existe una segunda clase de *spliceosoma*, el cual en vez del snRNP U2 contiene al snRNP U12, que procesa el 0,1% de los eventos de *splicing* (Hall and Padgett, 1994). Esta pequeña familia de intrones posee secuencias consenso diferentes a las discutidas (Nilsen, 1996).

Explosión de intrones y la emergencia del *splicing* alternativo

Una de las características que define a los eucariotas, a diferencia de los reinos Archea y Bacteria, es que los genes codificantes para proteínas están interrumpidos por intrones que sufren *splicing* mediado por el *spliceosoma*: hasta la fecha, todos los eucariotas cuyo genoma fue secuenciado contienen intrones y un complemento completo o cuasi completo de la maquinaria para llevar a cabo el *splicing* (Collins and Penny, 2005).

Dentro del reino eucaria, la genómica comparada nos muestra un amplio rango de densidades intrónicas: desde 140.000 intrones en los humanos (8.4 intrones por gen) a sólo 15 en el microsporidio *E. cuniculi* (0.0075 intrones por gen) (Mourier and Jeffares, 2003).

¿Cómo aparecieron, penetraron y se propagaron los intrones en los genomas? se conoce bastante poco acerca de esto, pero existe consenso en que los intrones y la maquinaria molecular del *splicing* estaban ya presentes en los estadios tempranos de la eucariogénesis y, quizás, su auge esté ligado a la emergencia del núcleo (Martin and Koonin, 2006).

Los últimos trabajos, en donde se utilizan bases de datos genómicos recientemente expandidas por los proyectos genomas de diferentes especies, presentan evidencias sobre la evolución de los mismos y sugieren que la alta densidad intrónica inicial en eukarya fue seguida por gran cantidad de ganancia y pérdida de intrones en la evolución subsiguiente (Carmel et al., 2007) . Muchos clados secuenciados presentaron pérdidas posiblemente asociadas a procesos de transcripción reversa de mRNAs maduros, seguidos de recombinación homóloga de los cDNAs con la copia genómica del gen (Derr and Strathern, 1993; Jeffares et al., 2006). Algunos también presentaron ganancias en el número de intrones. Así, el camino evolutivo por el cual emergen los opistocontes, se generan los metazoos y aparecen los primeros deuterostomados, parece estar signado por una explosión en el número de intrones (Fig. i7).

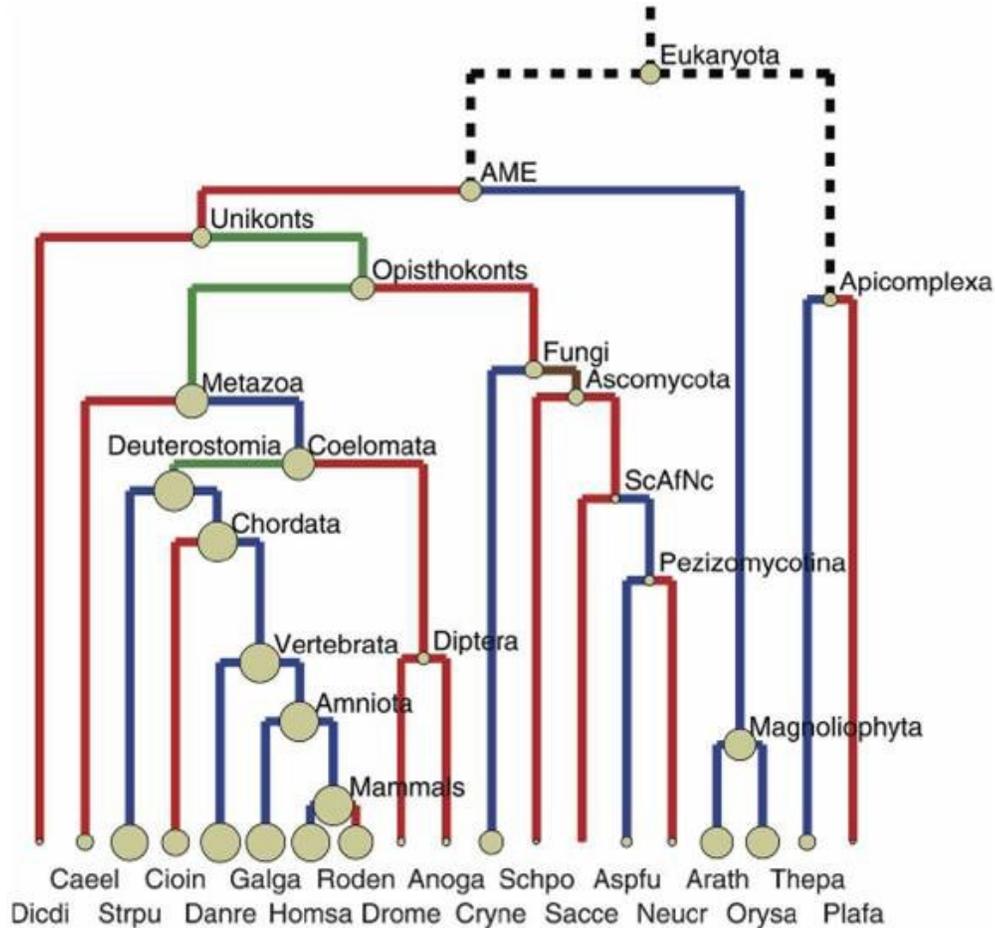


Fig.i7. Distribución de la ganancia y pérdida de intrones en el árbol filogenético de los eucariotas cuyos genomas fueron secuenciados. La medida de los nodos (círculos grises) es proporcional a la densidad intrónica conocida o calculada. Las ramas muestran un código de colores: verde=ganancia de intrones, rojo=pérdida de intrones, azul= pérdida y ganancia de intrones balanceada, marrón= pérdida y ganancia extensiva de intrones, abreviación de especies y linajes: (Caeel) *Caenorhabditis elegans*, (Strpu) *Strongylocentrotus purpuratus*, (Cioin) *Ciona intestinalis*, (Danre) *Danio rerio*, (Galga) *Gallus gallus*, (Homsa) *Homo sapiens*, (Roden) *Mus musculus* and *Rattus norvegicus* combined, (Drome) *Drosophila melanogaster*, (Anoga) *Anopheles gambiae*, (Cryne) *Cryptococcus neoformans*, (Schpo) *Schizosaccharomyces pombe*, (Sacce) *Saccharomyces cerevisiae*, (Aspfu) *Aspergillus fumigatus*, (Neucr) *Neurospora crassa*, (Arath) *Arabidopsis thaliana*, (Orysa) *Oryza sativa*, (Thepa) *Theileria parva*, (Plafa) *Plasmodium falciparum*, (Dicdi) *Dictyostelium discoideum*, (AME) Ancestro de los eucariotes multicelulares (tomado de Carmel et al., 2007).

La mayoría de la evidencia indica que probablemente el barajado de exones (*exon shuffling*), por el cual diferentes exones “iniciales” conteniendo la información de módulos funcionales conservados se multiplican, “saltan” e irrumpen el repertorio intrónico generando mayor riqueza intrónica y proteica, esté más asociado a eventos anteriores a la aparición de los eucariotas (Gilbert, 1978) y sea el determinante del esplendor del *pool* de genes que moldearon los principios de la vida como la conocemos (Roy and Gilbert, 2006).

El terreno de este fenómeno más tardío de generación de diversidad intrónica fue probablemente abonado, o por lo menos estabilizado, por la asociación de los intrones a una basta red de regulación de la expresión génica, típica de los organismos radiados a partir de esta rama evolutiva:

- a) una variedad de RNAs no traducidos, incluyendo microRNAs (miRNAs), RNAs pequeños nucleolares (snoRNAs) y RNAs guías para el proceso de *editing* están codificados por intrones (Mattick and Gagen, 2001).
- b) los intrones proveen la posibilidad de generar nuevos productos genéticos por *splicing* alternativo, el cual tiene un rol preponderante en los animales, principalmente los deuterostomados.

Las evidencias sobre la presencia del *splicing* alternativo en ciertos linajes son reveladoras: el 60% de los genes humanos y murinos tiene al menos un evento de *splicing* alternativo (Fig.i8) (Kim et al., 2007).

Al comparar la complejidad de los organismos en la escala evolutiva con el número de genes presentes en cada organismo, se observa que no existe relación directa entre ambas variables. Luego de la vieja paradoja del valor C, nos enfrentamos a la nueva paradoja del valor N: no sólo la cantidad de DNA (C) no correlaciona directamente con nuestra percepción intuitiva de la complejidad de los organismos, sino que el número de genes (N) tampoco (Claverie, 2001).

Más allá del hecho de que la sofisticación biológica puede estar relacionada con la evolución de mecanismos de expresión génica finamente regulada más que por el incremento en el número de genes (lo cual es cierto para los promotores de los metazoos,

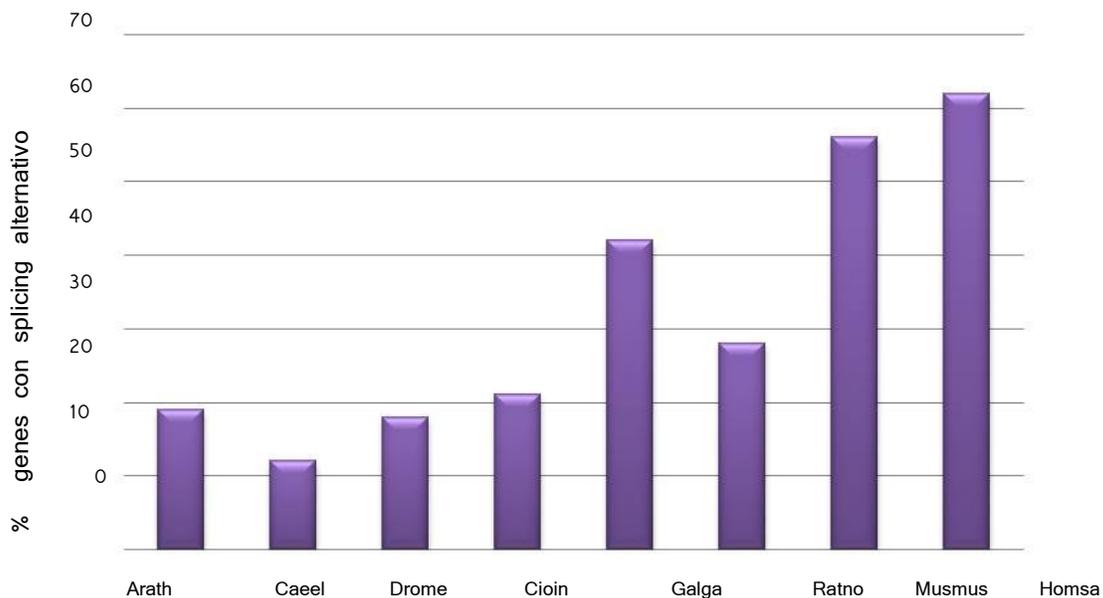


Fig.i8. Análisis a partir de ESTs del nivel de *splicing* alternativo en diferentes organismos eucariotas, utilizando la base de datos UniGene. (Caeel) *Caenorhabditis elegans*, (Cioin) *Ciona intestinalis*, (Galga) *Gallus gallus*, (Homsa) *Homo sapiens*, (Musmus) *Mus musculus*, (Ratno) *Rattus norvegicus*, (Drome) *Drosophila melanogaster*, (Arath) *Arabidopsis thalian* (tomado de Kim et al., 2007).

en donde existe un uso combinatorio de una plétora expandida de factores de transcripción con respecto a los protozoos (Fickett and Wasserman, 2000)), lo cierto es que el número calculado de proteínas en *H. sapiens* por lo menos cuadruplica al de genes (90.000 proteínas versus 20.000–25.000 genes (Harrison et al., 2002)).

Los mecanismos por los cuales se genera esta diversidad proteica incluyen las múltiples capas de eventos de modificación de los mRNAs: iniciación alternativa de la transcripción (Quelle et al., 1995), edición, corte y poliadenilación alternativos y, mayoritariamente, *splicing* alternativo, por el cual se combinan alternativamente los exones en diversos patrones que producen múltiples isoformas a partir de un único precursor, lo que en mamíferos se convierte en la regla más que la excepción.

Splicing alternativo

La posibilidad de que existan opciones en la elección de los sitios de *splicing* a ser procesados en un precursor tiene profundos efectos en los mRNAs y las proteínas de los genes.

Así, por ejemplo, los mRNAs de la proteína CSF-1 tienen menor estabilidad si se incluye el exón alternativo 10 que posee un elemento de desestabilización en la región 3' no traducida (Bonafe et al., 2005), mientras que la eficiencia traduccional del mRNA de la HMG-CoA sintetasa es afectada por el *splicing* alternativo del exón 2 en la región 5' no traducida (Gil et al., 1987).

Usualmente, la inclusión o exclusión de un exón alternativo en el mRNA altera la secuencia peptídica de las proteínas, lo cual tiene efectos en su estructura y función (Grabowski and Black, 2001): existen reportes de cambios en la localización celular (proteína de adhesión N-CAM), modificación de la actividad (factor de transcripción NF- κ B, receptor FGFR-2), y generación de actividades completamente diferentes (sistema calcitonina-CGRP) reguladas por eventos de *splicing* alternativo.

Las decisiones que conciernen a la elección de los sitios a ser utilizados están fuertemente reguladas dinámicamente en los diferentes tipos celulares, estadios del desarrollo y en respuesta a diferentes estímulos externos.

El rol fisiológico del *splicing* alternativo es amplio y su importancia biológica está ilustrada en los siguientes ejemplos:

- La determinación del sexo en *Drosophila* consiste en una cascada de eventos de *splicing* alternativo en al menos 5 genes (*Sex-lethal (Sxl)*, *Male-specific lethal (Msl2)*, *DoubleSex (dsx)*,

Fruitless(fru) y *transformer (tra)*) que culmina en la expresión diferencial de rasgos físicos y comportamentales (Black, 2003).

- En nuestros oídos, en la membrana basilar de la cóclea, existe un arreglo de neuronas ciliadas que responden a diferentes frecuencias. Cada célula individual responde a una frecuencia determinada diferente y uno de los componentes principales de esta “sintonía fina” de respuesta es el canal de potasio *slo*, el cual tiene 8 regiones de *splicing* alternativo. Las múltiples isoformas del canal de potasio *slo* difieren en su sensibilidad al calcio y en su cinética de desactivación. Se demostró que cada célula tiene una mezcla única de variantes de *slo* que se relaciona con la frecuencia a la cual responde (Black, 1998).

Existen diferentes maneras en las cuales un evento de *splicing* puede ser alternativo, entre las que la retención de intrones es la menos frecuente en mamíferos (Fig.i9).

Los factores que afectan y hacen posibles las diferentes formas de *splicing* alternativo son elementos en *cis* (secuencias en la molécula del precursor) y factores *trans* (proteínas reguladoras) que usualmente se unen a los primeros.

La eficiencia con la cual el *spliceosoma* reconoce y selecciona sitios de *splicing* presentes en los extremos de un exón depende del balance entre estos factores *cis* y *trans*. Entre los primeros se incluye la longitud del exón, la presencia de secuencias auxiliares (*enhancers* o *silencers*) y la fuerza misma de los sitios de *splicing*, que disminuye cuando sus secuencias se alejan del consenso y en el caso del sitio de *splicing* 3' hasta puede presentar anomalías como *switches* de posición entre el punto de ramificación y el trazo de polipirimidinas.

Como fue expuesto anteriormente, los *splicing enhancers* regulan positivamente y los *splicing silencers* negativamente la inclusión de un exón alternativo.

Estos elementos auxiliares son pequeños, con secuencia variable y usualmente de una sola hebra, aunque, en algunos, su función depende de su estructura secundaria (Blanchette and Chabot, 1997; Muro et al., 1999). Pueden estar dentro, usualmente próximos (menos de 100 nucleótidos), o distantes (más de 1000 nucleótidos) del exón que regulan. Si son intrónicos pueden estar río arriba o río abajo del exón a regular. Fueron descubiertos usando análisis delecional o mutaciones sitio específicas de secuencias, en minigenes conteniendo regiones alternativas, en transfecciones transitorias (Cooper, 1999), por comparación de secuencias exónicas mediante bioinformática y mediante selección de secuencias que activan el *splicing* (SELEX) (Blencowe, 2000; Fairbrother et al., 2002; Hertel et al., 1997). En sistemas naturales, la primera descripción correspondió al ESE del exón alternativo EDI de la fibronectina humana (Mardon et al., 1987).

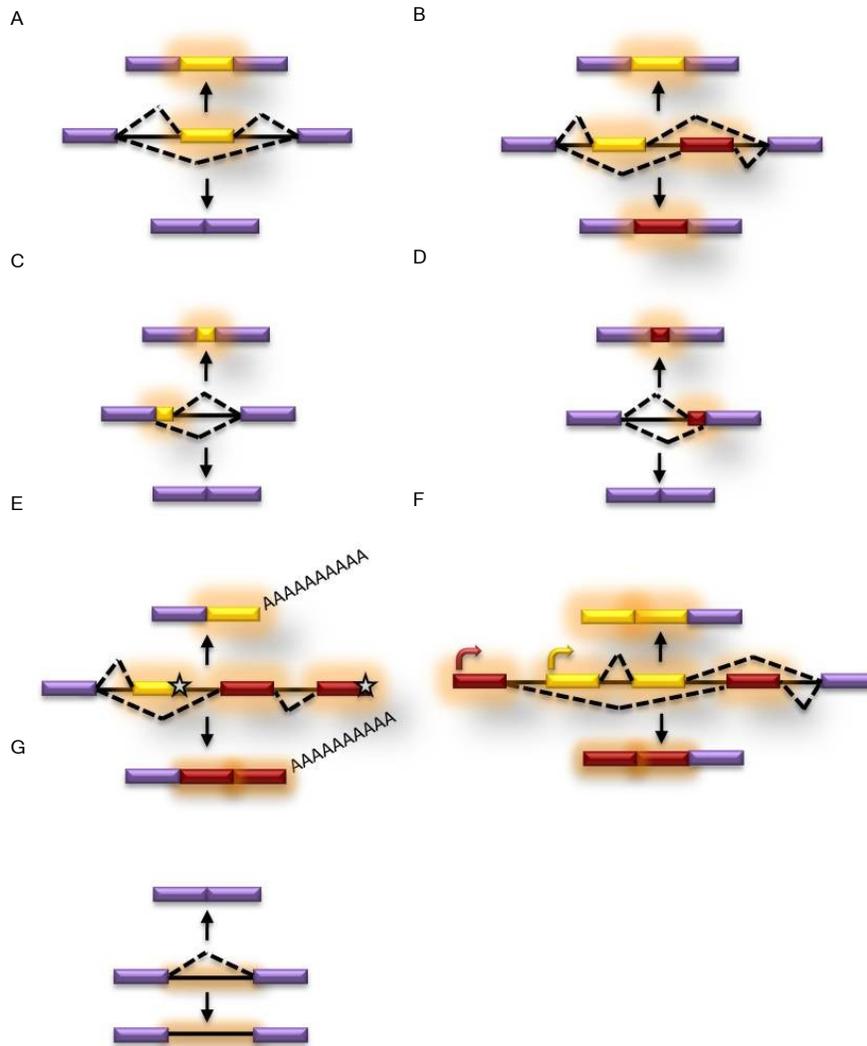


Fig.i9. Ejemplos de diferentes tipos de *splicing* alternativo. A) la inclusión alternativa de un *cassette* exónico N1 en el proto oncogen c-src., B) exones mutuamente excluyentes, IIIb y IIIc, en el receptor 2 del factor de crecimiento fibroblástico (FGFR-2), C) sitios de *splicing* 5' alternativos en el gen *Wt1*, D) sitios de *splicing* 3' alternativos en el gen *transformer (tra)* de *Drosophila*, E) exones terminales alternativos en el gen que codifica para la calcitonina y para el péptido CGRP, F) promotor-primer exón alternativo en el gen de la cadena liviana de miosina (MLC), G) retención intrónica en el gen del canal de cloro específico de músculo (ClC-1) (Ladd and Cooper, 2002).

Aunque los exones constitutivos tienen sitios de *splicing* fuertes que son reconocidos eficientemente por el *spliceosoma*, muchos de ellos utilizan también estas secuencias auxiliares para reasegurar su reconocimiento (McCullough and Berget, 1997). En el caso de los exones alternativos, los sitios de *splicing* son subóptimos y los elementos auxiliares requeridos para su reconocimiento son utilizados, además, para modular la selección de estos sitios de *splicing* de manera contexto dependiente.

Existe un creciente número de casos en donde sitios de *splicing* alternativos están asociados tanto a *enhancers* como a *silencers*, lo cual sugiere que es el antagonismo dinámico entre estos lo que regula el *output* de procesamiento de una región alternativa (Smith and Valcarcel, 2000).

Proteínas SR versus hnRNPs

La mayoría de los ESEs se une a proteínas SR. Éstas promueven la inclusión de exones alternativos con sitios de *splicing* débiles por medio de interacciones con la maquinaria basal de *splicing* (Cartegni et al., 2002). La sobreexpresión de estos factores tiene efecto sobre las regiones alternativas que regulan y se especula acerca del rol de la concentración de los mismos en la regulación del fenómeno a nivel fisiológico (Ge and Manley, 1990; Krainer et al., 1990).

Constan de uno o dos motivos de reconocimiento del RNA (RRM) seguidos de un dominio RS que contiene repeticiones del dipéptido arginina-serina. Los RRM determinan la especificidad de unión a los ESEs y el RS participa en las interacciones proteína-proteína que determinan el reclutamiento de la maquinaria de *splicing* y el apareamiento de los sitios de *splicing* (Tacke and Manley, 1999), siendo los estados de fosforilación de las serinas de este dominio los reguladores primarios de la actividad de las mismas (Prasad et al., 1999). La familia incluye a SRp20, SRp30c, SRp38, 9G8, SRp40, SRp55, SRp70, SF2/ASF y SC35 (Graveley, 2000).

Las proteínas SR no sólo activan el *splicing* por medio de ESEs sino que también juegan un rol clave en el *splicing* constitutivo *in vitro*, donde al menos un miembro de la familia es necesario para que éste pueda ser llevado a cabo.

La evidencia sugiere que las proteínas SR actúan mediante distintos factores, aunque las interacciones con los mismos no están del todo claras: pueden estimular la unión de U2AF o U2 a sitios de *splicing* 3' débiles (Graveley et al., 2001; Lavigne et al., 1993; Zuo and Maniatis, 1996), reclutar U1 al sitio de *splicing* 5' (Kohtz et al., 1994; Wang et al., 1995) o inhibir el efecto de *hnRNPs* unidas a ESSs o ISSs (Kan and Green, 1999).

Se supone que se unen a los ESEs como parte de complejos multiproteicos, que incluyen a otros factores de *splicing* con dominios RS como SRm160/300, hTra2 y U1snRNP e inclusive *hnRNPs* (Blencowe et al., 1998; Eldridge et al., 1999; Li and Blencowe, 1999).

Las proteínas *hnRNPs* están asociadas a los precursores no procesados (hnRNAs = heterogeneous nuclear RNAs = precursores) y dentro de la familia coexisten tipos muy diferentes de proteínas, la mayoría de ellas involucradas en la inhibición de la inclusión de exones alternativos (Krecic and Swanson, 1999).

El miembro más estudiado, hnRNPA1, contiene dos dominios de unión al RNA (RNP-cs) y un dominio auxiliar de interacción proteica rico en glicina (Mayeda et al., 1994), siendo estructuralmente distinta a hnRNPH y F, las cuales cuentan con tres dominios de unión al RNA que reconocen dominios ricos en guaninas (Caputi and Zahler, 2001).

hnRNPA1 se une a **ESSs e ISSs** en diversos transcriptos y contrarresta el efecto positivo de SF2/ASF y SC35 en la inclusión de exones alternativos en los mismos (Caceres et al., 1994; Hanamura et al., 1998; Mayeda et al., 1993).

Existen evidencias de múltiples posibles mecanismos de funcionamiento de hnRNPA1: puede interferir directamente en el ensamblado del **spliceosoma**, bloquear las interacciones que promueven la definición exónica e inhibir estéricamente la unión de una proteína SR a un **ESE o ISE** adyacente (Zhu et al., 2001).

Cabe destacar que el rol de las proteínas SR y **hnRNPs**, en ciertos casos, se invierte: hnRNP H actúa como un activador del **splicing** cuando se une a un ESE en el exón 2 del transcripto Tat del HIV (Caputi and Zahler, 2002), mientras que SRp20 inhibe la inclusión del exón EDI de la fibronectina humana (Cramer et al., 1999b; de la Mata and Kornblihtt, 2006).

Las proteínas SR se almacenan, y probablemente se reciclan y ensamblan, en complejos macromoleculares, estructuras nucleares llamadas “puntos” o “botones” de **splicing** (**splicing speckles**) cuya identidad se asoció a los gránulos inter cromatínicos identificados anteriormente por microscopía electrónica (Lamond and Spector, 2003).

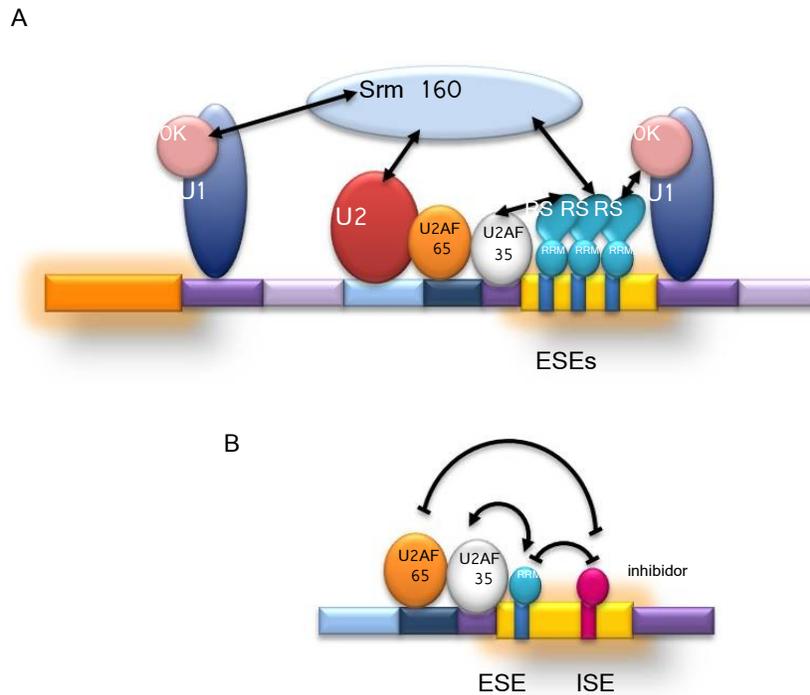


Fig.10. Modelos de acción de las proteínas SR unidas a ESEs. (A) Dependiente del dominio RS: Reclutamiento de U2AF35 y/o U1-70K a los sitios de *splicing* adyacentes 3' y 5' respectivamente e interacciones de reclutamiento indirectas (flechas negras) mediante puentes proteicos formados por Srm160, que interacciona con U2 snRNP. (B) Independiente del dominio RS: antagonismo del efecto negativo de proteínas inhibitorias unidas a ESSs yuxtapuestos y posible reclutamiento (en cooperación con el dominio RS) de U2AF65 por el dominio RRM (Cartegni et al., 2002).

La cinética de difusión de las proteínas SR presentes en los *speckles* hacia y desde el resto del núcleo es asombrosamente alta (Lamond and Spector, 2003; Phair and Misteli, 2000). Este dinamismo está íntimamente ligado a los niveles de transcripción de la RNA pol II: cuando la transcripción es inhibida con α -amanitina (un inhibidor específico de la transcripción por RNA pol II) o por shock térmico, los *splicing speckles* se vuelven más grandes y redondos y dejan de estar asociados a los sitios de transcripción de los mRNAs (Mintz et al., 1999), mientras que en contextos celulares en donde la transcripción de mRNAs es alta, las proteínas SR se distribuyen en el nucleoplasma (Lamond and Spector, 2003).

En experimentos de sobreexpresión de una mutante de RNA pol II cuyo dominio carboxilo terminal (*CTD= Carboxi Terminal Domain*) está truncado, la relocalización de los factores de *splicing* no ocurre (Misteli and Spector, 1999). La hipótesis dominante es que los factores de *splicing* son reclutados desde los *speckles* hacia los sitios de transcripción (Misteli et al., 1997) aunque existen pruebas de que algunos genes son reclutados en las vecindades de los *speckles* y sus precursores son transcritos dentro de estas estructuras, en donde serían procesados (Moen et al., 2004; Shopland et al., 2003). Lo cierto es que entre las proteínas asociadas a los *speckles* se encuentran las formas hiperfosforiladas de la RNA pol

II (Bregman et al., 1995), asociadas a la etapa de elongación transcripcional (Chesnut et al., 1992; Payne et al., 1989), así como diversos factores asociados a la transcripción (Saitoh et al., 2004).

Transcripción

Cada uno de los procesos celulares coincide con, y es dependiente de, patrones específicos de expresión génica precisos en el espacio y en el tiempo. El primer evento en la cadena de procesos celulares que lleva a la expresión génica consiste en la síntesis de moléculas de RNA a partir del molde de DNA, la cual en eucariotas es llevada a cabo por tres polimerasas, RNA polimerasa I, II y III, y sus respectivos factores de transcripción asociados. Las RNA polimerasas I y III son probablemente las más activas y están afectadas a la transcripción de genes de mantenimiento (*housekeeping*) asociados a la traducción: la RNA polimerasa I transcribe los precursores de los RNAs ribosomales mayores (28S, 18S y 5.8S) mientras que la RNA polimerasa III transcribe el RNA ribosomal 5S, los RNAs de transferencia (tRNAs) y el snRNA U6.

La RNA polimerasa II transcribe los mRNAs, a partir de los cuales se traducen casi todas las proteínas celulares, todos los snRNAs que llevan a cabo el proceso de *splicing*, excepto el U6 y la mayoría de los micro RNAs (miRNAs, que se encuentran en los intrones) (Kim and Kim, 2007). Dado que controla el primer paso de producción de toda la maquinaria enzimática celular, cualquier falla en la transcripción de la RNA polimerasa II se magnifica. Por ende, todas sus etapas están finamente reguladas y son el punto final de la mayoría de las cascadas de transducción de señales: la expresión errónea de programas genéticos causa desde malformaciones en el desarrollo hasta enfermedades como el cáncer.

Más regulación significa mayor complejidad. El conjunto de proteínas que conforman el aparato transcripcional asociado a la RNA polimerasa II ciertamente es el más sofisticado de los tres: solo en la pre-iniciación, un complejo gigante de alrededor de 60 proteínas que suman 3.100 kDa se une a los promotores de RNA polimerasa II.

La polimerasa II en sí misma consta de 12 subunidades que suman 500 kD, de las cuales la subunidad mayor (RPB1) posee rasgos distintivos: tiene un dominio carboxi terminal (CTD) que no tiene homólogos en las subunidades mayores del resto de las polimerasas, el cual le confiere propiedades especiales que exceden la transcripción y juega roles en el procesamiento de los transcritos (Neugebauer and Roth, 1997).

La transcripción mediada por polimerasa II consta de cuatro etapas: preiniciación, iniciación, elongación y terminación, y en cada una de estas etapas se suceden diferentes complejos de proteínas que asisten a la polimerasa II en sus diferentes funciones.

La polimerasa II es capaz de desenrollar DNA, sintetizar RNA y volver a enrollar el DNA nuevamente por sí misma, pero no así de reconocer un promotor e iniciar la transcripción. Para ello es necesaria la participación de los factores generales de transcripción (Kornberg, 2007).

Historia de un mRNA

La transcripción de un mRNA comienza con el reclutamiento de la RNA polimerasa II a un promotor, la secuencia de DNA que especifica el armado del complejo transcripcional y la transcripción. La polimerasa II es reclutada como un complejo de pre-iniciación (*PIC= Pre Initiation Complex*) que comprende también a las 26 subunidades de los factores generales de transcripción (*GTFs: TFIIB, D, E, F y H*) y a las 21 del complejo Mediador, que es el procesador de señales provenientes de proteínas activadoras y represoras unidas a secuencias *enhancers y silencers* transcripcionales respectivamente (Kornberg, 2005; Malik and Roeder, 2005).

En el epicentro del PIC, la proteína de unión al TATA (TBP), una subunidad del factor general de transcripción II D (TFIID), se une a esta secuencia presente en el promotor, dobla el DNA del promotor y lo adapta y configura físicamente a la superficie de la RPB1, que interactúa con TBP a través de dominio c-terminal de TFIIB (Nikolov et al., 1995). El DNA es acomodado de tal manera por el dominio N-terminal de TFIIB que el nucleótido ubicado 25-30 pb. río abajo del TATA box queda yuxtapuesto al sitio activo de la polimerasa, indicando el lugar en donde comenzará la transcripción. TFIIE entra en escena y recluta a TFIIH, cuya subunidad ATPasa/helicasa produce una tensión negativa en la hélice del DNA. Este desenrollamiento de la molécula del DNA produce una estructura de “burbuja transitoria” con separación de ambas hebras. TFIIH captura y se une a la hebra que no es molde de transcripción. La hebra de DNA molde se curva nuevamente y desciende hasta el sitio activo de la polimerasa II.

La síntesis del precursor de mRNA comienza y es estabilizada por la estructura *B finger* del extremo N-terminal de TFIIB, que puede cohabitar con el híbrido RNA-DNA el sitio activo de la RNA pol II hasta una elongación de 6 nucleótidos: si la transcripción es exitosa, el *B finger* es desplazado y TFIIB sale del complejo transcripcional, provocando el “escape del promotor” (Bushnell et al., 2004). TFIIH fosforila el CTD de la RPB1, liberando a la pol II de la mayoría de los GTFs e iniciando la fase de elongación.

La CTD de la RNA polimerasa II consta de 52 repeticiones de la héptada consenso YSPTSPS, cuyo estado de fosforilación cambia constantemente en las diferentes etapas del ciclo transcripcional: mediante experimentos de ChIP se demostró que la fosforilación en la serina 5 de las héptadas de la CTD se encuentra asociada al extremo 5' de los genes y contribuye al reclutamiento de la maquinaria de capping (McCracken et al., 1997), mientras que la fosforilación de las héptadas de la CTD en serina 2 se encuentra a través de todas las regiones codificantes y ayuda al reclutamiento del complejo de formación del extremo 3' del transcripto (Ahn et al., 2004; Hirose and Manley, 1997).

El complejo enzimático de *capping* (CEC), requerido para agregar la estructura *cap* en el extremo 5' del transcripto, está unido al CTD de la RNA polimerasa II antes de iniciarse la transcripción. Al emerger el extremo 5' del nuevo transcripto, el CEC se une al mismo y agrega el *cap* en una reacción que involucra la conversión del extremo 5' de trifosfato a difosfato por la actividad 5' - trifosfatasa, la adición de un monofosfato de guanosina por la mRNA guanililtransferasa para formar la unión 5' - 5' trifosfato y la metilación de la guanina por 2'-O-metiltransferasas. El *cap* metilado es reconocido por el CBC (*Cap Binding Complex*) que protege al transcripto de la degradación por nucleasas 5' - 3' y es requerido para la posterior exportación del núcleo e iniciación de la traducción (Proudfoot et al., 2002; Shatkin and Manley, 2000).

Luego de polimerizadas las primeras 20-40 bases, las probabilidades de abortar el transcripto y el fenómeno de *slippage* (en donde el transcripto se desliza río arriba, se re-aparea con la hebra molde y comienza nuevamente la polimerización) disminuyen dramáticamente (Pal and Luse, 2002, 2003). En algunos genes, en las vecindades del inicio de la transcripción existe un pausado de la RNA polimerasa II a la espera de señales que finalmente disparen la elongación plena (Kim et al., 2005).

La elongación, es decir, la polimerización de nucleótidos por unidad de tiempo en el proceso por el cual la RNA polimerasa II atraviesa un gen mientras sintetiza RNA, está regulada por múltiples factores. Entre los que interactúan directamente con RNA polimerasa II está la proteína TFIIIS, que en casos de arresto transcripcional por *backtracking*, activa la capacidad endonucleasa de la RNA polimerasa II y realinea el sitio catalítico con el extremo 3'-OH del RNA (Conaway et al., 2003; Fish and Kane, 2002); DSIF y NELF, que actúan positiva y negativamente en la regulación ante las pausas transcripcionales reversibles (Saunders 2006); TFIIIF, ELL, elongina y el complejo *Elongator*, que están involucradas en la regulación de la tasa catalítica de RNA polimerasa II mediante la alteración de su Km y/o su Vmax; y SWI-SNF, SpT6, FACT, PAF1, Set1 y Set2, que actúan indirectamente en la modificación y facilitación del pasaje de la polimerasa a través de la cromatina a ser transcripta por remodelación de nucleosomas o modificación de histonas (Arndt and Kane, 2003; Shilatifard et al., 2003). La actividad de estas últimas proteínas es particularmente importante en la modificación de los extremos N-terminales de las histonas H3 y H4, los cuales, según estén metilados, acetilados o fosforilados, confieren diferentes propiedades a los nucleosomas y pueden promover o inhibir la elongación de la RNA polimerasa II (Saunders et al., 2006). Todos estos factores contribuyen a la formación de un complejo de elongación posicionado alrededor de la RNA polimerasa II unida al híbrido DNA-RNA. Aunque usualmente el complejo es altamente procesivo y, una vez iniciada la transcripción, puede polimerizar un gran número de ribonucleótidos (Izban and Luse, 1991), existen situaciones en donde es específicamente disociado del templado: durante la condensación de los cromosomas, en la fase M del ciclo celular, el factor TTF2 se asocia a los complejos transcripcionales y los separa de la cromatina (Jiang et al., 2004), mientras que en la terminación de la transcripción, Pcf11 forma parte del grupo de proteínas que

cambia la conformación del complejo transcripcional y disocia a la RNA polimerasa II. Pcf11 está involucrado también en la terminación transcripcional prematura (Rosonina et al., 2006; Zhang and Gilmour, 2006; Zhang et al., 2007).

En el final de la transcripción, el complejo multiproteico de factores específicos de poliadenilación (CPSF) y clivaje (CstF), que está unido a la RNA polimerasa II, se transfiere al transcrito, uniéndose a dos secuencias lindantes, altamente conservadas, recién sintetizadas: AAUAAA y GU o ricas en U respectivamente. El transcrito se corta 35 bases río abajo de la secuencia AAUAAA, y este proceso es dependiente de la unión de la poli-adenilato polimerasa (PAP) al complejo. La PAP comienza la síntesis de una cola de poli-adenosinas de aproximadamente 200 bases. La proteína nuclear de unión a poli-adenosinas (PABPN1) se une con una estequiometría particular a la cola de poliA y actúa como calibre molecular, determinando su longitud. CstF, CPSF y PAP se disocian secuencialmente y PABPN1 permanece unido e interactúa junto con el *cap* 5' para promover la exportación nuclear y el inicio de la traducción.

El RNA sintetizado luego de la señal de corte y poliadenilación parece ser alcanzado por la proteína Xrn2, con actividad 5'-3' exonucleasa, que degrada al mismo, alcanzando a la polimerasa y llevándola a la desestabilización (Connelly and Manley, 1988; Kornblihtt, 2004; Luo and Bentley, 2004; Proudfoot, 1989).

Splicing cotranscripcional y acoplado a la transcripción

Existen evidencias que muestran que el *splicing* puede ocurrir durante la elongación transcripcional de la RNA polimerasa II. A finales de los 80, Beyer y Osheim realizan experimentos que descubren la cotranscripcionalidad del *splicing*. Observando preparaciones de genes coriónicos en embriones de *Drosophila* bajo el microscopio electrónico, encuentran loops en los transcritos, que coinciden con las estructuras de los *lariats* intrónicos del proceso de *splicing*, mientras éstos aún están siendo transcritos (Beyer and Osheim, 1988).

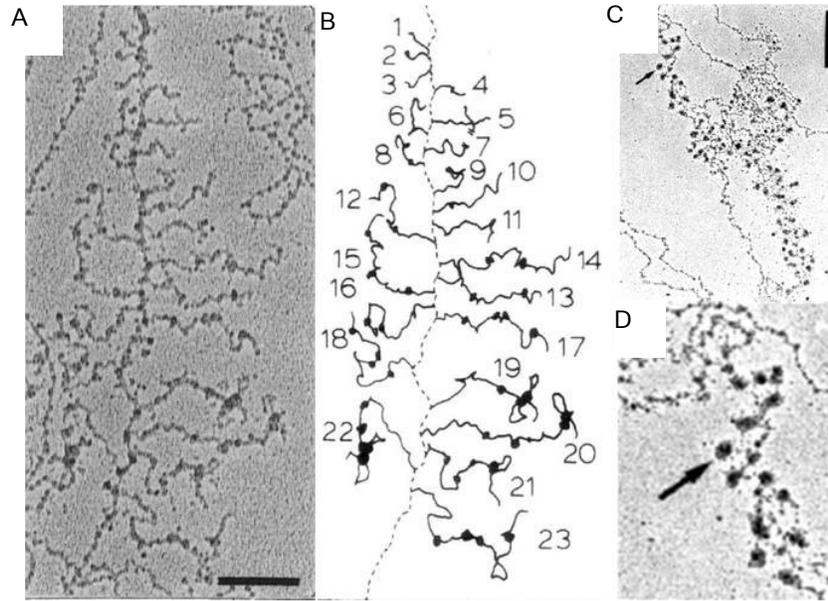


Fig. i11. Deposición de partículas de ribonucleoproteínas (RNPs) luego de la síntesis de sitio 3' de splicing de intrones cortos y largos. (A) Microfotografía electrónica (ME) de una unidad transcripcional del embrión de *Diosiphila*. La barra representa 0,2 μM . (B) Interpretación de la ME A. (C) Preparados de cromatina a pH fisiológico muestra los RNAs transcritos como gránulos ribonucleoprotéicos compactos de 60 nm. La barra representa 0,5 μM . (D) Ampliación de C (tomado de Osheim et.al., 1988).

También observan que los eventos de *splicing* en la parte 5' del transcripto se procesan, usualmente, con anterioridad a los de la región final. Esto es consistente con que cuentan con más tiempo para procesarse hasta que la polimerasa llega al final de la unidad transcripcional. Cuando las fibras de cromatina son expuestas a un *buffer* con pH fisiológico, a diferencia del pH básico en donde se despliegan las estructuras observadas de "árbol de navidad" (Fig. i11A), el RNA naciente se encuentra empaquetado en gránulos compactos de 60 nm. que se mantienen firmemente asociados a las fibras de cromatina (Fig. i11C-D), lo cual sugiere que partes lejanas en la estructura primaria del transcripto pueden estar en contacto directo en estas estructuras.

Las evidencias de cotranscripcionalidad del *splicing* también se observaron en genes humanos (Wuarin and Schibler, 1994) y en *Chironomus tentans* (Bauren and Wieslander, 1994). En cromosomas "plumulados" de anfibios, transcripcionalmente activos, tinciones con anticuerpos contra factores de *splicing* revelan marcas uniformes a lo largo de los mismos, que coinciden con las marcas de tinción de la RNA polimerasa II (Gall et al., 1999).

Al coincidir el *splicing* temporalmente con la transcripción, los intrones no pueden servir de sustrato hasta que ambos sitios de *splicing* son sintetizados, con lo cual el tiempo mínimo en el cual pueden reclutarse los factores de *splicing* y formarse los *spliceosomas* es el que lleva sintetizar cada intrón. El grupo de Neugebauer et. al. observó que el reclutamiento de los factores de *splicing* se realiza co-

transcripcionalmente, en coincidencia con la emergencia de los intrones (Gornemann et al., 2005; Kotovic et al., 2003; Listerman et al., 2006). Dado que el ensamblado del *spliceosoma* está ligado a la aparición de los sitios de *splicing*, la competencia entre sitios de *splicing* alternativos estaría regulada, por sobre todas las cosas, por el orden de aparición y el tiempo intermedio entre la emergencia de sitios que compiten. Esto abre la posibilidad de que la competencia entre sitios de *splicing* alternativos esté íntimamente ligada a la velocidad de elongación de la polimerasa. En el *splicing* del exón alternativo EDI, de la Mata et. al. observaron que utilizando una mutante lenta de la RNA polimerasa II la inclusión de este *cassette* exónico aumenta, lo cual es coherente con este modelo (de la Mata et al., 2003b). El parámetro clave de este tipo de regulación es el tiempo necesario para que un sitio de *splicing* nascente se una a los factores de *splicing* en *trans* que induzcan la formación de un *spliceosoma* funcional antes de la aparición del próximo sitio competitivo. La inclusión de pausas transcripcionales o la mutación de factores de elongación también tienen efectos sobre la selección de sitios de *splicing* alternativos en otros sistemas (Roberts et al., 1998), mientras que la demostración de que ciertos activadores transcripcionales influyen en el *splicing* alternativo mediante la modulación de la elongación de la RNA polimerasa II provee relevancia fisiológica para esta relación cinética, demostrando que el *splicing* alternativo puede regularse transcripcionalmente (Kadener et al., 2001). El fenómeno parece ser recíproco, ya que la adición de un intrón o el reclutamiento de snRNPs a un molde de DNA estimula la elongación de la RNA polimerasa II mediante la unión de los snRNPs a TAT-SF1, el cual se une y activa al complejo PTEF-b, cuya quinasa CDK9 fosforila en serina 2 a la CTD de la polimerasa (Fong and Zhou, 2001).

La remoción de intrones no siempre ocurre en su orden de síntesis: algunos intrones se procesan más rápidamente que otros. Estos últimos sufren *splicing* en el nucleoplasma, una vez terminada la transcripción (LeMaire and Thummel, 1990; Wetterberg et al., 1996).

Además de poder llevarse a cabo co-transcripcionalmente, existen evidencias de que el *splicing* puede estar acoplado físicamente a la transcripción, mediante interacciones de unión de factores de *splicing* a la RNA polimerasa II. Apoyando este argumento están las observaciones de que la CTD de la RNA polimerasa II estimula el *splicing* independientemente de su efecto sobre la formación del *cap* y el extremo 3' (Fong and Bentley, 2001) y de que la adición de la RNA polimerasa II o su CTD estimulan el *splicing* en ensayos *in vitro* (Hirose et al., 1999; Zeng and Berget, 2000). A su vez, la CTD es requerida por el factor de *splicing* SRp20 para ejercer su inhibición sobre la inclusión del exón EDI *in vivo* (de la Mata and Kornblihtt, 2006).

Recientemente, estudios de los proteomas asociados a inmunoprecipitaciones de la RNA polimerasa II y el *spliceosoma* por el grupo de Reed, demostraron que estos sistemas comparten un gran número de proteínas, entre las que se encuentran proteínas del snRNP U1, proteínas SR, factores de transcripción y proteínas de unión al RNA (Fig. i12) (Das et al., 2006; Das et al., 2007; Maniatis and Reed, 2002; Zhou et al., 2002). La hipótesis actual del modelo de acoplamiento implica una “fabrica de mRNA” en el cual el mRNA naciente emerge por el surco de salida de la RNA polimerasa II, donde se encuentra con los factores de *capping*, *splicing*, clivaje y poliadenilación unidos a la CTD de la polimerasa, de manera que el reclutamiento a los transcriptos sea eficiente (Cramer et al., 2001). Los factores de *splicing* podrían reclutarse en el promotor (Cramer et al., 1999b; Cramer et al., 1997). Este proceso podría estar mediado por interacciones con co-activadores transcripcionales como p52, que se une al factor de *splicing* SF2/ASF (Ge et al., 1998).

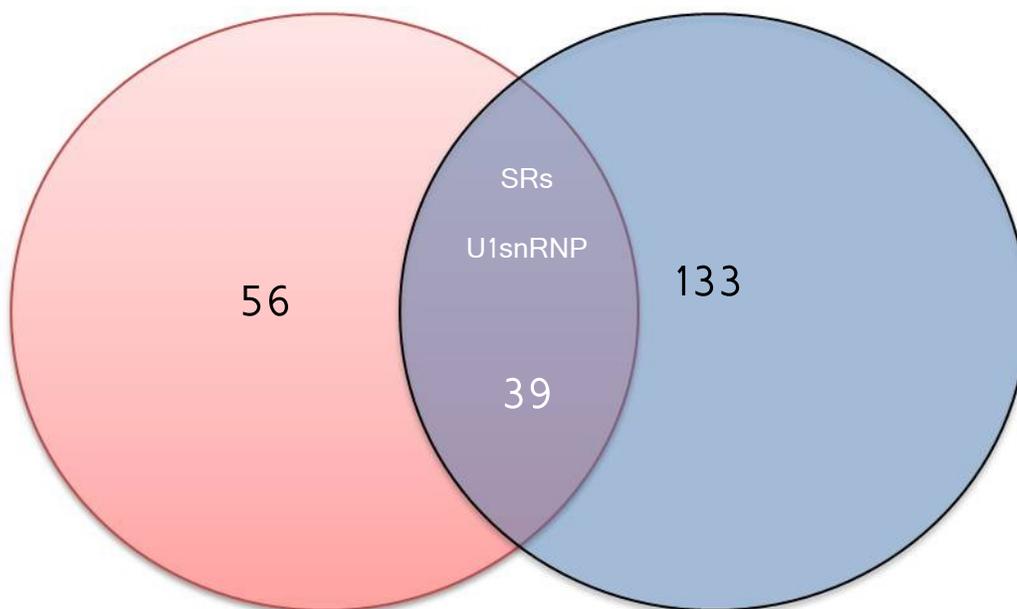


Fig.i12. Los proteomas asociados a la RNA polimerasa II y al spliceosoma comparten proteínas en común. Rosa: Conjunto de proteínas asociadas a la RNA polimerasa II, inmunoprecipitadas con el anticuerpo 8WG16 que reconoce las repeticiones YSPTSPS de la CTD (Das et al., 2007), Azul: conjunto de proteínas asociadas al spliceosoma (Zhou et al. 2002). Letras en blanco, conjunto intersección: proteínas inmunoprecipitadas en ambos complejos entre las cuales se encuentran proteínas SR y proteínas asociadas al U1 snRNA.

Splicing combinatorio: múltiples eventos, un transcripto

A través de la combinatoria de múltiples eventos de *splicing* alternativo, una unidad transcripcional puede generar una gran variedad de isoformas, expandiendo aún más el proteoma sin incrementar el número de genes (Smith and Valcarcel, 2000). Aunque teóricamente el número de posibles combinaciones en una unidad transcripcional con múltiples eventos de *splicing* alternativo a veces es enorme, la realidad es que en la mayoría de los genes estudiados, la cantidad de isoformas generadas es bastante menor (aunque no deja de ser importante): el caso quizás más conocido es el del gen CD44, en donde 20 isoformas dan cuenta de la mayoría de los transcritos, sobre una posible diversidad al azar de más de 1000 isoformas (Bell et al., 1998; Zhu et al., 2003); mientras que en el gen de la tropomiosina 1 (TPM1), que tiene un evento de exones mutuamente excluyentes (exones 6A versus 6B) y otro de exones terminales alternativos (exones 9+10 versus 11), las combinaciones 6A - 9+10 y 6B - 11 suman el 90% de las variantes (Xing et al., 2006). Estos casos plantean la hipótesis de la existencia de mecanismos de coordinación en el procesamiento de diferentes regiones de *splicing* alternativo. Una coordinación de este tipo tendría consecuencias en la estructura y función de las proteínas. La estructura terciaria funcional de una proteína se logra mediante un plegamiento complejo en donde residuos lejanos en la estructura primaria usualmente participan de estructuras en común, por ejemplo en el gen de la piruvato quinasa (PK), en donde la mayoría de los exones contribuye en la formación del core central de la misma (valores bajos de centralidad, Fig. i13) (Stoltzfus et al., 1994).

La expansión mencionada de las funciones proteicas de un gen, puede alcanzarse mediante la combinación lineal de dominios modulares independientes, o por la interacción funcional de dominios que cooperan. En vertebrados, los dominios proteicos están usualmente codificados por varios exones, siendo consistente con el hecho que la longitud exónica promedio, de 170 bases (o ≈ 57 aminoácidos) (Sakharkar et al., 2004), corresponde a la mitad de la longitud de un dominio proteico pequeño (Chothia and Gerstein, 1997; Murzin et al., 1995).

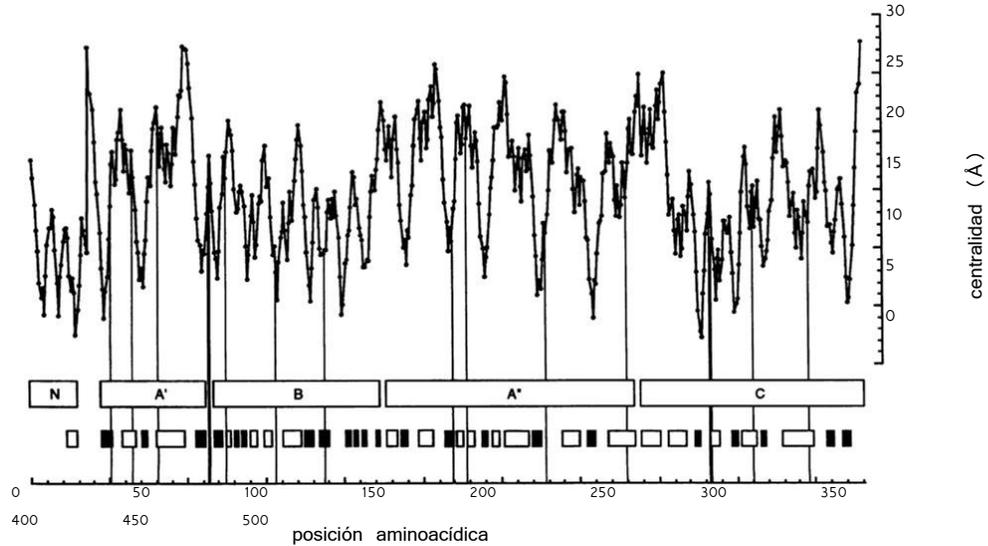


Fig.113. Posición de los intrones en el gen PK. Los dominios son codificados por varios exones y residuos muy distantes en la estructura primaria de la proteína, codificados por exones distantes, mantienen posiciones lindantes en la estructura terciaria. (A) Gráfico de centralidad que indica la posición de cada aminoácido con respecto al centro de la molécula de PK en su estructura terciaria. (B) los cinco dominios de PK, con uniones definidas por elementos estructurales secundarios: dominios A' y A (dos barriles β de 8 tiras), dominio B, y dominios N y C terminales (N y C). (C) Elementos de estructura secundaria: negro: Cintas β , blanco: α hélices. Líneas verticales negras: posición de los intrones.

¿Qué sucede cuando los exones involucrados, que codifican para dominios proteicos cooperativos, son alternativos? En este caso, la interacción funcional de dominios cooperativos podría conllevar la demanda de la coordinación de diferentes regiones de *splicing* alternativo.

En esta tesis se analizará la existencia de coordinación entre diferentes regiones de *splicing* alternativo, principalmente en el gen de la fibronectina, una proteína de la matriz extracelular de aproximadamente 250 kDa que tiene 3 regiones alternativas distantes bien definidas: EDB o EDII, EDA o EDI y V o IIICS. Estos exones alternativos están altamente regulados en diferentes procesos como la embriogénesis, la cicatrización y la formación de ciertos tumores, y pueden potencialmente generar 20 isoformas en humanos, 12 en roedores y 8 en pollo (Kornbliht et al., 1996).

La fibronectina está constituida de secuencias repetidas tipo I, II y III. Mediante los carbohidratos que la modifican, la fibronectina se une a una variedad de receptores celulares llamados integrinas. Estas interacciones le permiten jugar un rol importante en la organización del citoesqueleto y en el guiado de la migración celular, particularmente en la embriogénesis (Yamada, 2000). La estabilización de la estructura de la matriz extracelular es el resultado de su interacción con otros componentes de la misma, como el heparan sulfato, la fibrina y el colágeno (Lyon et al., 2000).

EDII y EDI son cassettes exónicos alternativos que codifican repeticiones peptídicas tipo III, mientras que IIICS es una región 3' alternativa en ratón (3 isoformas) y presenta retención intrónica más una combinación de

sitios de *splicing* 3' y 5' alternativos en humanos, lo que da como resultado 5 isoformas diferentes (Gutman and Kornblihtt, 1987).

El exón EDI es incluido en el transcripto maduro principalmente en la embriogénesis temprana (Norton and Hynes, 1987), luego de lo cuál su inclusión decae, sobreviviendo en ciertos tipos celulares como los fibroblastos (Kornblihtt et al., 1984) y procesos tisulares como la cicatrización (Muro et al., 2003).

La ventaja de utilizar el sistema fibronectina es que la regulación de sus regiones alternativas es una de las más estudiadas y, particularmente para el exón EDI, se conocen las secuencias en *cis* y los factores en *trans* que modifican su inclusión. El *splicing* alternativo de EDI es reproducido con fidelidad en transfecciones transitorias de minigenes, lo cual permite analizar cambios en el mismo producto de modificaciones en *cis* sobre secuencias específicas (Caputi et al., 1994a; Lavigueur et al., 1993; Vibe-Pedersen et al., 1984).

Objetivos

- Investigar la existencia de coordinación en el procesamiento de diferentes regiones alternativas en un mismo transcrito.
- Evaluar el nexo entre la transcripción de la RNA polimerasa II y la coordinación en el procesamiento de diferentes regiones alternativas en un mismo transcrito.
- Investigar la coordinación en el genoma humano, mediante análisis de bases de datos de mRNAs, ESTs y proteínas.

Resultados

Parte I - Coordinación en el gen de la fibronectina

Minigenes con dos regiones de *splicing* alternativo para estudiar la coordinación

Para investigar la coordinación entre dos regiones de *splicing* alternativo en un mismo transcripto, construimos minigenes reporteros con dos regiones de *splicing* alternativo de la fibronectina (FN) humana, los cuales transfectamos transitoriamente en células Hep3B. Reporteros similares, con sólo una región de *splicing* alternativo, fueron utilizados extensivamente para estudiar el control del *splicing* alternativo del exón EDI (Caputi et al., 1994b; Cramer et al., 1999b; Vibe-Pedersen et al., 1984). Las construcciones utilizadas en esta tesis contienen dos regiones de *splicing* alternativo EDI que están separadas por tres exones constitutivos del gen de alfa-globina humana y sus correspondientes intrones, bajo el control de diferentes promotores (Fig. r1).

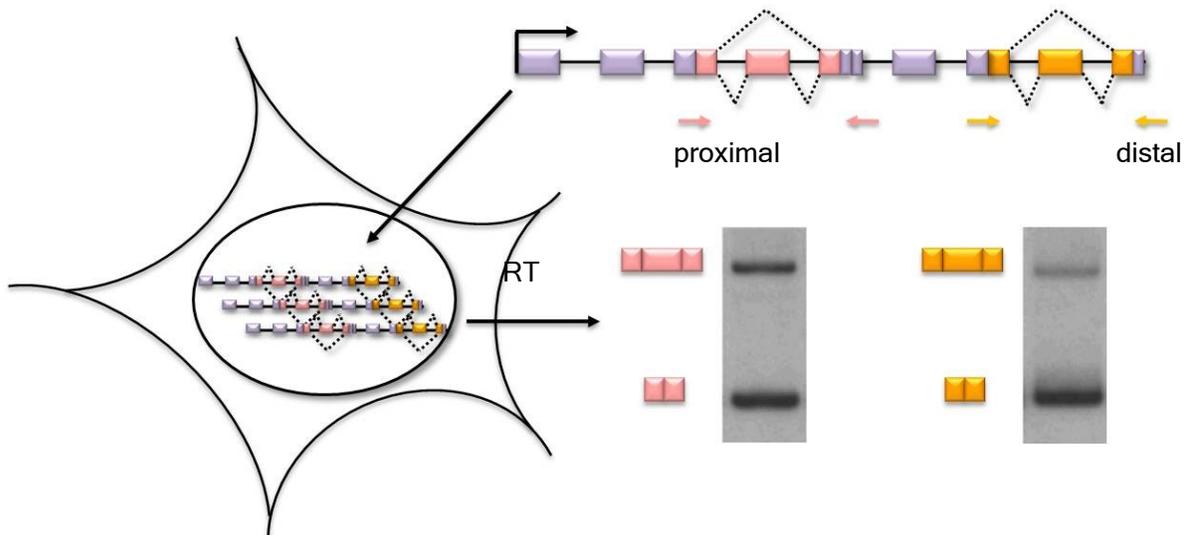


Fig. r1. Esquema de las construcciones tándem utilizadas en la tesis. 48hs. después de transfectadas las células Hep3B, se extrajo el RNA de las mismas y se analizó por RT-PCR específica la inclusión de cada una de las regiones alternativas.

La longitud de los exones alternativos EDI presentes en los minigenes es de 270 bases, de manera que las isoformas de *splicing* que contienen y excluyen al exón alternativo tienen el mismo marco de lectura traducional. El análisis del *splicing* alternativo de los transcritos provenientes de los minigenes

transfectados se realizó mediante la técnica de RT-PCR, con pares de *primers* específicos para cada una de las regiones alternativas, proximal y distal.

Coordinación entre eventos de *splicing* alternativo en un mismo transcripto

Los transcriptos provenientes de la expresión transitoria de la construcción conteniendo dos exones EDI WT bajo el control del promotor de fibronectina (pFN-EDI^{WT}/EDI^{WT}) fueron analizados por RT-PCR con pares de *primers* específicos para cada región alternativa. Dado que la inclusión del exón alternativo EDI en transcriptos de minigenes regulares transfectados en células Hep3B es muy baja (Cramer et al., 1997), los minigenes se co-transfectaron con pequeñas cantidades (2 ng.) de un plásmido que expresa la proteína SR SF2/ASF para estimular levemente la inclusión y permitirnos ver los efectos sobre la misma de una manera más clara. Bajo estas condiciones los cocientes de inclusión para la zona alternativa proximal y distal fueron respectivamente de 0.78 ± 0.04 y 0.24 ± 0.02 respectivamente.

El exón alternativo EDI posee un *enhancer* de *splicing* exónico (ESE) que contiene la secuencia núcleo 5'-GAAGAAGAC-3', la cual es una secuencia blanco para las proteínas SR SF2/ASF y 9G8 (Buratti and Baralle, 2004; Cramer et al., 1999b). Sorprendentemente, la disrupción del ESE en el exón EDI proximal (construcción pFN-EDI^{ΔESE}/EDI^{WT}) no sólo suprime su propia inclusión, sino que provoca una disminución de 8 veces en la inclusión del exón EDI distal, localizado aproximadamente a 3400 bases río abajo en el transcripto (Fig. r2).

proximal

distal

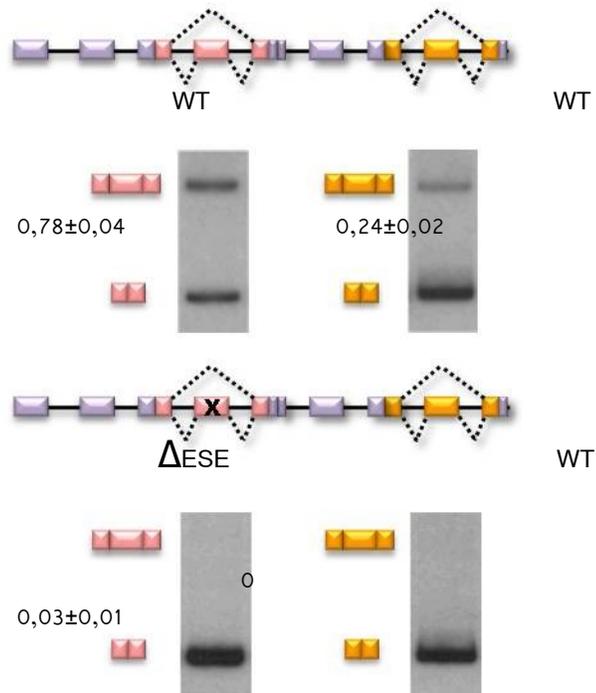


Fig. r2. Coordinación en el procesamiento de dos regiones de *splicing* alternativo en un mismo transcrito. Efecto de la disrupción del ESE (*Exonic Splicing Enhancer*) en la región de *splicing* alternativo proximal del minigen tándem. Células Hep3B fueron transfectadas con 800 ng. de plásmido pFN-EDI^{WT} /EDI^{WT} (arriba) o pFN-EDI^{ΔESE} /EDI^{WT} (abajo) y 2 ng. del vector de expresión de la proteína SF2/ASF g10SF2/ASF (Caceres et al., 1997). Los patrones de *splicing* alternativo de la región alternativa EDI proximal (rosa) y distal (naranja) fueron analizados por RT-PCR radioactiva con pares de *primers* específicos para cada sitio, seguida por electroforesis en geles de poliacrilamida nativos, como se describe en la sección Procedimientos Experimentales. Los datos representan el promedio de 3 transfecciones independientes en pocillo5 de 35mm., con triplicados para cada condición. Los valores \pm representan el desvío estándar.

La coordinación en el procesamiento de los eventos de *splicing* alternativo tiene polaridad

Al contrario de la observación anterior, la disrupción del ESE en el exón EDI distal (construcción pFN-EDI^{WT} /EDI^{ΔESE}; Fig. r3 abajo) suprime su propia inclusión, pero tiene un efecto mucho menor sobre la inclusión del exón EDI proximal. Estos experimentos indican fuertemente que ambos eventos de *splicing* alternativo están condicionados a través de un mecanismo que exhibe una polaridad con respecto al promotor (Fig. r3).

proximal

distal

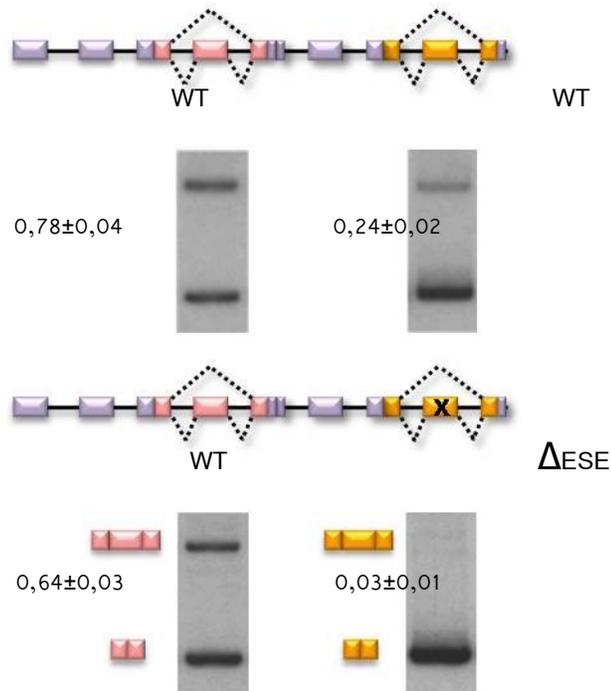


Fig. r3. Polaridad en la coordinación de dos regiones de *splicing* alternativo en un mismo transcrito. Efecto de la disrupción del ESE (*Exonic Splicing Enhancer*) en la región de *splicing* alternativo distal del minigen tándem. Células Hep3B fueron transfectadas con 800 ng. de plásmido pFN-EDI^{WT} /EDI^{WT} (arriba) o pFN-EDI^{WT} /EDI^{ΔESE} (abajo) y 2 ng. del vector de expresión de la proteína SF2/ASF g10SF2/ASF (Caceres et al., 1997). Los patrones de *splicing* alternativo de la región alternativa EDI proximal (rosa) y distal (naranja) fueron analizados como se describe en la sección Procedimientos Experimentales. Los datos representan el promedio de 3 transfecciones independientes en pocillos de 35mm., con triplicados para cada condición. Los valores ± representan el desvío estándar.

El efecto de coordinación es observado tanto en la ausencia como en presencia de la sobreexpresión de SF2/ASF (0, 20 y 60 ng. de plásmido de expresión de SF2/ASF co-transfectado; Fig. r4). La figura también muestra que la sensibilidad de inclusión del exón alternativo EDI distal a SF2/ASF es menor cuando el ESE del exón EDI proximal está mutado. Dado que ambos exones EDI (proximal y distal) son blanco de SF2/ASF, es difícil saber si la menor sensibilidad de los transcritos provenientes de la construcción pFN-EDI^{ΔESE}/EDI^{WT} es la consecuencia de que la unión de SF2/ASF está reducida en el exón EDI proximal o si lo está en ambos, causada por un efecto de coordinación.

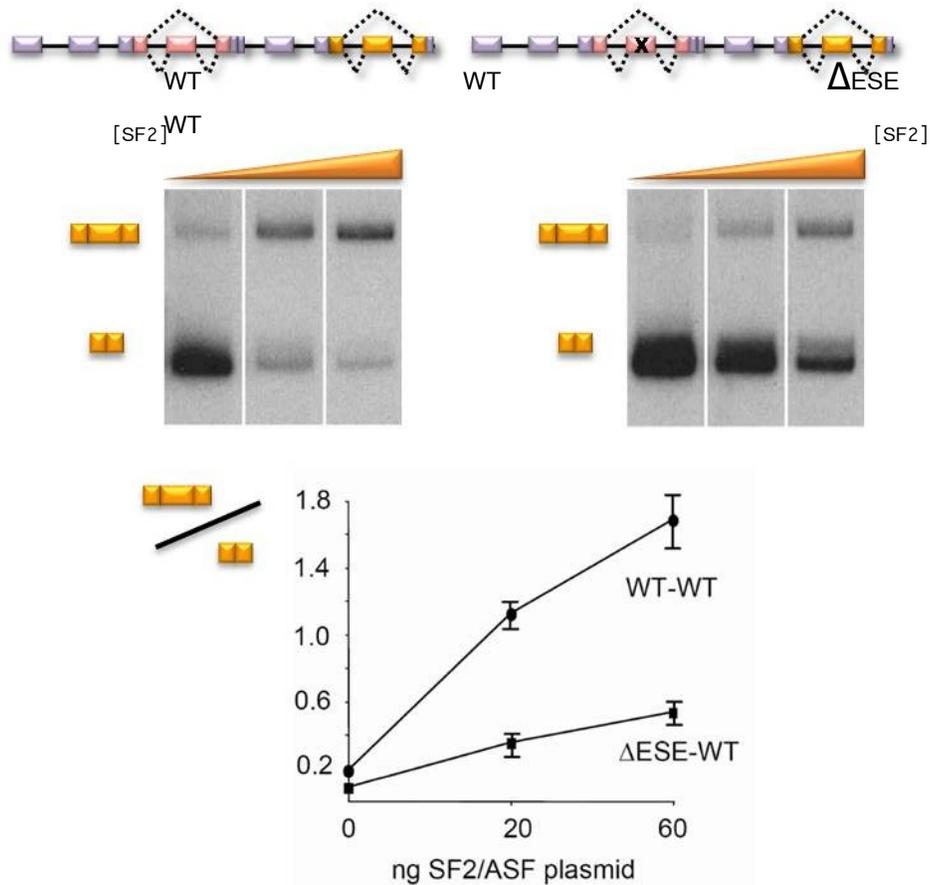


Fig. r4. Curva dosis respuesta de SF2/ASF sobre el *splicing* alternativo de la región EDI distal cuando la región proximal es WT o tiene el ESE (*Exonic Splicing Enhancer*) disrumpido. Células Hep3B fueron transfectadas con 800 ng. de plásmido pFN-EDJ^{WT} /EDJ^{WT} (izquierda) o pFN- EDJ^{ΔESE} /EDJ^{WT} (derecha) y cantidades crecientes del vector de expresión de SF2/ASF g10SF2/ASF (Caceres et al., 1997). Los patrones de *splicing* alternativo de la región alternativa EDI distal (naranja) fueron analizados como se describe en la sección Procedimientos Experimentales. Los datos representan el promedio de 3 transfecciones independientes en pocillos de 35mm., con triplicados para cada condición. Los valores ± representan el desvío estándar.

La coordinación polar depende del nivel de inclusión y no del pegado de la proteína SF2/ASF al ESE del exón proximal

Dado que la disrupción del ESE en el EDI proximal previene la unión de SF2/ASF al mismo y a la vez afecta la inclusión del exón EDI distal (Fig. r2), parece importante definir si el efecto de coordinación es causado por la capacidad de SF2/ASF de unirse al ESE del exón EDI proximal o a la inclusión del exón EDI proximal *per se*. Para dilucidar este problema, se construyeron minigenes tándem en donde la inclusión

de uno de sus exones EDI es constitutiva, mediante mutación del sitio de *splicing* subóptimo 3' (Nogues et al., 2003) en vez de sobreexpresar SF2/ASF. En ausencia de la sobreexpresión de SF2/ASF y luego de la transfección de células Hep3B con pFN-EDI^{WT}/EDI^{WT}, el cociente de inclusión para los exones EDI proximales y distales, fue respectivamente de 0.32 ± 0.08 y 0.13 ± 0.01 (Fig. r5A). Cuando el exón proximal es hecho constitutivo (pFN-EDI^C/EDI^{WT}), la inclusión del exón distal se incrementa más de 7 veces (Fig. r5B). Sin embargo, cuando el exón distal es convertido en constitutivo (pFN-EDI^{WT}/EDI^C), la inclusión del exón EDI proximal sólo se incrementa 1.8 veces (Fig. r5C). Este comportamiento diferencial nos indica nuevamente, y de forma independientemente, la existencia de polaridad en los efectos de coordinación.

Los experimentos arriba analizados muestran que a mayor inclusión del exón EDI proximal, mayor es la inclusión del evento EDI distal. Sin embargo, no resuelven si en este efecto está implicada la unión de SF2/ASF al exón proximal. Esta pregunta se abordó expresando construcciones tándem con un exón proximal EDI constitutivo que a su vez contenía el ESE mutado (pFN-EDI ^{Δ ESE}/EDI^{WT}).

Los niveles de inclusión de los eventos proximal y distal del minigen pFN-EDI ^{Δ ESE}/EDI^{WT} son similares a los de la construcción pFN-EDI^C/EDI^{WT}, en donde el exón EDI proximal tiene un ESE WT y es constitutivo.

Este nivel de inclusión en el evento distal de dichas construcciones es aproximadamente 6 veces mayor que en la construcción en donde ambos exones EDI son WT (Fig. r6C).

proximal

distal

A

B

C

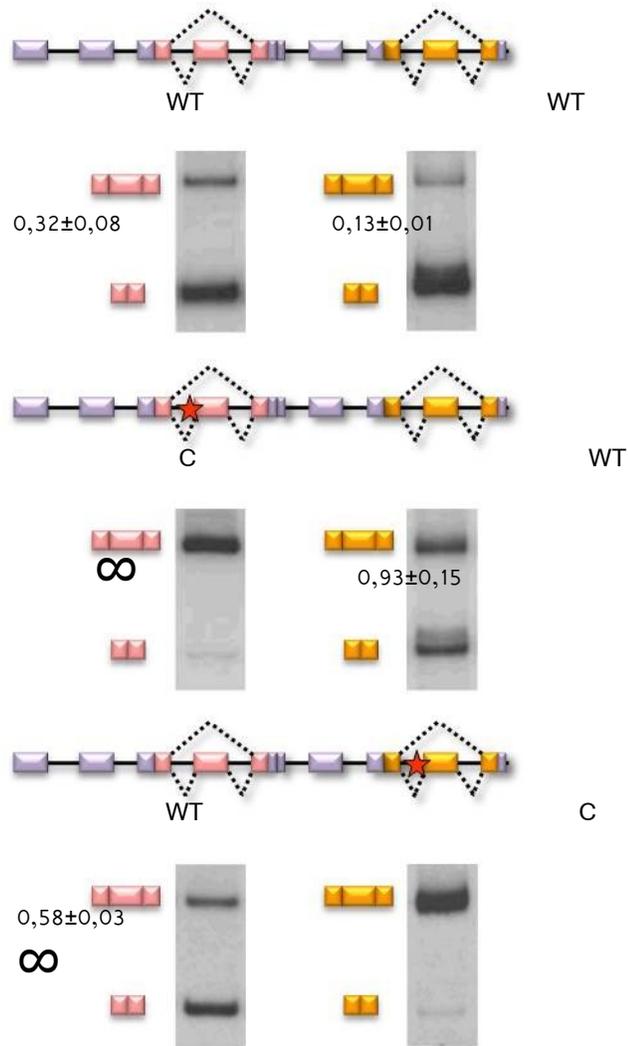


Fig. r5. Efecto de la inclusión constitutiva de uno de los exones EDI en la coordinación por optimización del tracto de polipirimidinas en el sitio de *splicing* 3' (Nogues et al., 2003). Células Hep3B fueron transfectadas con los plásmidos pFN-EDI^{WT}/EDI^{WT} (A), pFN-EDI^C/EDI^{WT} (B) y pFN-EDI^{WT}/EDI^C (C) en las mismas condiciones que en la figura 2. Los datos representan el promedio de 3 transfecciones de 3 pocillos de 35 mm. para cada condición. Los valores ± representan el desvío estándar.

Esto indica que son los niveles de inclusión del exón EDI proximal y no la unión de proteínas SR al ESE del mismo los que causan los efectos de coordinación.

Dependencia del ESE en el exón distal para la coordinación polar

Para analizar si el efecto polar observado sobre el exón distal es mediado por un mecanismo en el cual está involucrado su ESE, construimos un minigen en donde la inclusión del primer exón EDI es constitutiva y la segunda región alternativa contiene una delección del ESE en el exón EDI distal. Al analizar los niveles de inclusión de la construcción pFN-EDI^C/EDI^{ΔESE}, observamos que el ESE del exón alternativo distal es necesario para que su inclusión sea afectada por los niveles de inclusión del exón EDI proximal (Fig. r6D).

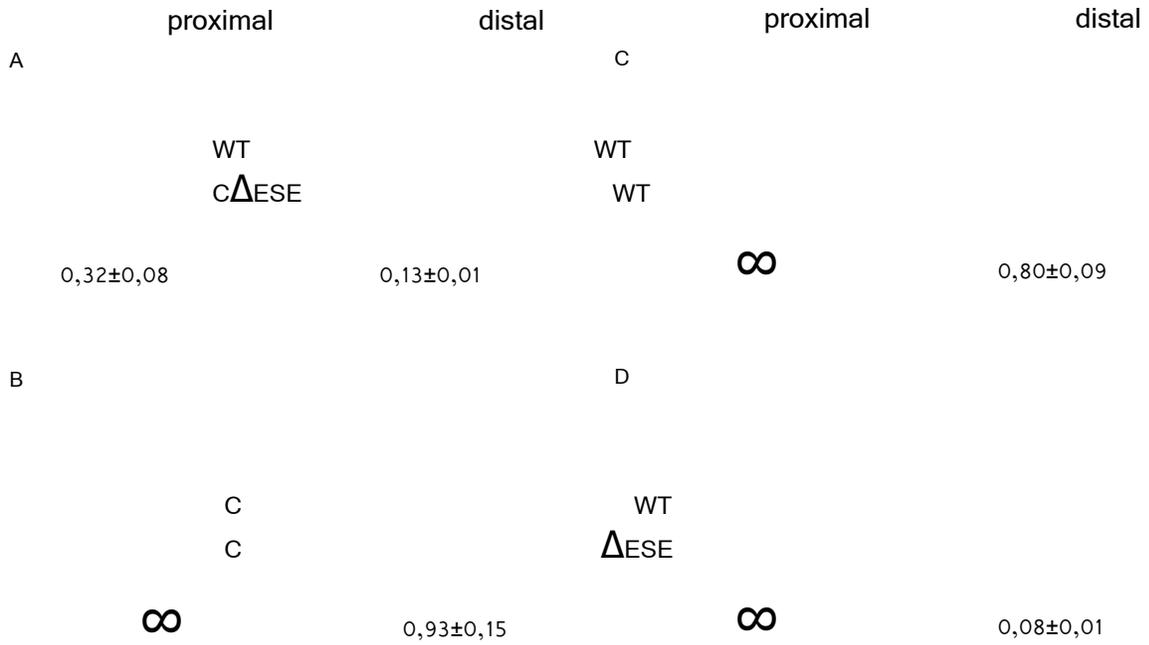
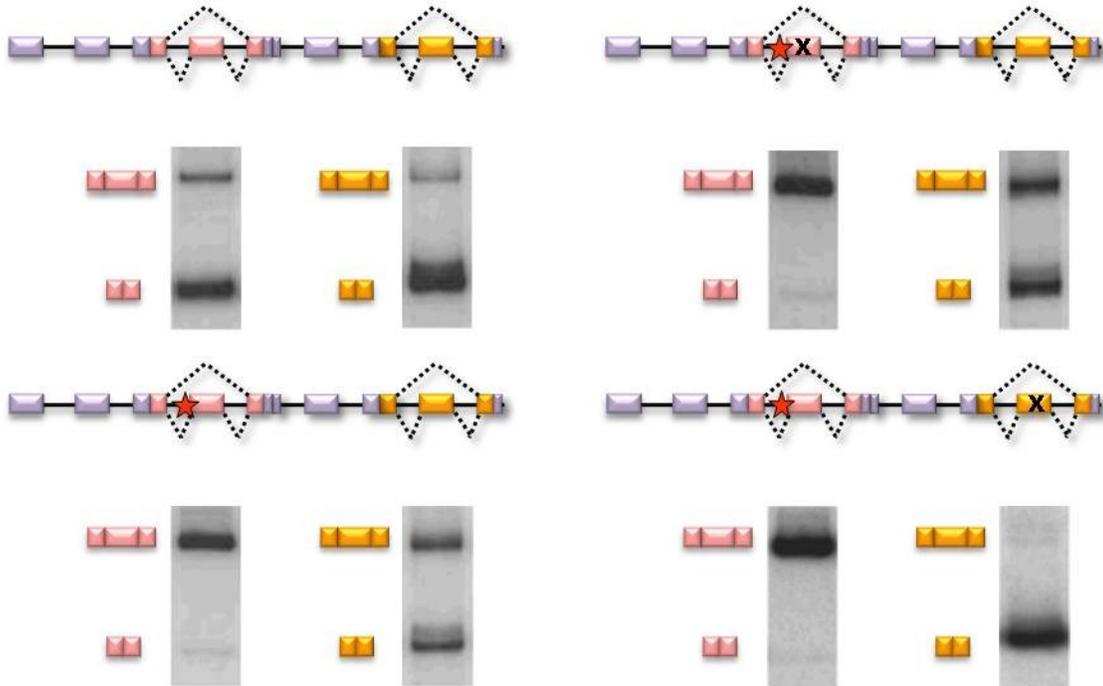


Fig. r6. Dependencia del nivel de inclusión proximal y del ESE en el exón distal para la coordinación polar. Células Hep3B fueron transfectadas con los plásmidos pFN-EDI^{WT}/EDI^{WT} (A), pFN-EDI^C/EDI^{WT} (B), pFN-EDI^{c Δ ESE}/EDI^{WT} (C), y pFN-EDI^C/EDI ^{Δ ESE} (D) en las mismas condiciones de cotransfección que en la figura r2, aunque sin el agregado de vector de expresión de ASF/SF2. Los datos representan el promedio de 3 transfecciones de 3 pocillos de 35 mm. para cada condición. Los valores \pm representan el desvío estándar.



El grado de inclusión de los exones río arriba de la región alternativa distal no es el causante del efecto de coordinación

Los experimentos descritos hasta ahora dejan abierta la duda de si los efectos de coordinación observados son causados por la inclusión en la región proximal del exón EDI o si la inclusión constitutiva de cualquier otro exón podría causar el mismo efecto. Para responder a esta pregunta transfectamos células Hep3B con una serie de diferentes construcciones con distintas configuraciones proximales, que difieren en el número de exones constitutivos río arriba de la región alternativa EDI distal, las cuales tienen o no la secuencia de EDI proximal (Fig. r7).

Una primera observación relevante realizada sobre estos nuevos minigenes es que los niveles de inclusión del exón EDI distal no correlacionan con el número de exones constitutivos que se hallan río arriba: construcciones con seis (Figs. r7Aa, r7Ab, yr7Af-r7Ah), tres (Fig. r7Ai), o un (Fig. r7Aj) exón constitutivo río arriba de la región alternativa distal muestran niveles de inclusión del EDI distal que son mucho menores que aquellos de la construcción con inclusión constitutiva en el exón proximal (Fig. r7Ac).

Rol de la definición exónica de EDI en la coordinación

Hasta este punto, la información indicaba que el efecto de coordinación sobre la zona distal del transcripto era mediado específicamente por la inclusión del exón alternativo EDI en el procesamiento de la zona proximal,

dado que la inclusión constitutiva de cualquier otro exón en dicha zona no ejercía ningún efecto sobre el grado de inclusión distal.

La pregunta que surgió fue la siguiente: ¿es la presencia de la secuencia del exón EDI en la porción procesada de la región proximal la que ocasiona el efecto de coordinación sobre la región alternativa distal o es el proceso de inclusión, es decir de definición exónica, del exón EDI en la región proximal la que origina el fenómeno? Para responder esto se analizaron los resultados provenientes de construcciones en donde la secuencia del exón EDI proximal deja de ser un exón en si mismo sino que está insertada como parte de un exón proximal constitutivo híbrido.

La presencia del exón EDI en este exón constitutivo híbrido no es suficiente para provocar altos niveles de inclusión del exón EDI distal (comparar figs. r7Ad y r7Ae) y esto es independiente de la posición (comparar figs. r7Ak y r7Aj). El interrogante que se planteó entonces es si alguno de los intrones proximales flanqueantes del exón EDI jugaba *per se* un rol en promover la inclusión distal del exón EDI. Ninguno de los dos intrones, el intrón -1 o el +1 por sí solos (Figs. r7Af y r7Ah) son capaces de reproducir los efectos observados en la figura r7Ac. Por otra parte, las construcciones f y g (ver figs. r7Af y r7Ag respectivamente) provocan ambas bajos niveles de inclusión del exón EDI distal, lo cual descarta un rol intrínseco de la mutación usada en la construcciones c y g para optimizar y llevar al consenso el trecho de polipirimidinas del intron -1 proximal. Esta optimización del trecho de polipirimidinas es entonces sólo relevante si contribuye al proceso de inclusión del exón EDI proximal. Estos resultados, a su vez, indican que lo que afecta la inclusión del exón alternativo distal es el proceso de definición exónica del exón alternativo EDI proximal.

La definición exónica es el mecanismo por el cual los sitios de *splicing* son seleccionados mediante interacciones entre factores de *splicing* a través del exón antes del ensamblado del *spliceosoma* y la remoción de los intrones (Berget, 1995). Dado que otros exones constitutivos (que no son EDI) bien definidos, presentes en los minigenes, no parecen estimular la inclusión del exón alternativo EDI río abajo en la posición distal (Figs. r7Ab y r7Ac-r7Ak), podemos concluir que la definición específica del exón alternativo EDI juega un rol fundamental en el reclutamiento de factores de *splicing* u otras proteínas que influyen en la coordinación.

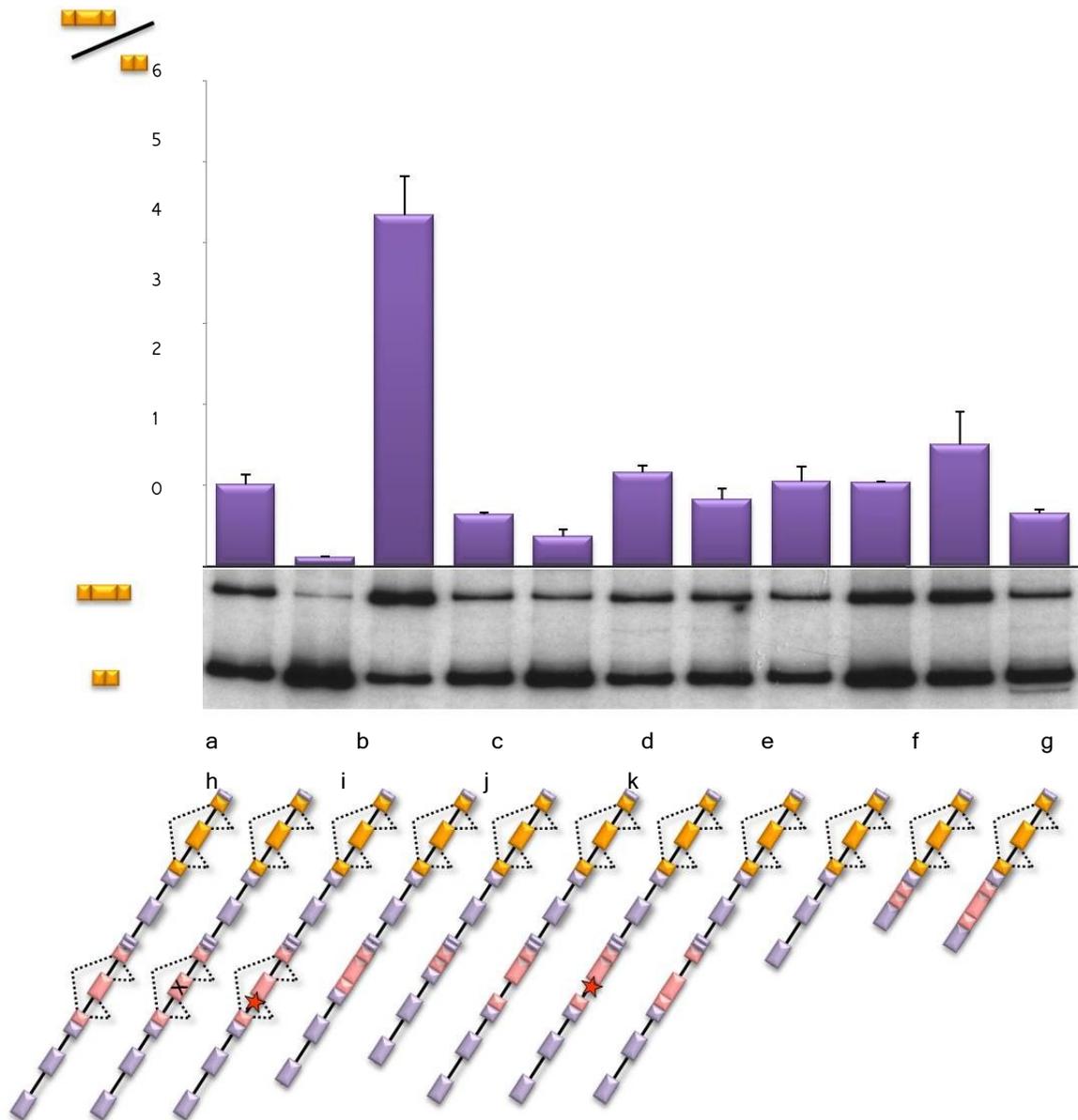


Fig. r7. Efecto sobre la inclusión del exón distal EDI, de las configuraciones proximales que difieren en el número de exones constitutivos y en la presencia de la secuencia EDI. Las construcciones a (pFN-EDI^{WT}/EDI^{WT}), b (pFN-EDI^{ΔESE}/EDI^{WT}) y c (pFN-EDI^C/EDI^{WT}) ya descritas fueron utilizadas como controles. En las construcciones d, f-h y k, la secuencia EDI no está presente como parte de exones proximales constitutivos. Células Hep3B fueron transfectadas con las construcciones descritas y el *splicing* alternativo se analizó como se describe en la sección Procedimientos Experimentales. Los datos representan el promedio de tres transfecciones independientes en pocillos de 35 mm., con triplicados para cada condición. Las barras de error representan el desvío estándar.

El promotor alfa-globina suprime la polaridad

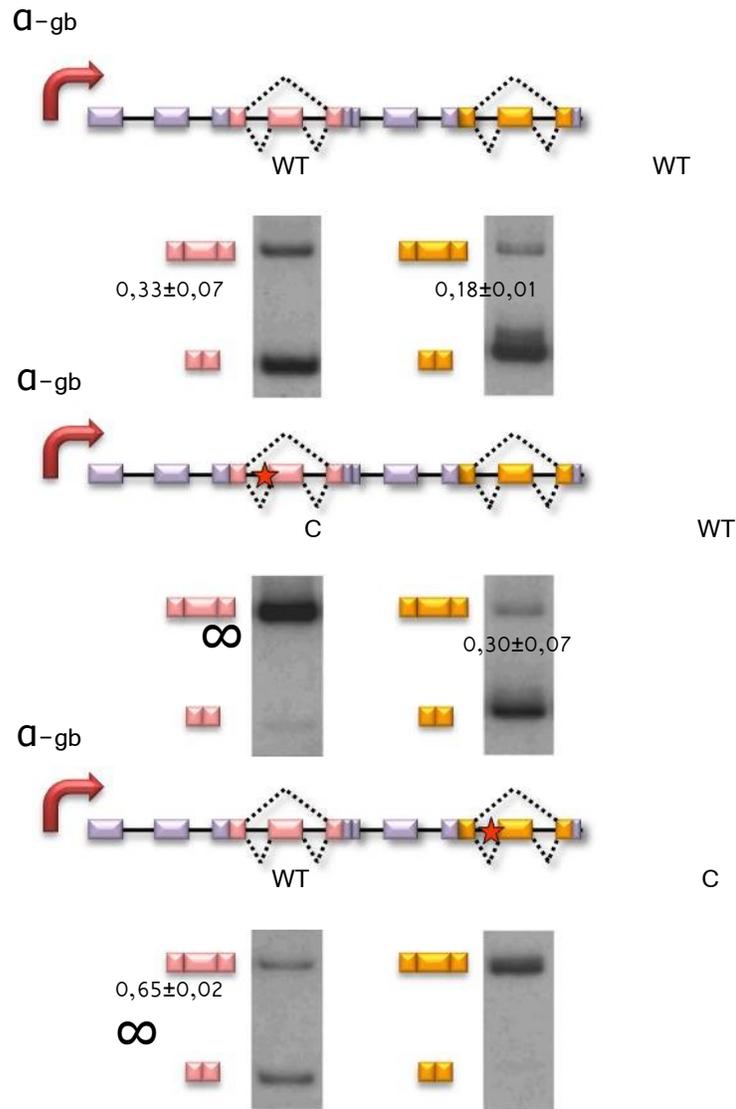
Los experimentos de las figuras r2-r7 fueron realizados con construcciones bajo el control del promotor de FN humana. Dado que el *splicing* alternativo del exón EDI es afectado por la identidad del promotor (Cramer et al., 1997), se evaluó si la coordinación y su comportamiento polar eran afectados por el tipo de promotor. Para este propósito se prepararon minigenes tándem en donde el promotor de FN humana fue reemplazado por el de la alfa globina humana (α -gb). La estimulación de la inclusión del EDI distal fue mucho menor cuando la transcripción fue controlada por el promotor de α -gb: los niveles de inclusión del exón EDI distal se incrementaron solo 1,6 veces cuando al exón EDI proximal se lo hizo en constitutivo (Fig. r8A), versus las 6 o 7 veces de observado con el promotor de FN (Fig. r5 y r6). Más importante, un nivel similar de estimulación sobre el exón EDI proximal WT (1,9 veces) se observó cuando se convirtió el exón EDI distal en constitutivo mejorando el trecho de polipirimidinas en su intrón -1. Estos resultados indican que bajo el control del promotor de α -gb la polaridad desaparece, dado que ambas regiones pueden influenciarse en niveles similares una a la otra.

La coordinación es recuperada por la inhibición de la elongación de la **RNA** polimerasa II

Anteriormente Kadener et al. demostraron que al cambiar el promotor que controla la transcripción de los minigenes EDI, también cambia la densidad de la RNA polimerasa II a través del DNA molde, mostrando que la elongación de la polimerasa era varias veces mayor en aquellos minigenes en donde la transcripción era gobernada por el promotor de α -gb (Kadener et al., 2002). Dado que las construcciones muestran alteraciones en la polaridad al cambiar el promotor, es posible que la elongación tenga un rol sobre el efecto polar. La fosforilación de la polimerasa II en el residuo serina 2 de las repeticiones del CTD por la actividad quinasa de PTEF-b (factor positivo de elongación transcripcional b) estimula la elongación transcripcional. El análogo de nucleótido DRB (5,6-dichloro-1- β -D-ribofuranosylbenzimidazole), un inhibidor de la subunidad quinasa Cdk-9 de PTEF-b, inhibe la elongación de la polimerasa II (Price, 2000). Si el promotor de α -gb humana inhibe el efecto polar por controlar la transcripción de manera de hacer la polimerasa “más elongadora”, al inhibir la elongación de la polimerasa II con DRB se debería recobrar el efecto polar en dichas construcciones.

proximal

distal



Células transfectadas con las construcciones tándem bajo el control del promotor α -gb fueron tratadas con DRB. La figura r9 muestra que cuando la transcripción es controlada por el promotor de α -gb, los niveles de inclusión del exón alternativo EDI distal provocados por la constitutividad del EDI proximal se incrementan a 3 veces en la presencia de DRB.

La replicación de los minigenes provoca una estructura cromatínica más compacta en los mismos, acompañada de una menor elongación de la RNA polimerasa II que a su vez provoca niveles de inclusión del exón EDI más altos (Kadener et al., 2001). Las construcciones tándem utilizadas en esta tesis tienen origen de replicación del virus SV40, el cuál se activa en presencia de la versión larga del antígeno T del virus SV40 (T-Ag), provocando la replicación de los mismos y una cromatinización concomitante. Nuestro grupo demostró que esta cromatinización inhibe la elongación de la RNA polimeración II, aumentando los niveles de inclusión de un minigen EDI (Kadener et al., 2001). Cuando la replicación de estos minigenes tándem bajo el control del promotor de α -gb es activada mediante la co-transfección con un plásmido que expresa el antígeno T, el exón constitutivo proximal provoca un efecto 1.7 veces mayor en la inclusión del exón EDI distal wt,

comparado con el efecto en moldes que no se han replicado (Fig. r9). Así, dos maneras independientes de inhibir la elongación de la polimerasa II (DRB y replicación mediada por T-Ag) revierten la ausencia de coordinación observada con el promotor de α -gb.

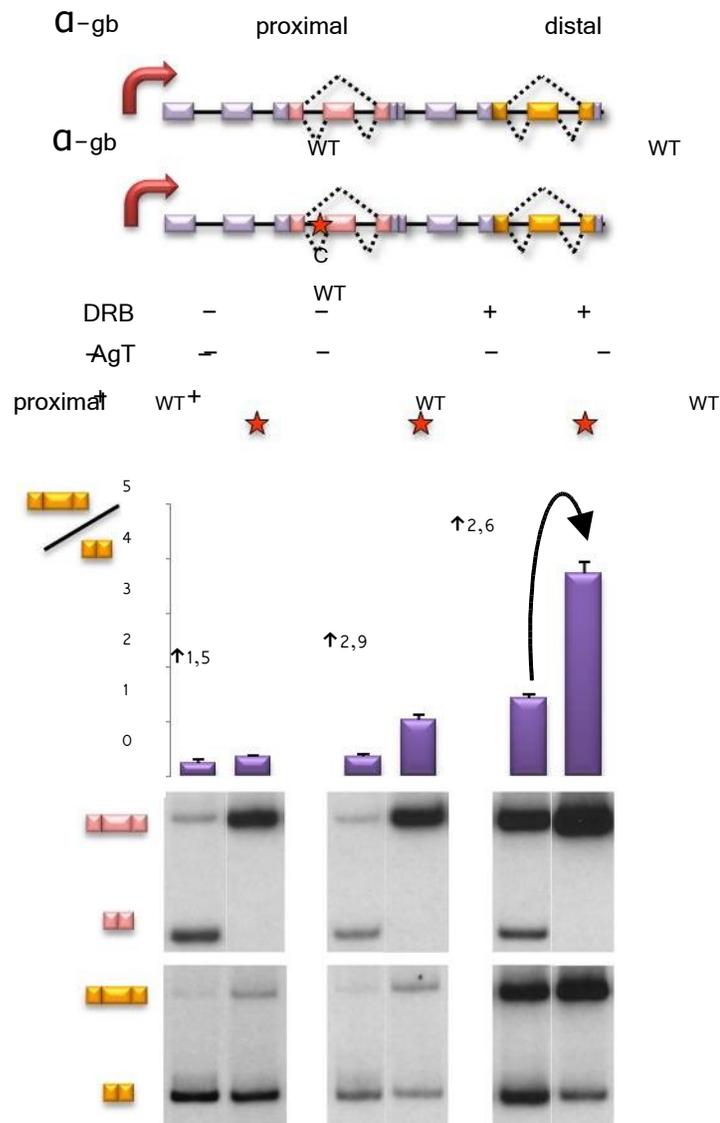


Fig. r9. La coordinación está ligada a la elongación de la polimerasa II. La coordinación es recuperada cuando se trata a las células transfectadas con DRB. Células Hep3B transfectadas con los minigenes tándem bajo el control del promotor de α -gb cuyo exón EDI proximal era alternativo ($p\alpha$ -gb-EDI^{WT}/EDI^{WT}) o constitutivo ($p\alpha$ -gb-EDI^C/EDI^{WT}). Los pocillos tratados con DRB fueron expuestos a una concentración de 50 μ M durante 24 hs., 24 hs. después de la transfección. En todos los casos el *splicing* alternativo fue analizado de la misma manera que en la figura 2 (Procedimientos Experimentales). Los datos representan el promedio de 2 transfecciones independientes de 3 pocillos de 35 mm. para cada condición. Las barras de error y los valores \pm representan el desvío estándar.

El efecto de coordinación es específico de exón

Para investigar la especificidad de exón del efecto de coordinación, transfectamos células Hep3B con minigenes tándem bajo el control del promotor de FN, que contenían exones EDI WT o constitutivos en la región de *splicing* alternativo proximal y al exón alternativo EDII en la región alternativa distal. La figura r10 nos

muestra que cuando el exón EDI es alternativo (pFN-EDI^{WT}/EDII^{WT}), el cociente de inclusión para el exón EDII es 0.32 ± 0.05 . Cuando el exón proximal EDI es constitutivo (pFN-EDI^C/EDII^{WT}), los niveles de inclusión de EDII permanecen no afectados (cociente 0.28 ± 0.04), evidenciando que el *splicing* de EDII (al contrario de EDI) no responde a cambios en la inclusión de una región alternativa EDI proximal.

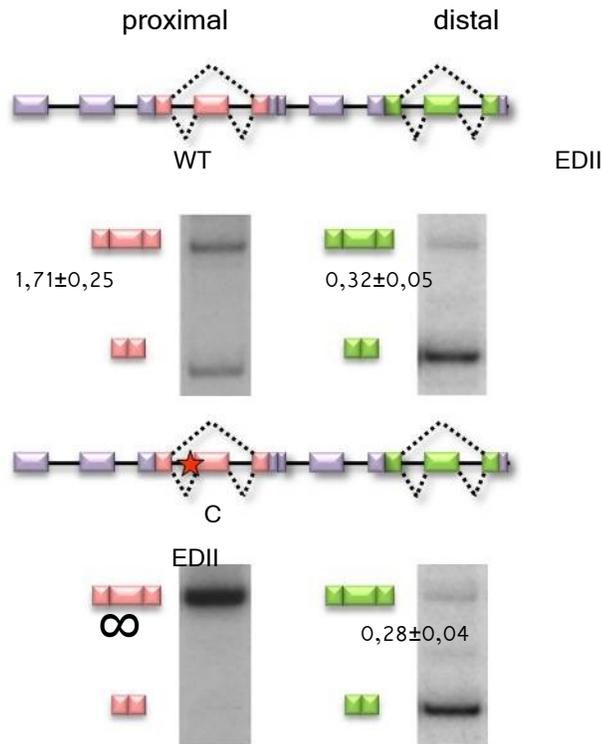


Fig. r10. El efecto de coordinación es exón específico. La inclusión del exón EDII distal no es afectada por el nivel de inclusión de la región alternativa EDI proximal. Células Hep3B fueron transfectadas con 500 ng. de minigenes tándem conteniendo una región alternativa EDII distal WT, bajo el control del promotor de FN humana, con una región alternativa EDI WT (pFN-EDI^{WT}/EDII) o constitutiva (pFN-EDI^C/EDII). Las transfecciones fueron suplementadas con 300 ng. de un vector de expresión de la proteína SRp40, para aumentar y hacer visible la inclusión del exón EDII (Du et al., 1997). Los datos representan el promedio de 2 transfecciones independientes, cada una conteniendo duplicados de pocilloS de 35 mm para cada condición. Los valores \pm representan el desvío estándar.

Aunque ambos, EDI y EDII, son *cassetes* exónicos de similar longitud (270 y 273 nt. respectivamente) y codifican repeticiones de FN tipo III similares, los determinantes en *cis* y los factores *trans* que actúan regulando la inclusión de los mismos son diferentes (Muro et al., 1999). Para explorar más detalladamente la especificidad de exón, se realizaron construcciones en donde la región alternativa murina de FN IIICS fue colocada río abajo de la región EDI, lo cual refleja las posiciones relativas en el gen endógeno de la FN (ver Fig. r12). La región alternativa IIICS murina tiene 3 sitios de *splicing* 3' alternativos y su procesamiento puede dar 3 isoformas: IIICS 120 cuando es utilizado el primer sitio 3' de *splicing* alternativo, IIICS 95 con el segundo y IIICS 0 con el tercero. Cuando el exón proximal EDI es alternativo (pFN-EDI^{WT}/IIICS), el cociente de isoformas IIICS120/IIICS0 es de 4.14 ± 0.76 . Cuando el exón EDI

proximal es constitutivo (pFN-EDI^C/IIICS), hay un pequeño pero significativo descenso en el cociente IIICS120/IIICS0 a 2.97 ± 0.21 (Fig. r11).

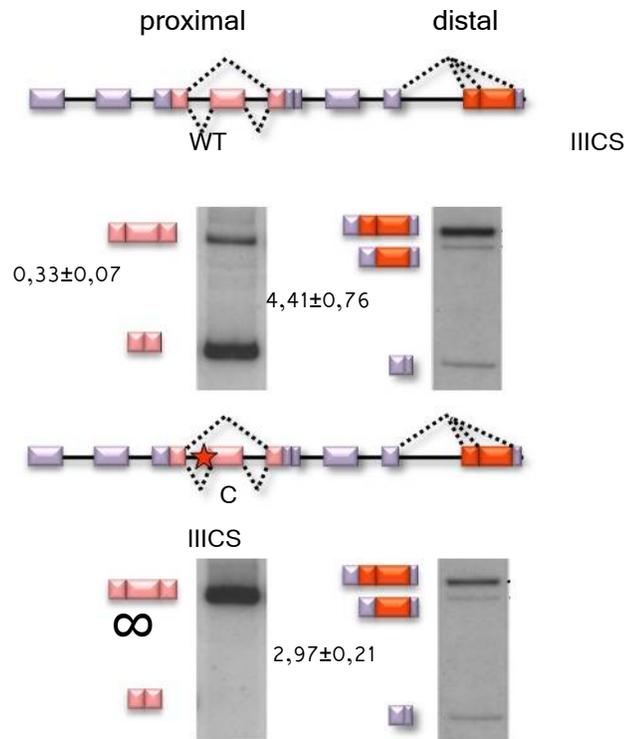


Fig. r11. La inclusión del exón IIICS distal es afectada por el nivel de inclusión de la región alternativa EDI proximal. Células Hep3B fueron transfectadas con 500 ng. de minigenes tándem conteniendo una región alternativa IIICS distal WT, bajo el control del promotor de FN humana, con una región alternativa EDI WT (pFN-EDI^{WT}/IIICS^{WT}) o constitutiva (pFN-EDI^C/IIICS^{WT}). Los datos representan el promedio de 3 transfecciones independientes, cada una conteniendo duplicados de pocillos de 35 mm. para cada condición. Los valores \pm representan el desvío estándar.

Coordinación de dos regiones de *splicing* alternativo en un gen endógeno

En el gen de la fibronectina, EDI es la región de *splicing* alternativo más cercana, río arriba, de IIICS (Fig. r12A). La región IIICS de ratón resulta en tres isoformas alternativas de mRNA generadas a partir de un sitio de *splicing* 5' y tres sitios de *splicing* 3' alternativos. Como ya se mencionó, estas isoformas son IIICS120, IIICS95 y IIICS0, nombradas de acuerdo al número de aminoácidos codificados por los segmentos proteicos. Para validar los resultados obtenidos en un contexto más fisiológico, se investigó la coordinación putativa entre las regiones EDI y IIICS en el gen endógeno de FN de ratón. Para este propósito se utilizaron ratones mutantes generados en el laboratorio del Dr. Andres Muro. Usando *gene targeting*, Muro et al. 2003, desarrollaron una cepa de ratones con expresión constitutiva (EDA^{+/+}) o nula (EDA^{-/-}) del exón EDI. En el gen endógeno, las regiones EDI y IIICS están separadas por aproximadamente 5400 bp,

abarcando seis exones constitutivos y sus intrones. Se investigó el *splicing* alternativo de la región IIICS por RT-PCR en mRNAs de fibroblastos embrionicos (MEFs) en cultivo, los cuáles derivan de los ratones EDA^{+/+} y EDA^{-/-}. En los MEFs EDA^{-/-}, el cociente IIICS120:IIICS0 es casi 5.5 veces más alto que en los MEFs EDA^{+/+} (Fig. r12B). El gen de la FN está ubicado en el cromosoma 1. Como controles, no encontramos cambios entre los genotipos EDA^{+/+} y EDA^{-/-}, en el *splicing* alternativo de genes que se encuentran en otros cromosomas, por ejemplo CD44 (exón 9, cromosoma 2) y FGFR2 (receptor 2 del factor de crecimiento fibroblástico, exones mutuamente excluyentes IIIb y IIIc, cromosoma 7). Por otra parte, la inclusión endógena de EDII es similar en los MEFs EDA^{+/+} y EDA^{-/-} (Fig. r12B). Sin embargo, no se puede interpretar que la falta de cambios en el *splicing* de EDII refleje un efecto polar (EDII está localizado río arriba de EDI en el gen endógeno, Fig. r10B) dado que demostramos que EDII no responde a cambios en los niveles de inclusión de EDI aún cuando éste se situaba río abajo en los minigenes (Fig. r10).

A

fibronectina

EDII
(EDB)
IIICS

EDI
(EDA)

120 aa

95 aa

B

0 aa

EDII
FGFR2

EDI

IIICS

CD 44

Genotipo EDA -/-

EcoRV

- EcoRV

∞

0,31±0,01

0

15,4±3,20

0,12±0,01

Genotipo EDA +/+

- EcoRV

∞

0,37±0,01

∞

2,7±0,29

0,13±0,01

Fig. r12. *Splicing* alternativo coordinado en el gen endógeno de FN murina. (A) Variaciones de la estructura primaria de la FN murina. El esquema superior representa el polipéptido más largo posible, el cual está compuesto de una estructura de dominios repetitivos tipo II y III. Las regiones alternativas EDII (verde), EDI (rosa) y IIICS (rojo) y sus variantes están también representadas. (B) El *splicing* de la región alternativa IIICS es afectado por la inclusión del exón EDI que se encuentra río arriba de la misma. Se muestra el patrón de *splicing* alternativo en las regiones EDI y IIICS del gen endógeno de FN, en células MEFs (*murine embryonic fibroblasts* = fibroblastos embrionicos murinos) de ratones mutantes con expresión nula (EDA^{-/-}) o constitutiva (EDA^{+/+}) del exón EDI (EDA). Los patrones de *splicing* de los genes CD44 (exón alternativo 9) y FGFR2 (exones mutuamente excluyentes IIIb y IIIc), localizados en otros cromosomas, no son afectados por la inclusión de EDI. Las cifras correspondientes a la región alternativa IIICS representan el cociente entre la isoforma 120 y la 0. Los datos representan 18 aislamientos de RNA independientes de pocillos de 35 mm. para cada condición. Los valores ± representan el desvío estándar. Las isoformas del gen FGFR2 se diferenciaron por medio de la digestión de los productos de PCR con EcoRV, que solo corta los productos que contienen el exón IIIc.

Elongación de la RNA polimerasa II y los efectos en el gen endógeno de FN

De manera similar a lo encontrado en los minigenes transfectados (Fig. r9), el tratamiento de los MEFs con el inhibidor de la elongación DRB aumenta el efecto de la configuración proximal de EDI sobre el cociente IIICS120:IIICS0 del gen endógeno (Fig. r13).

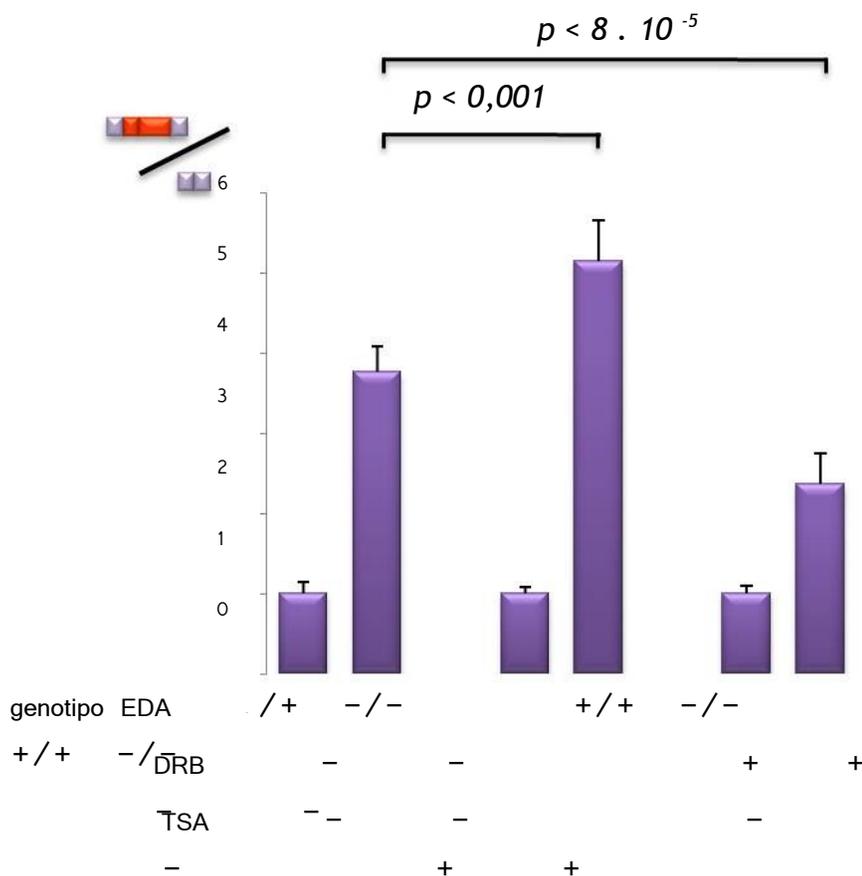


Fig. r13. El efecto polar está ligado a la elongación de la polimerasa II. DRB aumenta y TSA inhibe el efecto polar causado por las configuraciones de la región alternativa EDA sobre el *splicing* de la región IIICS. Células MEFs fueron tratadas con 50 μ M de DRB o 5 ng./ml. de TSA durante 24 hs. antes de la preparación de *RNA* total. El *splicing* alternativo de la región IIICS fue analizado de la misma manera que en la figura r12 (Procedimientos Experimentales). Los datos representan el promedio del cociente de las isoformas de *splicing* 120/0 de la región IIICS de 12 aislamientos de *RNA* independientes de pocillos de 35 mm. para cada condición. Las barras de error representan el desvío standard. En cada par de columnas, los datos están relativizados a la condición del genotipo EDA +/+.

A la inversa, la activación de la elongación mediante la promoción de un incremento en la acetilación de histonas tendría efectos opuestos. La figura r13 muestra que al tratar los MEFs con tricostatina A (TSA), un

potente inhibidor de la desacetilación de histonas, se inhibe el efecto de la inclusión constitutiva de EDI sobre el cociente IIICS120:IIICS0 .

Efectos alelo-específicos

Dado que la FN EDI⁺ tiene actividad de factor de crecimiento (Manabe et al., 1999), las diferencias observadas en el *splicing* de IIICS podrían ser causadas por diferencias en el estado metabólico o en la tasa de crecimiento determinadas por la presencia de FN EDI⁺ en el medio condicionado EDA^{+/-}. Para descartar esta posibilidad, se realizaron mediciones del *splicing* alternativo de la región IIICS en un cultivo de MEFs derivado de un ratón heterocigoto para los alelos constitutivo y nulo de (cepa EDA^{+/-}). La identificación de las bandas alelo específicas de la RT-PCR en la electroforesis en gel de poliacrilamida permitió estimar que el cociente IIICS120:IIICS0 era 3 veces mayor en los mRNAs transcritos del alelo EDA⁻ comparados con los mRNAs transcritos del alelo EDA⁺ (Fig. r14C). Estos resultados confirman las diferencias en IIICS observadas en la figura r11 y, al utilizarse un solo cultivo celular, disipa la posibilidad de que las diferencias observadas sean el resultado de diferentes ambientes celulares. Además, la especificidad alélica indica que la coordinación entre los *splicing* alternativos de EDI y IIICS ocurren en *cis*, lo cual es coherente con la naturaleza polar observada con la transfección de minigenes. La coordinación endógena es consistente con la tendencia observada en transfecciones con minigenes en donde IIICS fue puesto río debajo de EDI (Fig. r11).

A

C

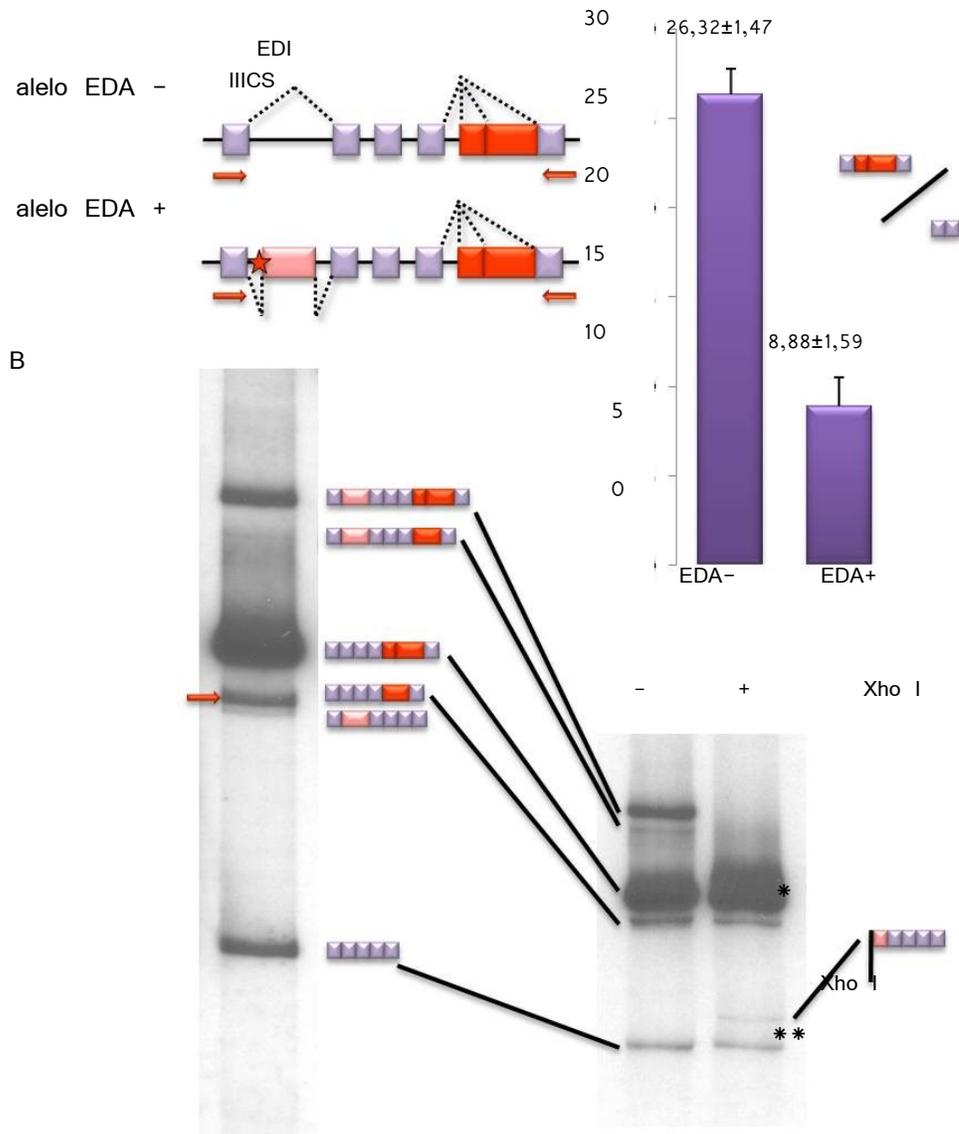


Fig. r14. RT-PCR alelo específica de RNA proveniente de MEFs de ratones mutantes heterocigotas EDI^{+/−} para analizar los patrones de *splicing* alternativo de la región IIICS. (A) Detalle conteniendo las regiones alternativas EDI (EDI) y IIICS de los alelos del gen de FN en los ratones mutantes heterocigotas. (B) En el panel de la derecha los productos de PCR EDI⁺ fueron digeridos con Xho I para generar un fragmento pequeño EDI⁺/IIICS0 distinguible de la variante EDI[−]/IIICS95 (flecha roja en el panel izquierdo). Los cocientes */** son idénticos digiriendo o no el producto de PCR con Xho I (aproximadamente 26,5). (C) Los datos representan el promedio de 24 diferentes pocillos de 35mm. conteniendo células MEFs EDI^{+/−} en 3 pasajes diferentes. Los valores ± representan el desvío estándar.

Parte II - Bioinformática

Parámetros de frecuencia de *splicing* alternativo

Para evaluar la generalidad de nuestros hallazgos, se estudió mediante el análisis de bases de datos la frecuencia de múltiples eventos de *splicing* alternativo en un mismo transcripto y también evidencias de coordinación entre los mismos en colaboración con el grupo del Dr. Gelfand del ITIP, Moscú. El número de regiones alternativas por gen fue determinado por alineamiento con la isoforma más larga de todas, satisfaciendo las condiciones descritas en Procedimientos Experimentales. Las regiones presentes en todas las isoformas fueron consideradas constitutivas, mientras que aquellas que sólo aparecían en una fracción de las isoformas fueron consideradas alternativas.

La fracción de genes que contienen regiones sujetas a *splicing* alternativo (de acuerdo a la definición anterior) en el genoma humano depende de la cobertura de ESTs (*tags* de secuencias expresadas), y se extiende desde un 40% (si sólo las variantes que están apoyadas por datos de secuencias proteicas son consideradas) hasta un 60% si también son considerados los ESTs (Fig. r15A). Aún así, el número estimado de regiones alternativas por gen que sufre *splicing* alternativo (Fig. r15B) permanece similar, independientemente de la estimación del número de genes con *splicing* alternativo: 52%-72% de todos los genes que contienen regiones alternativas contienen sólo una región alternativa, el menor valor proveniente de la estimación por ESTs y el mayor de la estimación teniendo en cuenta secuencias proteicas, con el valor proveniente de mRNAs en una posición intermedia (62%); 20%-29% contienen dos regiones alternativas; 6%-12%, tres regiones y 1%-6% más de tres regiones. Así, aproximadamente el 25% de todos los genes humanos contiene más de una región de *splicing* alternativo, una fracción significativa que justifica el estudio de una posible coordinación.

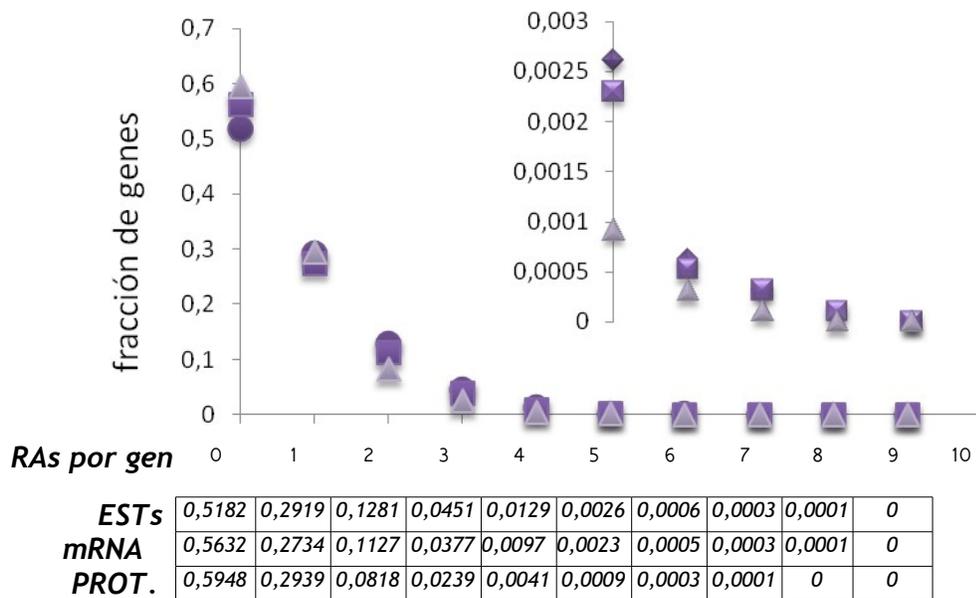
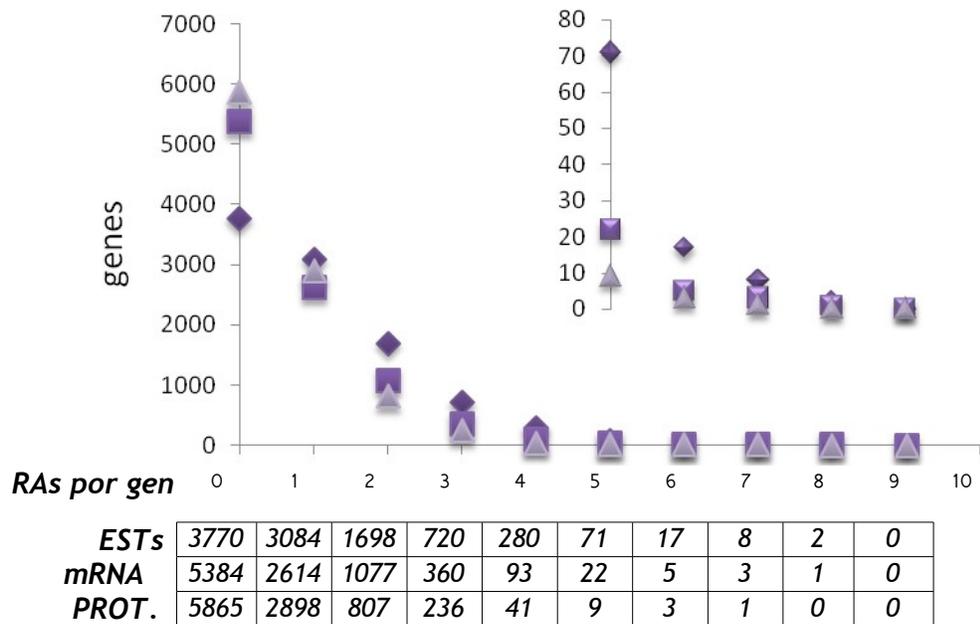


Fig. r15. Estudio y anotación de transcritos con múltiples eventos de *splicing* alternativo. (A) anotación del número de regiones alternativas por gen realizada según las técnicas descritas en Procedimientos Experimentales. (B) fracción de genes humanos conteniendo diferente número de regiones alternativas, entre todos los genes con regiones codificantes de *splicing* alternativo. Rombos, cuadrados y círculos denotan las variantes de análisis apoyadas por bases de datos de ESTs, mRNAs y proteínas respectivamente, RAs: número de regiones alternativas (ver Procedimientos Experimentales).

Coordinación en el genoma humano

Para estudiar la coordinación, se recogieron todos los pares de regiones de *splicing* alternativo adyacentes presentes en la base de datos EDAS (<http://www.belozersky.msu.ru/edas/>). Por cada par, se colectaron todos los ESTs que atravesaban ambas regiones y se compilaron tablas de contingencia en las cuales las filas y columnas se correspondían a variantes de las regiones proximales y distales, respectivamente, y las celdas contenían el número de ESTs *spliceados* según la combinación de variantes. Columnas y filas con menos de 2 ESTs totales fueron ignoradas. Esto resultó en tablas para un total de 630 genes. Se aplicó el test exacto de Fisher (<http://www.unc.edu/wpreacher/fisher/fisher.htm>) a estas tablas en orden de identificar casos de *splicing* alternativo coordinado. La mayoría de las tablas eran estadísticamente no significativas, debido a un número reducido de ESTs cubriendo ambas regiones. En 60 genes las variantes correlacionaban con una significancia $<0,1$, la cual es la esperada para distribuciones al azar. Sin embargo, en muchos casos, las dependencias fueron significativas aun después de la corrección de Bonferroni por testeo múltiple. La figura r16 muestra cinco ejemplos de genes con combinaciones no al azar significativas al nivel de $<0,007$. En tres casos (CNOT8, PCBP2, y BRRN1), la diferencia de longitud entre variantes corresponde a un múltiplo de tres. El exón alternativo que se sitúa río arriba en CNOT8 se encuentra en la región 5' no traducida. Aquellos de PCBP2 y BRRN1 están en la región codificante sin introducir codones *stop* prematuros. En el caso de MGAT4B, la inclusión del exón alternativo río arriba introduce un codón *stop* prematuro en el penúltimo exón del gen, pero a una distancia menor a 50 nt. del sitio 5' de *splicing*. Estas observaciones excluyen la posibilidad de que combinaciones ausentes o representadas en forma baja en cuatro ejemplos fueran en realidad eliminadas por NMD (*non sense mediated mRNA decay*).

CNOT8	a	<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="border-right: 1px solid black; padding: 5px;"></td> <td style="padding: 5px;">c</td> <td style="padding: 5px;">d</td> <td rowspan="3" style="padding-left: 10px; vertical-align: middle;">$p < 0,007$</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">a</td> <td style="padding: 5px;">4</td> <td style="padding: 5px;">1</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">b</td> <td style="padding: 5px;">0</td> <td style="padding: 5px;">8</td> </tr> </table>		c	d	$p < 0,007$	a	4	1	b	0	8
	c	d	$p < 0,007$									
a	4	1										
b	0	8										

<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="border-right: 1px solid black; padding: 5px;"></td> <td style="padding: 5px;">c</td> <td style="padding: 5px;">d</td> <td rowspan="3" style="padding-left: 10px; vertical-align: middle;">$p < 0,003$</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">a</td> <td style="padding: 5px;">5</td> <td style="padding: 5px;">17</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">b</td> <td style="padding: 5px;">5</td> <td style="padding: 5px;">0</td> </tr> </table>		c	d	$p < 0,003$	a	5	17	b	5	0
	c	d	$p < 0,003$							
a	5	17								
b	5	0								

<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="border-right: 1px solid black; padding: 5px;"></td> <td style="padding: 5px;">c</td> <td style="padding: 5px;">d</td> <td rowspan="3" style="padding-left: 10px; vertical-align: middle;">$p < 0,001$</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">a</td> <td style="padding: 5px;">0</td> <td style="padding: 5px;">12</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">b</td> <td style="padding: 5px;">44</td> <td style="padding: 5px;">48</td> </tr> </table>		c	d	$p < 0,001$	a	0	12	b	44	48
	c	d	$p < 0,001$							
a	0	12								
b	44	48								

<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="border-right: 1px solid black; padding: 5px;"></td> <td style="padding: 5px;">c</td> <td style="padding: 5px;">d</td> <td rowspan="3" style="padding-left: 10px; vertical-align: middle;">$p < 2,11 \cdot 10^{-9}$</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">a</td> <td style="padding: 5px;">10</td> <td style="padding: 5px;">2</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">b</td> <td style="padding: 5px;">2</td> <td style="padding: 5px;">68</td> </tr> </table>		c	d	$p < 2,11 \cdot 10^{-9}$	a	10	2	b	2	68
	c	d	$p < 2,11 \cdot 10^{-9}$							
a	10	2								
b	2	68								

<table style="border-collapse: collapse; margin: 0 auto;"> <tr> <td style="border-right: 1px solid black; padding: 5px;"></td> <td style="padding: 5px;">c</td> <td style="padding: 5px;">d</td> <td rowspan="3" style="padding-left: 10px; vertical-align: middle;">$p < 2,4 \cdot 10^{-8}$</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">a</td> <td style="padding: 5px;">11</td> <td style="padding: 5px;">2</td> </tr> <tr> <td style="border-right: 1px solid black; padding: 5px;">b</td> <td style="padding: 5px;">0</td> <td style="padding: 5px;">28</td> </tr> </table>		c	d	$p < 2,4 \cdot 10^{-8}$	a	11	2	b	0	28
	c	d	$p < 2,4 \cdot 10^{-8}$							
a	11	2								
b	0	28								

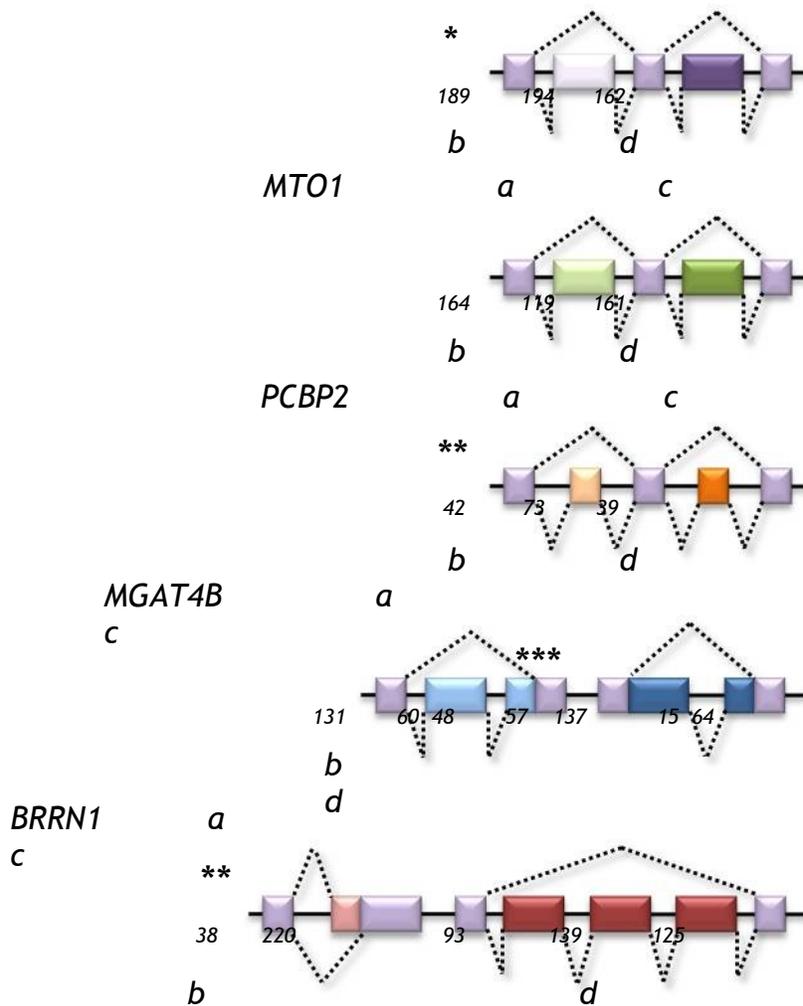


Fig. r16. Evidencia bioinformática de coordinación entre diferentes regiones de *splicing* alternativo dentro de un mismo gen. Ejemplos de genes humanos con dos regiones de *splicing* alternativo consecutivas con dos patrones alternativos para cada región (a o b para el evento río arriba, c o d para la región río abajo). Las tablas muestran el número de ESTs para cada posible combinación encontrada en la base de datos EDAS. * exón localizado en la región 5' no traducida del transcripto. ** Exones localizados en regiones codificantes, que no introducen codones stop prematuros. *** Exones localizados en regiones codificantes que introducen un codón stop en el penúltimo exón, en una posición que no desencadena *NMD*. CNOT8: Subunidad 8 del complejo transcripcional CCR4-NOT (LocusLink ID: 9337); MTO1: homólogo de la proteína de optimización de traducción mitocondrial 1 (*S. cerevisiae*) (LocusLink ID: 25821); PCBP2: proteína de unión al poli(rC) 2 (LocusLink ID: 5094); MGAT4B : manosil (alfa-1,3-)-glicoproteína beta-1,4-Nacetilglucosaminiltransferasa isoenzima B (LocusLink ID: 11282); BRRN1: barren homolog (*Drosophila*) (LocusLink ID: 23397).

Eventos de *splicing* coordinados pueden solo reflejar factores específicos de tipo celular actuando en ambas regiones en cada transcripto. La tabla 1 descarta esta posibilidad dado que muestra que excepto por BRRN1, hay poca especificidad de tejido en la distribución de ESTs para las diferentes combinaciones.

GEN	Combinación de isoformas	N° de ESTs	Distribución de Tejidos/Órganos
CNOT8	a/c	4	1 músculo, 3 nervioso.
	a/d	1	1 genitales.
	b/c	0	-
	b/d	8	2 sangre y sistema inmune, 1 cerebro, 1 genitales, 1 pulmones, 2 sistema nervioso, 1 piel.
MTO1	a/c	5	1 cerebro, 4 mamas.
	a/d	17	1 sangre y sistema inmune, 1 glándulas mamarias, 1 genitales, 1 <i>pool</i> de glándulas, 1 pulmones, 7 músculo, 1 <i>pool</i> de tejidos, 1 piel, 2 colon, 1 cabeza-cuello.
	b/c	5	5 músculos.
	b/d	0	-
PCBP2	a/c	0	-
	a/d	12	5 tracto gastrointestinal, 1 glándulas, 1 hígado, 2 músculo, 1 placenta, 1 pulmones/macrófagos alveolares, 1 desconocido/ posibles rombomeros anteriores 1-4 (<i>pooleados</i>).
	b/c	44	3 sangre y sistema inmune, 2 glándulas mamarias, 5 ojo, 4 tracto gastrointestinal, 5 genitales, 1 glándulas, 2 gónadas, 1 riñón, 1 hígado, 7 pulmones, 1 sistema nervioso, 4 placenta, 3 pool, 3 cabeza-cuello, 1 corazón-neonatal.
	b/d	48	3 sangre y sistema inmune, 1 cerebro, 1 tejido conectivo, 2 ojo, 7 tracto gastrointestinal, 9 genitales, 2 glándulas, 2 gónadas, 6 hígado, 2 pulmones, 2 músculo, 2 <i>pool</i> , 3 piel, 1 cabeza-cuello, 1 meduloblastoma, 1 glándula mamaria, 1 cabeza-normal, 1 útero-tumor, 1 colon-normal.
MGAT4B	a/c	10	9 cerebro, 1 gónadas.
	a/d	2	1 cerebro, 1 genitales.
	b/c	2	2 tracto gastrointestinal.
	b/d	68	2 cerebro, 1 glándulas mamarias, 9 ojo, 9 tracto gastrointestinal, 8 genitales, 10 glándulas, 2 gónadas, 2 riñón, 1 hígado, 5 pulmones, 2 músculo, 1 sistema nervioso, 1 placenta, 1 pool, 14 piel.
BRRN1	a/c	11	10 cerebro, 1 riñón.
	a/d	2	1 glándula mamaria, 1 ojo.
	b/c	0	-
	b/d	28	27 tracto gastrointestinal, 1 genitales.

Tabla r1. Distribución en tejidos/órganos de los ESTs de la figura 16, con todas las posibles combinaciones de procesamiento de las dos regiones de *splicing* alternativo.

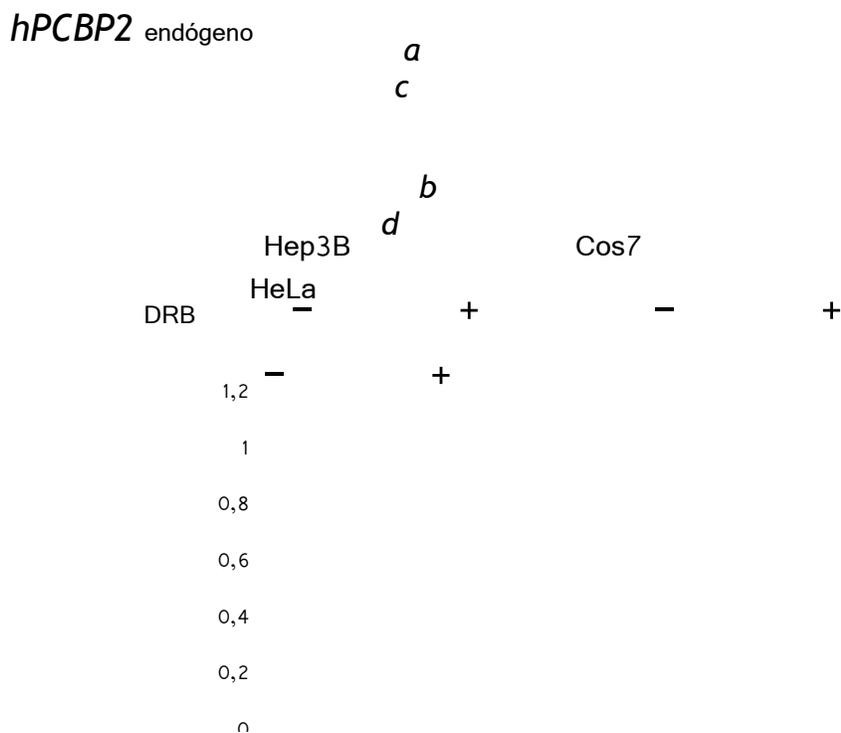
Splicing alternativo coordinado en el gen PCBP2

Se eligió uno de los ejemplos bioinformáticos de la figura r16, la proteína de unión al poli-ribo-C 2 (PCBP2), y se estudió si la distribución de los ESTs correlacionaba con las proporciones relativas de las diferentes isoformas de mRNAs detectadas por RT-PCR en varios tipos celulares y cómo esta distribución era afectada por la elongación de la polimerasa II.

Se encontró una alta correlación entre los datos provenientes de los ESTs y las abundancias relativas de las diferentes isoformas de PCBP2. En las tres líneas celulares ensayadas (Hep3B, Cos7 y HeLa) la isoforma más abundante fue la b-d, seguida por b-c y a-d (Fig. r17). La isoforma que carece de ambos exones alternativos, a-c, es solo detectable en exposiciones de geles muy prolongadas y representa menos del 1% de los mRNAs de PCBP2, lo cual explica en parte la ausencia de ESTs para esta isoforma en particular en las

bases de datos. Los datos de ESTs indican que cuando el exón alternativo que se encuentra río arriba en el par de exones analizados en PCBP2 es excluido, hay un 100% de inclusión del exón alternativo río abajo. Sin embargo, cuando el exón alternativo río arriba es incluido, la inclusión del exón río abajo cae un 52%. Por lo tanto, un incremento en los niveles de inclusión del exón alternativo río arriba (causado, por ejemplo, por una inhibición de la elongación de la polimerasa II) provocaría una caída en la inclusión del exón alternativo río abajo. Esto es observado en las tres líneas celulares estudiadas, en las cuáles el tratamiento con DRB causa una reducción de alrededor de hasta un 60 % en la inclusión del exón alternativo río abajo, dependiendo de la línea celular.

Para investigar si las correlaciones observadas en el gen endógeno eran originadas por un fenómeno de correlación, utilizamos minigenes que contenían la porción con ambos eventos de *splicing* del gen PCBP2, generados por el Dr. Marco Baralle. Cuando analizamos una construcción que contenía una deleción del exón alternativo proximal (12, b), se observó que la inclusión del exón alternativo distal (14, d) aumentaba casi dos veces con respecto a la construcción en la cual el exón b no estaba delecionado (Fig. r18).



*

Fig. r17. El patrón de *splicing* del gen PCBP2 correlaciona con la distribución de ESTs en el análisis bioinformático y es afectado por DRB, un inhibidor de la elongación de la RNA polimerasa II. Las isoformas de PCBP2 fueron detectadas por RT-PCR en 3 líneas celulares diferentes, tratadas o no con DRB. Las bandas de las isoformas b-c y a-d comigran dado que sus longitudes difieren en sólo 3 pb. La presencia de un sitio Hinc II en a-d, pero no en b-c, permite la cuantificación de ambas bandas luego de la digestión de las mismas con dicha enzima. La banda a-c se detecta solamente luego de exposiciones autoradiográficas prolongadas (*). Los datos representan el promedio de 9 preparaciones de mRNA provenientes de 9 cultivos celulares de pocilloS de 35 mm. de cada línea celular para cada tratamiento. Las barras representan el desvío estándar. Los datos están relativizados a la condición de cultivo sin DRB para cada línea celular.

Esto indica que existe coordinación en el *splicing* de los exones 12 y 14 y que la inclusión del exón distal 14 es dependiente negativamente de la inclusión del exón 12.

hPCBP2 minigen

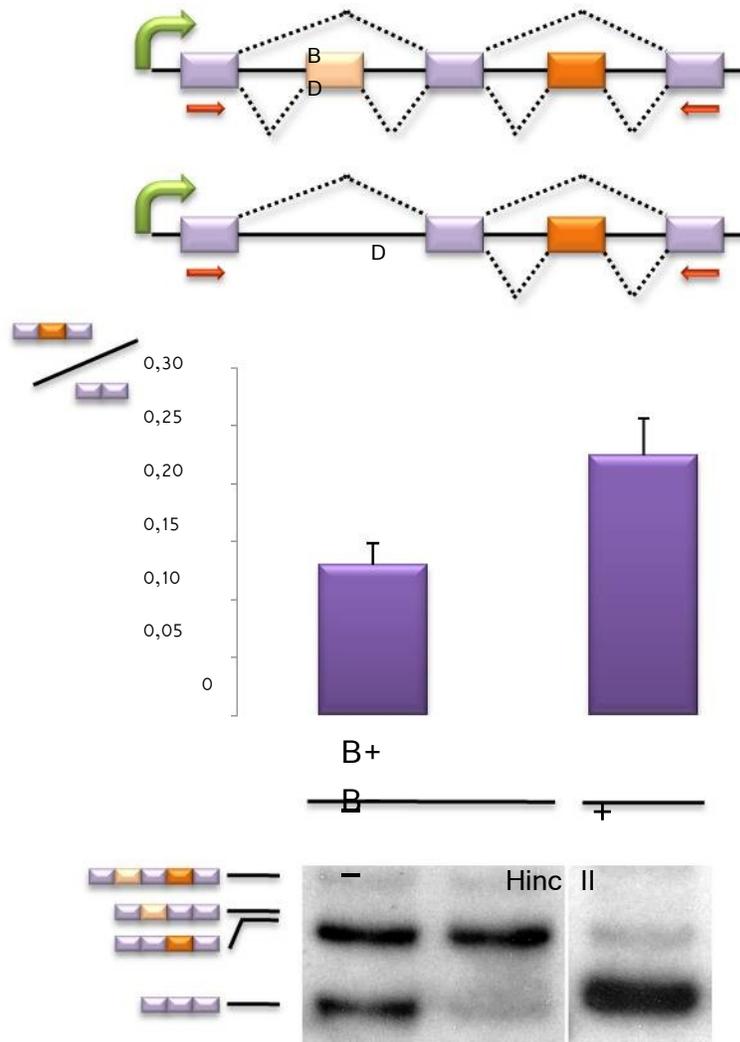


Fig. r18. Correlación en el splicing de los exones 12 (b) y 14 (d) de PCBP2. Se transfirieron construcciones conteniendo los eventos alternativos. Las isoformas fueron detectadas por RT-PCR en 3 líneas celulares diferentes, tratadas o no con DRB. Las bandas de las isoformas conteniendo solo el exón proximal b y no el d y viceversa (no b y si el d) comigran dado que sus longitudes difieren en sólo 3 pb. La presencia de un sitio Hinc II en la segunda, pero no en la primera, permite la cuantificación de ambas bandas luego de la digestión de las mismas con dicha enzima. Los datos representan el promedio de 2 preparaciones de mRNA provenientes de 4 cultivos celulares de pocilloS de 35 mm. para cada construcción en células Hep3B. Las barras representan el desvío estándar.

Conclusiones

Parte I – Coordinación en el gen de la fibronectina

- En minigenes con dos regiones de splicing alternativo del exón EDI de la FN, se observó coordinación en la inclusión de ambos exones alternativos: cuando se realizan mutaciones que aumentan o disminuyen la inclusión del exón alternativo proximal, la inclusión del exón alternativo distal cambia en el mismo sentido.
- Las mismas mutaciones aplicadas sobre la región alternativa distal, tienen un efecto menor sobre el nivel de inclusión del exón alternativo proximal.
- Esta polaridad observada en la magnitud de los efectos tiene el mismo sentido que la transcripción, 5'–3': el procesamiento del exón alternativo proximal influye más sobre la inclusión del exón alternativo distal que viceversa.
- La coordinación no depende de la interacción de proteínas SR al ESE del exón que ejerce el efecto.
- Sin embargo, la coordinación es dependiente de la interacción de proteínas SR al ESE del exón sobre el cual se ejerce el efecto.
- El factor que desencadena el efecto de coordinación no es el nivel de inclusión del exón que ejerce el efecto: en minigenes con regiones proximales que representan los cDNAs de inclusión y exclusión del exón alternativo EDI, éstas no ejercen efecto sobre el procesamiento del exón alternativo distal.
- Cuando los minigenes se transcriben a partir del promotor de α -gb, la polaridad se pierde: ambas regiones alternativas se influyen de manera similar.
- Tratamientos que inhiben la tasa de elongación de la RNA polimerasa II (DRB, AgT) aumentan los efectos de coordinación observados con el promotor de α -gb.
- El efecto de coordinación tiene especificidad de exón: cuando la región alternativa distal corresponde al exón EDII del gen de la FN, el nivel de inclusión de la región alternativa EDI proximal no afecta el nivel de inclusión del exón alternativo distal.
- Asimismo, cuando la región distal corresponde al exón alternativo IIICS del gen de la FN, se observa una leve pero consistente coordinación con la región alternativa EDI proximal.

- De modo similar a lo observado en los minigenes EDI-IIICS, estudiando la coordinación en el gen de la FN en MEFs de ratones con genotipo $EDA^{-/-}$ versus $EDA^{+/+}$, se observa procesamiento diferencial de la región IIICS, río abajo de EDA (EDI): la relación IIICS120/IIICS0 es 6,5 veces mayor en las MEFs $EDA^{-/-}$. El efecto de coordinación es gen específico, dado que los patrones de *splicing* de otros genes no cambian.
- Drogas que inhiben (DRB) o aumentan (TSA) la tasa de elongación de la polimerasa II influyen de forma positiva o negativa, respectivamente, sobre el efecto de coordinación observado en el gen endógeno.
- Las moléculas de mensajero que incluyen al exón EDI, presentan una relación IIICS120/IIICS0 3 veces menor en la región alternativa IIICS, en MEFs de ratones con genotipo $EDA^{+/-}$.

Parte II – Bioinformática y coordinación en el genoma humano

- El 60% de todos los genes humanos contiene al menos una región de *splicing* alternativo.
- El 25% de todos los genes humanos contiene más de una región de *splicing* alternativo.
- Los genes humanos CNOT8, MTO1, PCBP2, MGAT4B y BRRN1 presentan distribuciones de isoformas de *splicing* de dos regiones alternativas que sugieren fuertemente la existencia de coordinación entre las mismas.
- En el gen PCBP2, se confirma que cuando se excluye el exón b (12), la inclusión del exón d (14) está favorecida.
- Minigenes conteniendo los eventos de *splicing* alternativos estudiados del gen PCBP2, en los cuales se deleta el exón b, muestran un aumento de la inclusión del exón d, confirmando que el sistema de *splicing* alternativo de los exones b y d del gen PCBP2 está coordinado.

Discusión

Coordinación

Mediante la expresión de diferentes minigenes con diferentes configuraciones en dos regiones alternativas, se demostró que existe coordinación entre dos regiones de *splicing* alternativo en un mismo gen. En este mecanismo de coordinación están involucradas secuencias específicas del pre-mRNA, que actúan preferencialmente con una polaridad 5'-3': cuando dos regiones alternativas EDI son separadas por 3400 pb, mutaciones que inhiben o estimulan la inclusión del exón EDI proximal correlacionan directamente con los niveles de inclusión del exón distal (Figs. r2-r4), sin embargo, las mismas mutaciones aplicadas a la región alternativa distal tienen efectos mucho menores sobre la región proximal.

El grupo de Andrés Muro (Muro et al., 1998) demostró que la definición exónica gobierna la inclusión del exón EDI, realizando experimentos que muestran que tanto la optimización de los sitios de *splicing* 3' como 5' adyacentes al exón alternativo estimulan notablemente la inclusión. Por otro lado, la actividad del ESE presente en el exón EDI en el contexto de una estructura secundaria particular del RNA es determinante para la inclusión (Buratti and Baralle, 2004). Los resultados de la figura r7 indican que la definición exónica del exón alternativo EDI proximal, y no el hecho de que el mismo esté incluido en el transcripto maduro, estimula la inclusión del exón EDI distal.

¿Cómo puede la definición exónica del exón EDI proximal afectar el procesamiento de una región alternativa que se encuentra río abajo? Los mecanismos de reclutamiento de factores de *splicing* general y específicos necesarios para la definición de ambas regiones alternativas podrían actuar cooperativamente, involucrando factores que actúan de manera similar (en el caso de los minigenes EDI-EDI) o antagónica en regiones alternativas distantes a través de interacciones físicas entre ambas. Recientemente, el grupo de Tina Lenasi encontró un ISE en el intrón 1 del gen de la β -caseína equina que regula la inclusión de los exones alternativos 5 y 8, los cuales se ubican a una distancia mayor a 8 kb. de este *enhancer* intrónico. Los autores demuestran que este efecto de coordinación, mediado por un ISE a distancia, interconecta el uso de todos los sitios de *splicing* débiles en el transcripto, aún cuando utilizan minigenes híbridos en donde el evento alternativo es el exón EDI de FN (Lenasi et al., 2006). El hecho de que un segundo grupo haya encontrado que la inclusión del exón EDI puede estar coordinada a distancia en *Cis* reafirma nuestros hallazgos.

Al investigar la función del ESE del exón EDI proximal sobre el efecto de coordinación en el exón alternativo distal encontramos pruebas de que éste no es necesario para que suceda la misma (Fig. r6c). Más allá de este hecho, quizás otros elementos en *Cis* distribuidos en los intrones lindantes o en el mismo exón EDI, activados en o por la definición exónica, sean los efectores de la coordinación.

Luego de la definición exónica del exón alternativo EDI y el consiguiente *splicing*, EDI queda incluido en el transcripto maduro, depositándose EJC's (*EJC=Exon Junction Complex*) en las junturas con los exones lindantes. Recientemente, el grupo de Klemens Hertel encontró indicios de un rol de la deposición del EJC en el *splicing*: transcriptos que tienen un primer evento inicial de *splicing*, sobre los cuales se depositan EJC's, tienen una remoción de intrones y una definición exónica más eficiente en el resto de los eventos de *splicing* que aquellos que no tienen un evento de *splicing* inicial, sobre los cuales no se depositan EJC's (Crabb et al., presentación en el meeting Eucaryotic RNA Processing, CSH, 2007). Quizás suceda algo similar en nuestro sistema: al definirse e incluirse el exón alternativo EDI proximal en las construcciones tándem, la cantidad (2 EJC's versus 1 en la exclusión) o la calidad de los EJC's depositados pueda variar, teniendo efecto sobre el resto del procesamiento en el transcripto. Por lo tanto, una hipótesis a investigar es si esta deposición diferencial en las construcciones tándem tiene relación con el efecto de coordinación observado sobre el segundo evento alternativo.

Otra hipótesis a seguir es la posibilidad de que el procesamiento del evento de *splicing* proximal pueda cambiar la elongación de la RNA polimerasa II: Fong et.al. observó que cuando un transcripto contiene sitios de *splicing*, la elongación de la RNA polimerasa II está estimulada (Fong and Zhou, 2001). Si la exclusión del exón alternativo proximal, es decir la utilización de los sitios fuertes de *splicing*, activase la elongación de la RNA polimerasa II, esto podría cambiar el grado de inclusión del evento distal, dado que la inclusión del exón EDI es dependiente de la tasa de elongación de la misma (de la Mata et al., 2003b). Aunque las moléculas involucradas en la coordinación no fueron halladas, se puede concluir que este efecto cooperativo muestra polaridad cuando una elongación lenta de la RNA polimerasa II provoca un retraso en la emergencia de la región alternativa distal, generando tiempo extra para el reclutamiento y la definición exónica en la región alternativa proximal (ver La trampa transcripcional).

La existencia de la coordinación a nivel fisiológico fue investigada en dos eventos de *splicing* alternativo que ocurren naturalmente en el gen endógeno de la FN murina (Fig. r12-r14) y en el gen PCBP2 (Fig. r17-r18). En el caso del gen de la FN, se observó el patrón de *splicing* de la región IIICS en ratones mutantes deficientes en el *splicing* regulado de la región alternativa EDI. En MEFs provenientes de estos ratones, la inclusión constitutiva del exón EDI correlaciona inversamente con la inclusión del segmento IIICS120. Cabe destacar que el grupo del Dr. Muro estudió el *splicing* coordinado entre las regiones EDI y IIICS en diferentes órganos de los mismos ratones mutantes de los cuales provienen las células MEFs utilizadas en esta tesis (Chauhan et al., 2004), encontrando ausencia de coordinación en corazón, hígado, bazo, pulmones y cerebro de ratones neonatos de 1 día y adultos de 4 meses. Sin embargo, en riñón, aunque no existe coordinación en el ratón adulto, el ratón *EDA^{-/-}* recién nacido mostró un incremento en el cociente

IIICS120/IIICS0 con respecto al ratón EDA^{+/+}, en concordancia con los resultados obtenidos en esta tesis en MEFs provenientes de dichos ratones.

En esta tesis se observaron cambios en el *splicing* de IIICS en los neonatos EDA^{-/-} versus EDA^{+/+}. En las muestras a las cuales tuvimos acceso, se observó un incremento en la relación IIICS120/IIICS0 en riñón, corazón, pulmón y cerebro de los ratones de 1 día EDA^{-/-} versus los EDA^{+/+}, sin observar diferencias en hígado (Fig. d1). Más allá de la discordancia de resultados, la ausencia de coordinación observada en los órganos de ratones adultos y neonatos que no son riñón por parte del Dr. Muro, en comparación con los resultados obtenidos en MEFs, puede dar cuenta de la dependencia del contexto para el fenómeno de coordinación.

¿Cuáles pueden ser las consecuencias de la coordinación en el gen endógeno de la FN? Una regulación positiva de la inclusión IIICS120 en los embriones EDI^{-/-} podría compensar la ausencia de isoformas conteniendo el exón EDI. De hecho, ambos segmentos proteicos procedentes de los exones IIICS 120 y EDI contienen segmentos de unión a la integrina $\alpha 4\beta 1$ (Liao et al., 2002), lo cual significa que la sobreexpresión de uno de los segmentos alternativos podría compensar el defecto funcional provocado por la ausencia del otro.

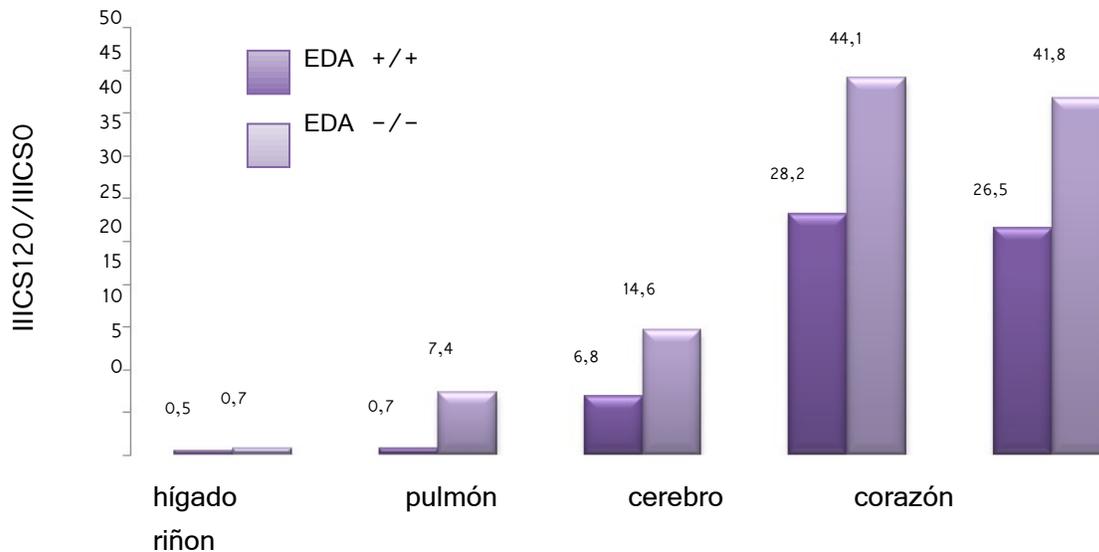


Fig. d1. El ratio IIICS120/IIICS0 del evento IIICS en los transcritos del gen de FN es más alto en los ratones EDA^{-/-} comparados con los EDA^{+/+}. No se detectaron cambios en hígado, pero sí en pulmón, cerebro, corazón y riñón. Las muestras pertenecen a ratones neonatos (1 día de vida).

Esto es consistente con el hecho de que los ratones EDI^{-/-} son viables, aunque tienen una expectativa de vida menor, cicatrización anormal (Muro et al., 2003), coordinación motora y capacidad exploratoria vertical reducidas (Chauhan et al., 2005). Siguiendo esta hipótesis, sería interesante generar y observar la viabilidad de ratones con genotipo EDA^{-/-}-IIICS⁰ vs. EDA^{+/+}-IIICS⁰. Si el fenotipo de los ratones EDA^{-/-}-IIICS⁰ fuese embrionario letal, como lo observado en los *knock-outs* de fibronectina (George et al., 1993), se demostraría la complementariedad de los segmentos EDI y IIICS 120 en la embriogénesis, en donde la interacción de la FN y la integrina $\alpha_4\beta_1$ es necesaria, por ejemplo, para la correcta migración de las células progenitoras del epicardio y por ende para el desarrollo del sistema circulatorio (Sengbusch et al., 2002). Esta función correlaciona con el hecho de que la inclusión de los exones EDI y IIICS 120 es particularmente alta en todos los tejidos del embrión (Peters and Hynes, 1996).

La evidencia bioinformática presentada en esta tesis indica que en el 40% de los genes con *splicing* alternativo, este proceso se lleva a cabo en múltiples regiones (Figs. 12A -12B). Esto significa que un cuarto de los genes humanos tienen más de una región de *splicing* alternativo. Nuevas estimaciones indican que existen en promedio 3,9 isoformas de *splicing* por gen humano (Stamm et al., 2006).

El análisis *in silico* también proporcionó evidencias de coordinación entre dos regiones en un mismo gen. Por ejemplo, en el gen BRRN1 (Fig. r15) las dos variantes en la región río arriba difieren en la inclusión de un segmento exónico de 33 bases como consecuencia de la utilización de un sitio de *splicing* 5' en común en combinación con dos sitios de *splicing* 3' alternativos. La inclusión del segmento de 33 bases (variante a) está fuertemente coordinada con la exclusión de los 3 exones alternativos que se encuentran en la región alternativa inmediatamente río abajo (variante c): en 11 de los 13 ESTs en los cuales se observa la variante a, también se observa la variante c en la región distal. De forma inversa, la exclusión del segmento de 33 bases (variante b) parece ser el factor determinante de la inclusión de los tres exones alternativos río abajo (variante d): todos los ESTs que presentan la variante b en el evento proximal, presentan la variante d en el distal. El hecho de que la longitud del segmento diferencial proximal (33 bases) es un múltiplo de 3, y que su inclusión no introduce ningún codón stop prematuro, descarta la posibilidad de una degradación preferencial por NMD como explicación para la virtual ausencia de isoformas con combinaciones "a-d" y "b-c". Consideraciones similares se pueden aplicar al análisis del procesamiento de los transcritos del gen PCBP2. En el caso de CNOT8, el NMD (Chang et al., 2007) puede descartarse también, dado que el evento alternativo está ubicado en la región 5' no traducida.

Es importante destacar que se buscaron en las bases de datos disponibles secuencias de ESTs que abarcaran las regiones EDI y IIICS del gen FN. Desafortunadamente, y probablemente debido a la gran distancia entre ambas regiones, no se encontraron ESTs que simultáneamente abarcaran ambas regiones.

En cuanto al mecanismo de coordinación gatillado por el exón EDI, cabe resaltar que la longitud del exón EDI (270 bases, múltiplo de 3) imposibilita cualquier implicación del NMD como causa de la coordinación

observada tanto en el gen endógeno como en los minigenes utilizados. Dos aspectos mecanísticos deben ser considerados separadamente o en conjunto: la influencia de un evento de *splicing* sobre el otro y el comportamiento polar. La activación de la inclusión distal del exón EDI no parece ser causada por la unión de proteínas SR al ESE del exón EDI proximal, dado que sobrevive a la disrupción del ESE mientras el exón EDI proximal siga incluyéndose constitutivamente (Fig. r6).

La trampa transcripcional

Se puede especular que la inclusión del exón EDI proximal puede afectar el *splicing* alternativo de la región distal por “sensibilización” del ensamblado del spliceosoma junto con factores de *splicing* específicos, listos para actuar sobre las secuencias exónicas distales del precursor inmediatamente son transcriptas. Este “sensibilizado” puede ocurrir en ambas direcciones, de hecho, es lo que se observa, en una menor medida, cuando un exón distal constitutivo es capaz de incrementar la inclusión de un exón EDI proximal alternativo (Fig. r5). Lo que hace polar al fenómeno de coordinación es el sentido de la transcripción, que va en dirección 5’-3’. Mecanismos similares, pero no idénticos, se observan en el gen de la trombopoyetina humana, en donde la eliminación de un intrón generado por sitios de *splicing* subóptimos flanqueantes del exón 6 parece depender del *splicing* de los intrones constitutivos que se ubican río arriba de dicho evento alternativo (Romano et al., 2001) o en el *splicing* del transcripto del virus del HIV-1, en donde el *splicing* de los intrones río abajo es dependiente de que se hayan procesado los intrones río arriba, y mutaciones en primer sitio 5’ de *splicing* anulan el *splicing* del resto de los intrones (Bohne et al., 2005). El primer caso sugiere la existencia de un efecto de sensibilización de la maquinaria de *splicing* por un intrón constitutivo río arriba que permite el reconocimiento de sitios subóptimos río abajo.

Los siguientes descubrimientos argumentan fuertemente a favor de que la polaridad del efecto de coordinación descrito en esta tesis es la consecuencia de la co-transcripcionalidad del *splicing*, y que está acoplado a la elongación de la RNA pol II: 1) La polaridad 5’-3’ es coincidente con la polaridad de la transcripción, 2) La especificidad de los efectos observados en células MEFs provenientes de ratones heterocigotas EDA+/- es consistente con un mecanismo de coordinación en *cis*, 3) El efecto polar observado cuando la transcripción de los minigenes es gobernada por el promotor de FN desaparece cuando la transcripción es llevada a cabo a partir del promotor de α -gb (Fig. r8). Nuestro laboratorio demostró con experimentos de inmunoprecipitación de la cromatina (ChIP) que el promotor de α -gb provoca mayor elongación de la RNA pol II que el de FN (Kadener 2002), 4) Al inhibir la elongación de la RNA pol II en *trans* mediante la utilización de drogas (DRB) o en *cis* por replicación del molde mediante la co-transfección de los minigenes

tándem (promotor de α -gb) con un plásmido de expresión del antígeno T, se revierte en cierta medida la ausencia de coordinación observada en los transcritos provenientes de dichas construcciones, 5) Las drogas DRB y TSA, que tienen efectos opuestos en la elongación de la RNA Pol II, también tienen efectos opuestos en la influencia de la región alternativa EDI sobre la región IIICS en el gen endógeno de fibronectina en MEFs de ratones mutantes.

Lo observado en el exón EDI podría suceder en otros genes: la elongación de la RNA polimerasa II podría estar jugando por lo menos dos roles sobre un evento de *splicing* alternativo, actuando directamente sobre la inclusión del mismo (de la Mata et al., 2003a; Kadener et al., 2002) y aumentando o disminuyendo el *gap* temporal que lo aislara o expondría a la influencia de posibles coordinaciones con eventos río arriba y río abajo del mismo.

Aunque los efectos observados en los MEFs son consistentes con un mecanismo co-transcripcional, otras posibilidades post-transcripcionales no pueden ser descartadas. Por ejemplo, cambios en la conformación de la estructura secundaria de los pre-mRNAs pueden persistir luego de que la transcripción haya culminado y afectar preferencialmente la generación de ciertas combinaciones de isoformas.

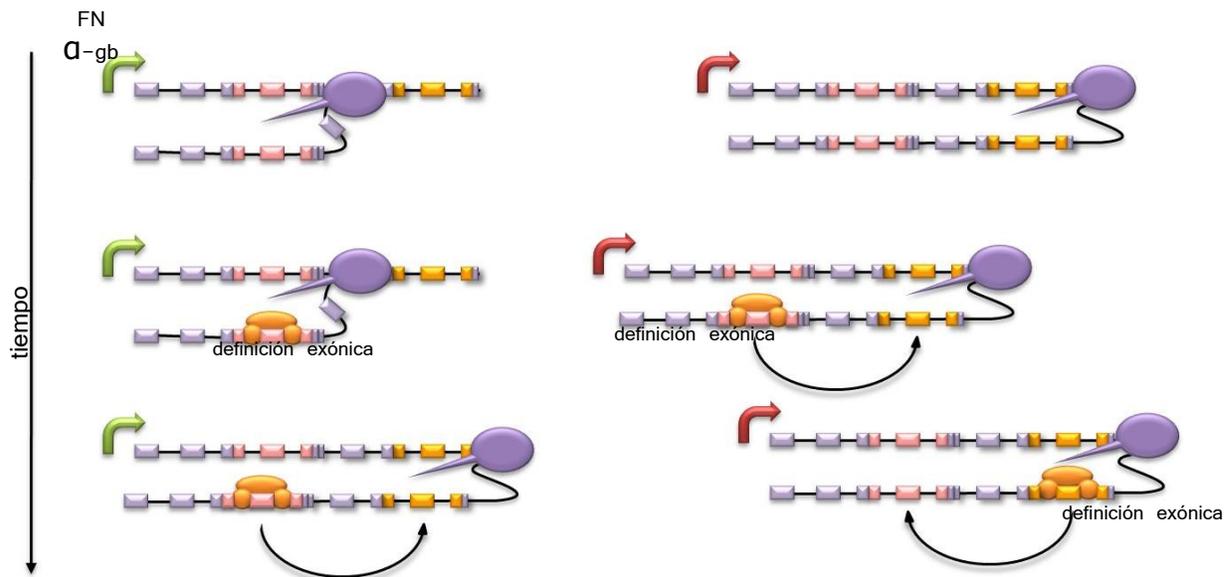


Fig. d2. Modelo del rol de los promotores sobre la polaridad.

Bajas tasas de elongación o pausas internas en el promotor de FN, abren una ventana temporal en donde las chances de que los complejos de splicing se ensamblen en el EDI proximal antes de que el EDI distal se transcriba aumentan. Al seguir la transcripción, los complejos de splicing que actúan en la definición exónica del exón EDI proximal estimulan la inclusión distal de manera polar. La alta procesividad o la falta de pausas internas, en el caso del promotor α -gb, permite que ambos exones alternativos sean expuestos simultáneamente a la maquinaria de splicing, lo cual resulta en la ausencia de polaridad.

Con respecto a las diferencias observadas en las construcciones tándem con promotores de FN versus α -gb, no pueden ser descartados efectos de acoplamiento dependientes del promotor que no estén relacionados con

la tasa de elongación de la RNA polimerasa II, sino con el reclutamiento diferencial de “factores” de coordinación. Podría tratarse de una carga (o descarga) diferencial de factores de *splicing* a la CTD de la RNA polimerasa II dependiente de promotor: Quizás, un primer evento con definición exónica muy eficiente (construcciones con evento constitutivo proximal) hace reclutar factores de regulación positiva de *splicing* a la CTD de la RNA polimerasa II (por ejemplo SF2/ASF) que, luego de actuar sobre este primer evento, mantienen las interacciones con la CTD y pueden cambiar la definición del evento distal. En ambos casos, podemos concluir que la coordinación está acoplada a la transcripción.

Implicancias de la Coordinación

Las observaciones de coordinación reportadas en esta tesis pueden en parte explicar la generación de complejos patrones de *splicing* alternativo en diversos genes.

Por ejemplo, los del gen murino CD44. Este gen tiene 10 exones alternativos (v1-v10) dispuestos en un arreglo central tándem. La combinación libre de estos exones puede producir, en teoría, hasta 1024 isoformas diferentes (2^{10}). Sin embargo, Zhu et al. (Zhu et al., 2003) han demostrado, utilizando una técnica de análisis de moléculas individuales de mRNA llamada perfilado exónico por polony (*polony exon profiling*), que ciertas combinaciones son más comunes que otras, observándose 15 isoformas que dan cuenta del 95% de todos los transcritos del gen CD44 con al menos un exón v. En coincidencia con lo observado en esta tesis, la inclusión de cualquiera de los exones alternativos provoca una cascada de inclusiones de los exones alternativos que se encuentran en una posición río abajo de éstos.

El modelo en la figura d2 muestra que la polaridad es observada con el promotor de FN, el cual provoca una menor elongación de la RNA pol II, mientras que con el promotor de α -gb, que promueve una mayor elongación de polimerasa, no se observa polaridad.

Los resultados presentados sugieren un nuevo rol para el *splicing* alternativo, en donde el valor adaptativo de una secuencia alternativa puede no residir exclusivamente en su potencial de codificar una secuencia proteica sino también en su rol regulatorio del procesamiento de regiones alternativas que se encuentran río abajo. Por lo tanto, dos regiones alternativas podrían funcionar autónomamente bajo ciertas condiciones, pero volverse interdependientes y coordinadas bajo otros estímulos o situaciones celulares, por ejemplo, por la ocupación diferencial de los promotores por diferentes complejos transcripcionales o modificaciones epigenéticas que afecten la configuración de la cromatina y posiblemente, la elongación de la polimerasa II.

La expansión del proteoma mediada por el *splicing* alternativo no sólo emerge por la combinación de dominios modulares individuales en una proteína, sino por efectos cooperativos de dominios alternativos que interactúan funcionalmente como en el caso de los canales de calcio tipo T (Emerick et al., 2006). Este

hecho puede suponer la existencia de una red de co-regulación de diferentes exones alternativos en un mismo gen, que de como resultado proteínas con diferentes combinaciones de dominios funcionalmente dependientes.

Procedimientos Experimentales

Bioinformática

Los datos de *splicing* alternativo del genoma humano provienen de la base de datos EDAS (<http://www.belozersky.msu.ru/edas/>). Los datos de los ESTs fueron adquiridos de la base de datos UniGene (<http://www.ncbi.nlm.nih.gov/UniGene/>). Esta base de datos contiene genes con al menos un intrón en la región codificante y cubiertos por al menos 20 ESTs. Las secuencias proteicas, de mRNAs y de ESTs que mapearon a secuencias genómicas sin intrones en la región codificante fueron ignoradas. Para eliminar artefactos, sólo aquellos exones apoyados por ESTs de por lo menos dos bibliotecas de clones fueron considerados.

Todas las posibles isoformas de *splicing* alternativo generadas fueron anotadas con los datos que pasaron los filtros descritos arriba. Aquellas isoformas que contenían marcos de lectura abiertos cortos (menos de la mitad de la longitud promedio de todas las proteínas descritas por la base de datos RefSeq para ese gen), múltiples codones ATG (más de 2), o intrones localizados a más de 55 nucleótidos río abajo de un codón *stop* también fueron filtradas. Las dos últimas condiciones de filtrado de datos fueron basadas, respectivamente, en el modelo de escaneo ribosomal para la iniciación de la traducción en eucariotas (Kozak, 2000) y en el modelo de vigilancia de transcritos con *splicing* erróneo por *NMD (non sense mediated decay)* (Green et al., 2003; Lewis et al., 2003).

Plásmidos

Reporteros de *splicing* alternativo con promotor de α -gb humana

Se digirió el plásmido pSVEDATot (Caputi et al., 1994b) o pSVEDATotDM (Nogues et al., 2003) con las enzimas SspI/BssHII, se purificó el fragmento de 7000 pb., se rellenaron los extremos 5' protruyentes con el fragmento Klenow de la RNA polimerasa I y se utilizaron como vectores para introducir los fragmentos rellenados de 4300 pb. provenientes de la digestión con SspI/BglI de los plásmidos pSVEDATot, pSVEDATot Δ 2E or pSVEDATotDM,

obteniendo las siguientes construcciones tandem: p α -gb-EDI^{WT}/EDI^{WT}, p α -gb-

EDI ^{Δ ESE}/EDI^{WT}, p α -gb-EDI^C/EDI^{WT}, y p α -gb-EDI^{WT}/EDI^C.

Reporteros de *splicing* alternativo bajo el control del promotor de FN humana

Se digirieron las construcciones p**Q**-gb-EDI^{WT}/EDI^{WT} y p**Q**-gb-EDI^{ΔESE}/EDI^{WT} con Xho I y se purificaron los fragmentos de 3738 bp para ser introducidos en el sitio Xho I de los plásmidos pSVEDATot/FNmut, pSVEDATot/FNmut^{Δ2E} (Cramer et al., 1999) o pSVEDATot/FNmutDM, para generar las construcciones tandem pFN-EDI^{WT}/EDI^{WT}, pFN-EDI^{ΔESE}/EDI^{WT}, pFN-EDI^{WT}/EDI^{ΔESE}, pFNEDI^{ΔESE}/EDI^{ΔESE}, pFN-EDI^C/EDI^{WT}, pFN-EDI^{CΔESE}/EDI^{WT}.

El fragmento de digestión con EcoR I/Xho I de 2663 pb. proveniente del vector pFN-EDI^C/EDI^{WT} fue clonado en el fragmento de 8600 pb. proveniente de la digestión total con EcoR I y parcial con Xho I de la construcción pFN-EDI^{WT}/EDI^{ΔESE} obteniendo el minigen pFN-EDI^C/EDI^{ΔESE}.

Los fragmentos de 3738 pb. provenientes de la digestión con Apa I de las construcciones pFN-EDI^{WT}/EDI^{WT} o pFN-EDI^C/EDI^{WT} fueron introducidos en el sitio Apa I del minigen pSVEDBXho (Muro et al., 1999) o pSVED**Q**1W/mIIICS (ver abajo), generando cuatro construcciones intermedias.

Los fragmentos de 7028 pb. de la digestión con Sfi I/Bcl II de las construcciones intermedias derivadas del vector pSVEDBXho, fueron utilizadas como vector para introducir los fragmentos de 3950 pb. provenientes de la digestión con Sfi I/Bcl II de pSVEDA/FNmut, obteniendo los minigenes tandem pFN-EDI^{WT}/EDI^{WT} and pFN-EDI^C/EDI^{WT}. Los fragmentos de 7575 pb. provenientes de la digestión con BamH I/Xho I de los dos intermediarios derivados de la construcción pSVED**Q**1W/mIIICS fueron utilizados como vectores para introducir el fragmento de 3303 pb. de la digestión con BamH I/Xho I del minigen pSVEDA/FNmut. Esto resultó en las construcciones tandem pFN-EDI^{WT}/IIICS y pFN-EDI^C/IIICS.

Construcción pSVED**Q**1W/mIIICS

Se amplificó por PCR un fragmento que atraviesa todo el exón alternativo murino IIICS, los intrones y parte de los exones flanqueantes con Pfu RNA polimerasa (Stratagene) utilizando los primers genIIICS-Idir (5'-AACCGAGTACACCATCTACGTC-3') y genIIICS-Irev (5'-TCAGAGTCGCACTGGTAGAAG-3'). El mismo se clonó en el sitio romo BstEII de pSVED**Q**1W (Higgs et al., 1983) para obtener pSVED**Q**1W/mIIICS.

Mutantes de delección en la figura r7

El fragmento de 176 pb. proveniente de la digestión con Pst I del producto de PCR de inclusión distal EDI fue clonado en el fragmento de 4260 pb. proveniente de la digestión con Pst I del vector pSVEDA/FNmut, generando el intermediario 1+.

El fragmento proveniente de la digestión con BssH II/Xho I del intermediario 1+, de 724 pb. , fue insertado en el fragmento de 5726 pb. proveniente de la digestión con BssH II/Xho I del vector pSVEDA/FNmut para obtener el intermediario 2+.

El producto de PCR de inclusión distal de EDI de 682 pb. fue pulido con T4 RNA polimerasa y clonado en el sitio EcoR V del vector pBluescript KS+, generando el intermediario 1b+.

El vector pFN-EDI^{WT}/EDI^{WT} se digirió con EcoR V y el fragmento de 2589 pb. fue clonado en el sitio EcoR V del intermediario 1b+ para obtener la construcción intermedia 2b+.

El intermediario 2b+ se digirió con Xho I y el fragmento resultante de 2252 pb. fue insertado en el sitio Xho I del intermediario 2+, generando pFN-cRNAEDI⁺/EDI^{WT}.

El fragmento de 6763 pb. proveniente de la digestión con Apa I del vector pSVEDATot fue rellenado por el fragmento Klenow de la RNA polimerasa I y religado para generar el intermediario 1-.

El fragmento como de 325 pb. proveniente de la digestión con Pst I del producto de PCR de exclusión distal fue clonado en el sitio Pst I/Dra I del fragmento -vector de 3065 pb. proveniente del intermediario 1-, generando el intermediario 2-.

Se digirió la construcción pFN-EDI^{WT}/EDI^{WT} con BstEII y el fragmento de 1976 pb. fue clonado en el sitio BstEII del intermediario 2-, generando el intermediario 3-.

El intermediario 3- se digirió con ApaI y el fragmento resultante de 852 pb. fue clonado en el sitio Apa I de pSVEDATot, generando el intermediario 4-.

El fragmento de 1488 pb. BamHI/BssHII proveniente de pSVEDA/FNmut fue ligado con el fragmento BamHI / BssHII de 6949 pb. del intermediario 4- para generar la construcción pFN-cRNAEDI⁻/EDI^{WT}.

Se digirió la construcción pFN-cRNAEDI⁺/EDI^{WT} con Xho I y el fragmento de 2251 pb. obtenido se clonó en el sitio XhoI de pSVEDA/FNmut o pSVEDATot/FNmutDM para obtener los mutantes de delección f y g.

El fragmento de 3738 pb. proveniente de la digestión con XhoI del vector pFN-EDI^{WT}/EDI^{WT} fue ligado con el fragmento de 6450 pb. proveniente de la digestión con XhoI de la construcción pFN-cRNAEDI⁺/EDI^{WT} para obtener el mutante de delección h.

El mutante de delección j fue obtenido cambiando el promotor de la construcción pSVEDA/CMV short (Kadener et.al., 2001).

Se amplificó por PCR con polimerasa Pfu (Stratagene) un fragmento que contenía todo el exón EDI con los primers EDIA (5'-GCTCTAGAACATTGATCGCCCTAAAGGAC-3') y EDIB (5'-GCTCTAGACTGTGGACTGGGTTCCAATC-3'), luego se digirió con Xba I y se clonó en el sitio XbaI del mutante de delección j para obtener el mutante k.

Construcciones de *splicing* alternativo del gen PCBP2

La zona que contiene los eventos de *splicing* alternativo B y D (exones 11-15) del gen PCBP2 humano se amplificó por PCR con los siguientes primers: PCBP2FORBAMH1 (5'-CGCGGATCCGATGGA CAGGTACAGCACAGGCAGC-3') y PVBP2REXHO (5'-CCGCTCGAGCTTAATCTTTGGAATGG TGAGTTCATGAGAA-3'). El fragmento amplificado fue digerido con BamHI y XhoI, y clonado en el vector pcDNA 3.1 utilizando dichos sitios de restricción. La delección del exón 12 se realizó por PCR en el laboratorio del Dr. Marco Baralle.

Vectores de expresión de proteínas

El plásmido de expresión para SF2/ASF (g10SF2/ASF) fue descrito en reportes previos (Cramer et al., 1999a), como también el de AgT (Kadener et al., 2001).

Transformación bacteriana, minipreps, maxipreps

Transformación de bacterias

Para generar bacterias competentes se utilizó la cepa DH5 α de *Escherichia coli*, y un protocolo tradicional basado en CaCl₂ y otras sales. Con este protocolo, se generan bacterias con una eficiencia de aproximadamente 10⁶ UFC/ μ g. Estas bacterias fueron transformadas según métodos convencionales basados en el protocolo de Hanahan (Hanahan, 1983).

Preparación de plásmidos a pequeña escala (minipreps)

Se utiliza esta técnica para producir una cantidad pequeña de plásmido a partir de un cultivo bacteriano de 3 ml con el objetivo de verificar mediante digestiones por restricción o secuenciación la identidad de los diferentes clones plasmídicos generados. No se detalla esta técnica debido a que es realizada en cualquier laboratorio de biología molecular.

Preparación de plásmidos a gran escala (maxiprep)

Las columnas de intercambio aniónico de Qiagen permiten obtener unos 100 µg de RNA plasmídico (columnas tip 100) o 500 µg (columnas tip 500). Se realiza el protocolo siguiendo las recomendaciones de los fabricantes.

Cuantificación de los plásmidos

Para conocer la concentración aproximada, los plásmidos fueron cuantificados en todos los casos con un espectrofotómetro (*GeneQuant*) que mide la absorción a 260 nm (ultravioleta). Luego, las concentraciones fueron ajustadas corriendo alícuotas digeridas con la enzima EcoRI (se encuentra en el esqueleto de todas las construcciones tándem) en un gel de agarosa con bromuro de etidio y comparando con estándares de masa.

Líneas celulares

Mantenimiento de las líneas celulares

Para las transfecciones transitorias realizadas en esta tesis se utilizaron células de la línea, Hep3B. Esta línea celular es mantenida en placas de plástico de 10 cm de diámetro con medio completo, en un incubador para células con temperatura controlada (37 °C) y en presencia de 5% de CO₂. El medio completo consiste en D-MEM baja o alta glucosa para las células Hep3B o MEFs respectivamente (Invitrogen), al cual se le agrega 10% v/v de suero fetal bovino, antibióticos (100 U/ml final de penicilina y 100 µg/ml final de estreptomina). Ocasionalmente, si se utiliza el medio completo con posterioridad a las dos semanas de preparación, se agrega también L-glutamina en una concentración final de 2,9 g/ml. Antes de que las células alcancen la confluencia total, se aspira el medio de cultivo y se coloca 1 ml de tripsina 2,5 % p/v. Luego de incubar a temperatura ambiente unos dos o tres minutos, se detiene la reacción con unos mls. de medio completo, se replaquea entre 1/4 y 1/8 del volumen a una nueva placa de plástico de 10 cm y se lleva a 8 ml de volumen final con medio completo. Las células son congeladas en N₂ líquido y descongeladas siguiendo protocolos convencionales.

Células Hep3B

La línea celular de hepatoma humano Hep3B es utilizada desde hace muchos años en nuestro laboratorio. Proviene de un carcinoma hepatocelular y es una línea celular adherente con morfología epitelial (para más información ver <http://www.atcc.org>).

MEFs

Los fibroblastos embrionarios murinos (MEFs) de ratones mutantes con genotipo EDA^{+/+}, EDA^{-/-} y EDA^{+/-} fueron provistos por el laboratorio del Dr. Andres Muro (ICGEB, Trieste, Italia), quien generó dichas líneas celulares.

Transfecciones

Se transfectaron células Hep3B (derivadas de hepatoma) con Lipofectamina (Invitrogen) siguiendo las indicaciones del fabricante. Se plaquean las células en placas de plástico de 6 pocillos (cada pocillo tiene un diámetro aproximado de 3,5 cm) el día anterior a realizar la transfección. En el caso de las células Hep3B, se sembraron 200.000 células por pocillo. La cantidad de plásmido total transfectado se ajustó a 2 µg para una cantidad de 6 µL de lipofectamina.

Para poder medir la eficiencia de transfección, se cotransfecta un plásmido que expresa la proteína β-galactosidasa (β-gal) y se realizan ensayos de actividad de dicha enzima. Se realizaron transfecciones con un pocillo extra para cada condición, para obtener el extracto proteico para la medición de actividad β-gal.

Brevemente, el protocolo de transfección con lipofectamina que utilizamos en nuestro laboratorio es el siguiente:

- Se coloca 2 µg totales de RNA plasmídico en un tubo eppendorf que contiene 100 µl de D-MEM solo (sin suero ni antibióticos).
- Se coloca 6 µl de Lipofectamina en un tubo polipropileno de 5 ml con tapa que contiene 100 µl de D-MEM solo (sin suero ni antibióticos).
- Se transfiere el contenido del tubo eppendorf al tubo con Lipofectamina.
- Se incuba 30 minutos para que se formen los complejos entre el polícatión y el DNA.
- Se lava 2 veces con PBS 1X las células a transfectar para eliminar trazas de suero y antibióticos que inhiben la transfección.
- Se dejan las células en el incubador unas 4 hs aproximadamente para que ocurra la transfección.
- Luego se aspira el medio para eliminar la Lipofectamina, que de lo contrario tiene efectos tóxicos a largo plazo, y se agrega 2 ml de medio completo por pocillo.
- Se dejan las células en el incubador 48 hs, y luego se extrae RNA a partir de las mismas.

Tratamientos

Los tratamientos con DRB y TSA fueron realizados 24 hs. después de transfectar las células, a una concentración final de 50 μM y 5 ng./ml. respectivamente. Las células se levantaron con solución D 24 hs. después del tratamiento

Ensayo de actividad de β -galactosidasa

Para realizar este ensayo, primero se debe realizar un extracto proteico. Para ello, se lavan las células dos veces con PBS 1X, se escurre y se colocan 150 μl de buffer de lisis. Se transfiere el lisado a un tubo eppendorf y se elimina el detrito celular centrifugando 10 min a 12.000 rpm. Se transfieren 50 μl del sobrenadante, donde se encuentra la enzima β -gal a un tubo eppendorf nuevo, y se le agrega 50 μl de buffer β -gal 2X, que contiene ONPG como sustrato. La actividad de la enzima β -gal se evidencia como aparición de color amarillo por acumulación del producto ONP. Se detiene la reacción con 100 μl de Na_2CO_3 1M. Se cuantifica la actividad de la enzima midiendo en un espectrofotómetro a una longitud de onda de 415 nm. Soluciones utilizadas

- PBS 10X: 80 g NaCl, 2 g KCl, 14,4 g Na_2HPO_4 , 2,4 g KH_2PO_4 , H_2O c.s.p. 1 litro.
- Buffer de lisis 1X: 100 mM buffer fosfato de potasio pH 7,8 + 0,2 % Triton X- 100.
- Buffer β -gal 2X: 0,2 M buffer fosfato de sodio, 2 mM MgCl_2 , 0,1 M β -mercaptoetanol, 1,33 mg/ml ONPG.

Análisis del RNA

Preparación de RNA

Para obtener el RNA de las células transfectadas, se utiliza una adaptación del protocolo original de Chomczynski (Chomczynski y Sacchi, 1987). Con este protocolo se pueden obtener aproximadamente 5-10 μg de RNA a partir de un pocillo de 35 mm de diámetro. Cabe aclarar que todos los reactivos utilizados deben estar libres de RNAsas, lo que se logra tratando las soluciones con DEPC, y luego autoclavandolas o utilizando stocks con cuidado.

- Se descarta el medio de las células y se aspira bien para que no queden restos de líquido.

- Se agrega 500 μ l de solución D, que ya contiene β -mercaptoetanol y glucógeno, a cada pocillo con células.
- Se levanta el lisado de células con pipeta y se las transfiere a un tubo eppendorf.
- Se agrega 1/10 de volumen (50 μ l) de AcONa pH 4,2 2M y se mezcla por inversión.
- Se agrega un volumen (500 μ l) de fenol ácido saturado en agua y se agita con un vortex.
- Se agrega 1/5 de volumen (100 μ l) de una mezcla 49:1 de cloroformo:alcohol isoamílico y se agita vigorosamente con un .vortex..
- Se deja 30 min en hielo.
- Se centrifuga 30 min a 10.000 rpm en una centrifuga refrigerada a 4 °C.
- Se toma la fase acuosa (superior) y se la pasar a un nuevo tubo eppendorf.
- Se precipita con un volumen (500 μ l) de isopropanol y se agita con un .vortex..
- Se deja al menos 2 hs a -20°C.
- Se centrifuga 30 min a 10.000 rpm en una centrifuga refrigerada a 4 °C.
- Se vuelca el sobrenadante.
- Se agrega 500 μ l de etanol 80% v/v y se agita con un vortex para lavar bien el precipitado.
- Se deja unos 15 min a temperatura ambiente.
- Se centrifuga 20 minutos a 10.000 rpm a temperatura ambiente para disminuir la cantidad de sales precipitadas.
- Se vuelca el sobrenadante y se aspira bien para que no queden restos de etanol.
- Se seca el pellet al aire.
- Se agrega 20 μ l de agua y se congela a .20 °C o .80°C.
- Para resuspender bien el RNA, se descongela en un baño de agua a 50 °C durante 5 min y se agita vigorosamente con un vortex.

Se puede controlar la integridad del RNA extraído corriéndolo en un gel de agarosa 1% p/v, a un voltaje relativamente alto y por poco tiempo, para disminuir la degradación durante la corrida. En una preparación buena de RNA, deben observarse 3 bandas correspondientes, de mayor a menor tamaño, al rRNA 28S, al rRNA 18S y a los tRNA.

Soluciones utilizadas:

- Solución D o desnaturizante: A 234 ml de H₂O mQ se le agregan 200 g de isotiocianato de guanidinio, 14 ml de citrato de sodio pH7 0,75 M y 21,2 ml de sarcosil 10 %. Antes de usar se agrega 7 μ l de β -mercaptoetanol y 40 μ g de glucógeno por cada ml de solución desnaturizante.

RT-PCR

Con el objetivo de evaluar el *splicing* del mRNA transcripto, se realiza una RT-PCR, es decir, una reacción de transcripción reversa (RT) para producir cRNA a partir del RNA, que servirá a su vez como molde de una reacción posterior de reacción en cadena de la polimerasa (PCR) en la cual se amplificará el fragmento de interés utilizando *primers* específicos.

Transcripción reversa (RT)

Primer para RT oligo dT17: 5'-TTTTTTTTTTTTTTTT-3'

- Se desnaturaliza 5 µl de RNA colocándolo 5 min a 65 °C.
- Se pasa el tubo inmediatamente a hielo para impedir la renaturalización.
- Se agrega 15 µl de .mix. a cada tubo. La .mix. está compuesta de la siguiente manera:
 - 4 µl de buffer de RT 5X.
 - 0,2 µl de DTT 100 mM.
 - 0,24 µl de dNTPs 25 mM (mezcla equimolar de dATP, dCTP, dGTP y dTTP).
 - 0,5 µl de inhibidor de RNasas 40 U/µl.
 - 0,5 µl de oligo dT 100 µM (oligo de 12-18 nucleótidos).
 - 1,5 µl de enzima M-MLV RT 200 U/µl.
- H₂O c.s.p. 15 µl.
- Se deja 10 min a temperatura ambiente para que el oligo dT pueda aparearse a las colas de poliA de mRNAs poliadenilados.
- Se incuba 1h a 35 °C para que ocurra la reacción.
- Se detiene la reacción incubando 5 min a 95 °C.

PCRs del cDNA de minigenes tándem con dos eventos de *splicing* alternativo

Evento Proximal EDI

Primers:

pSV5^j: 5'-CACTGCCTGCTGGTGA CT CGA-3'

3^jNJ1: 5'-ATACTGCCAGCGTGCGGAG-3'

Reacción de PCR radiactiva:

- Se utilizan 4 µl de reacción de RT que contiene el cDNA que será el molde para la reacción de PCR.
- A cada tubo de cRNA se le agregan 46 µl de mezcla que estará compuesta de la siguiente manera:
 - 5 µl de buffer de PCR sin MgCl₂

3 µl de MgCl₂ 25 mM

1,5µl de DMSO (dimetil sulfóxido)

2,5 µl de *primer* forward 20 µM

2,5 µl de *primer* reverso 20 µM

1 µl de dNTPs 10 mM (mezcla equimolar de dATP, dCTP, dGTP y dTTP)

0,3 µl de Taq polimerasa 5 U/µl

1µCi [³²P]dCTP (aprox. 0,1 µl de [³²P]dCTP 10 µCi/µl, act. esp.: 3000 Ci/mmol)

H₂O c.s.p. 46 µl

— Se agrega una gota de aceite mineral a cada tubo.

Programa:

45" a 94 °C

1' a 63 °C 30 ciclos

30" a 72 °C

Evento Distal EDI

Primers:

pSV5'^j: 5'-CACTGCCTGCTGGTGA CT CGA-3'

3'polyA: 5'- CAAAGACCACGGGGGTACGGG-3'

Reacción de PCR radiactiva:

Idem anterior pero con 2 µl de MgCl₂ 25 mM

Programa:

45" a 94 °C

60" a 59 °C 30 ciclos

60" a 72 °C

Evento distal EDII (humano)

Primers:

23-glob: 5'- TTCAAGCTCCTAAGCCACTG-3'

3'polyA: 5'- CAAAGACCACGGGGGTACGGG-3'

Reacción de PCR radiactiva:

Idem evento proximal EDI.

Programa:

30" a 95 °C

60" a 58 °C 30 ciclos

60" a 72 °C

Evento distal IIICS (ratón)

Primers:

5'-IIICSsec: 5'- CCACTGCCTG CTGGTGAAC3'

P1W/IIICS3': 5'- CACAGAAGCCAGGAAGTTGTC-3'

Reacción de PCR radiactiva:

Idem evento proximal EDI.

Programa:

30" a 94 °C

45" a 60 °C 30 ciclos

45" a 72 °C

Ensayos de *splicing* alternativo de genes endógenos

PCBP2 endógeno (humano)

La PCR para amplificar el producto de *splicing* alternativo del gen endógeno PCBP2 consiste en 34 ciclos de 30 s a 94º, 30 s a 56º y 60 s a 72º, y fue realizado con 4 µl de la reacción de RT , con conc. de los siguientes primers 1µM: Pcbp2forBamHI (5' CGCGGATCCGATGGACAGGTACAGCACAGGCAGC-3') y Pcbp2revXhoI (5'- CCGCTCGAGCTTAATCTTTGGAATGGTGAGTTCATGAGAA-3'). La banda de la isoforma b-c fue diferenciada de la banda de la isoforma a-d por digestión con HincII por 4 hs. a 37°C, en condiciones especificadas por NEB.

Regiones alternativas endógenas de ratón

Se realizaron los ensayos de *splicing* alternativo por RT-PCR con los protocolos desarrollados por los siguientes autores:

EDII : Blaustein et al., 2004

EDI : Werbach et al., 1998

IIICS : Chauhan et al., 2004

CD44 : Hebbard et al., 2000

FGFR2 : Hajihosseini et al., 2001

La PCR alelo específica que abarca ambas regiones murinas, EDI y IIICS, consiste en: 30 ciclos de 30 s a 95°C, 60 s a 62°C y 120 s a 72°C en 1.5 mM de MgCl₂, usando los primers FNDir (Werbajh et al., 1998) y IIICSRev (Chauhan et al., 2004).

Observación y cuantificación de productos de PCR radiactiva

Al finalizar la reacción de PCR radiactivas, se procede a separar los productos corriendo en un gel de poliacrilamida 6% nativo. Luego de la corrida, se seca el gel sobre un papel Whatmann y se lo expone a una película con orientadores para generar una autoradiografía. De esta manera, se puede observar las bandas correspondientes al producto de PCR que incluye el exón y al que lo excluye. Para cuantificar estos productos, se orienta la película sobre el gel utilizando los orientadores, se recorta las bandas correspondientes a los productos de interés (con y sin exón alternativo), y se cuenta la radiactividad que poseen utilizando un contador beta. Debido a que utilizamos el isótopo ³²P, que es un emisor beta de alta energía, podemos evitar el uso de líquido de centelleo para amplificar la señal, y contamos la radiactividad con el método Cerenkov, que es en seco. La eficiencia de este método en seco es algo menor que utilizando un líquido de centelleo pero, por otra parte, es más rápido y se evita el manipuleo de tolueno.

Referencias

Ahn, S.H., Kim, M., and Buratowski, S. (2004). Phosphorylation of serine 2 within the RNA polymerase II C-terminal domain couples transcription and 3' end processing. *Mol Cell* **13**, 67-76.

Arndt, K.M., and Kane, C.M. (2003). Running with RNA polymerase: eukaryotic transcript elongation. *Trends Genet* **19**, 543-550.

Bauren, G., and Wieslander, L. (1994). Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell* **76**, 183-192.

Bell, M.V., Cowper, A.E., Lefranc, M.P., Bell, J.I., and Screatton, G.R. (1998). Influence of intron length on alternative splicing of CD44. *Mol Cell Biol* **18**, 5930-5941.

Berget, S.M. (1995). Exon recognition in vertebrate splicing. *J Biol Chem* **270**, 2411-2414.

Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc Natl Acad Sci U S A* **74**, 3171-3175.

Beyer, A.L., and Osheim, Y.N. (1988). Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev* **2**, 754-765.

Black, D.L. (1998). Splicing in the inner ear: a familiar tune, but what are the instruments? *Neuron* **20**, 165-168.

Black, D.L. (2003). Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* **72**, 291-336.

Blanchette, M., and Chabot, B. (1997). A highly stable duplex structure sequesters the 5' splice site region of hnRNP A1 alternative exon 7B. *RNA* **3**, 405-419.

Blencowe, B.J. (2000). Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends Biochem Sci* **25**, 106-110.

Blencowe, B.J., Issner, R., Nickerson, J.A., and Sharp, P.A. (1998). A coactivator of pre-mRNA splicing. *Genes Dev* **12**, 996-1009.

Bohne, J., Wodrich, H., and Krausslich, H.G. (2005). Splicing of human immunodeficiency virus RNA is position-dependent suggesting sequential removal of introns from the 5' end. *Nucleic Acids Res* **33**, 825-837.

Bonafe, N., Gilmore-Hebert, M., Folk, N.L., Azodi, M., Zhou, Y., and Chambers, S.K. (2005). Glyceraldehyde-3-phosphate dehydrogenase binds to the AU-Rich 3' untranslated region of colony-stimulating factor-1 (CSF-1) messenger RNA in human ovarian cancer cells: possible role in CSF-1 posttranscriptional regulation and tumor phenotype. *Cancer Res* **65**, 3762-3771.

Bregman, D.B., Du, L., van der Zee, S., and Warren, S.L. (1995). Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J Cell Biol* **129**, 287-298.

Brody, E., and Abelson, J. (1985). The "spliceosome": yeast pre-messenger RNA associates with a 40S complex in a splicing-dependent reaction. *Science* **228**, 963-967.

Buratti, E., and Baralle, F.E. (2004). Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol Cell Biol* **24**, 10505-10514.

Bushnell, D.A., Westover, K.D., Davis, R.E., and Kornberg, R.D. (2004). Structural basis of transcription: an RNA polymerase II-TFIIB cocrystal at 4.5 Angstroms. *Science* **303**, 983-988.

Caceres, J.F., Misteli, T., Sreaton, G.R., Spector, D.L., and Krainer, A.R. (1997). Role of the modular domains of SR proteins in subnuclear localization and alternative splicing specificity. *J Cell Biol* **138**, 225-238.

Caceres, J.F., Stamm, S., Helfman, D.M., and Krainer, A.R. (1994). Regulation of alternative splicing in vivo by overexpression of antagonistic splicing factors. *Science* **265**, 1706-1709.

Caputi, M., Casari, G., Guenzi, S., Tagliabue, R., Sidoli, A., Melo, C.A., and Baralle, F.E. (1994a). A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. *Nucleic Acids Res.* **22**, 1018-1022.

Caputi, M., Casari, G., Guenzi, S., Tagliabue, R., Sidoli, A., Melo, C.A., and Baralle, F.E. (1994b). A novel bipartite splicing enhancer modulates the differential processing of the human fibronectin EDA exon. *Nucleic Acids Res* **22**, 1018-1022.

Caputi, M., and Zahler, A.M. (2001). Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. *J Biol Chem* **276**, 43850-43859.

Caputi, M., and Zahler, A.M. (2002). SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D. *EMBO J* **21**, 845-855.

Carmel, L., Wolf, Y.I., Rogozin, I.B., and Koonin, E.V. (2007). Three distinct modes of intron dynamics in the evolution of eukaryotes. *Genome Res* **17**, 1034-1044.

Cartegni, L., Chew, S.L., and Krainer, A.R. (2002). Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* **3**, 285-298.

Claverie, J.M. (2001). Gene number. What if there are only 30,000 human genes? *Science* **291**, 1255-1257.

Collins, L., and Penny, D. (2005). Complex spliceosomal organization ancestral to extant eukaryotes. *Mol Biol Evol* **22**, 1053-1066.

Conaway, R.C., Kong, S.E., and Conaway, J.W. (2003). TFIIIS and GreB: two like-minded transcription elongation factors with sticky fingers. *Cell* **114**, 272-274.

Connelly, S., and Manley, J.L. (1988). A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev* **2**, 440-452.

Cooper, T.A. (1999). Defining pre-mRNA cis elements that regulate cell-specific splicing. *Methods Mol Biol* **118**, 391-403.

Cramer, P., Caceres, J.F., Cazalla, D., Kadener, S., Muro, A.F., Baralle, F.E., and Kornblihtt, A.R. (1999a). Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. *Mol. Cell* **4**, 251-258.

Cramer, P., Caceres, J.F., Cazalla, D., Kadener, S., Muro, A.F., Baralle, F.E., and Kornblihtt, A.R. (1999b). Coupling of transcription with alternative splicing: RNA pol II promoters modulate SF2/ASF and 9G8 effects on an exonic splicing enhancer. *Mol Cell* **4**, 251-258.

Cramer, P., Pesce, C.G., Baralle, F.E., and Kornblihtt, A.R. (1997). Functional association between promoter structure and transcript alternative splicing. *Proc Natl Acad Sci U S A* **94**, 11456-11460.

Cramer, P., Srebrow, A., Kadener, S., Werbajh, S., de la Mata, M., Melen, G., Nogues, G., and Kornblihtt, A.R. (2001). Coordination between transcription and pre-mRNA processing. *FEBS Lett.* **498**, 179-182.

Chang, Y.F., Imam, J.S., and Wilkinson, M.F. (2007). The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem* **76**, 51-74.

Chauhan, A.K., Iaconig, A., Baralle, F.E., and Muro, A.F. (2004). Alternative splicing of fibronectin: a mouse model demonstrates the identity of in vitro and in vivo systems and the processing autonomy of regulated exons in adult mice. *Gene* **324**, 55-63.

Chauhan, A.K., Moretti, F.A., Iaconig, A., Baralle, F.E., and Muro, A.F. (2005). Impaired motor coordination in mice lacking the EDA exon of the fibronectin gene. *Behav Brain Res* **161**, 31-38.

Chesnut, J.D., Stephens, J.H., and Dahmus, M.E. (1992). The interaction of RNA polymerase II with the adenovirus-2 major late promoter is precluded by phosphorylation of the C-terminal domain of subunit Ila. *J Biol Chem* **267**, 10500-10506.

Chiara, M.D., Palandjian, L., Feld Kramer, R., and Reed, R. (1997). Evidence that U5 snRNP recognizes the 3' splice site for catalytic step II in mammals. *EMBO J* **16**, 4746-4759.

Chothia, C., and Gerstein, M. (1997). Protein evolution. How far can sequences diverge? *Nature* **385**, 579, 581.

Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977a). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* **12**, 1-8.

Chow, L.T., Roberts, J.M., Lewis, J.B., and Broker, T.R. (1977b). A map of cytoplasmic RNA transcripts from lytic adenovirus type 2, determined by electron microscopy of RNA:DNA hybrids. *Cell* **11**, 819-836.

Das, R., Dufu, K., Romney, B., Feldt, M., Elenko, M., and Reed, R. (2006). Functional coupling of RNAP II transcription to spliceosome assembly. *Genes Dev.* **20**, 1100-1109.

Das, R., Yu, J., Zhang, Z., Gygi, M.P., Krainer, A.R., Gygi, S.P., and Reed, R. (2007). SR proteins function in coupling RNAP II transcription to pre-mRNA splicing. *Mol Cell* **26**, 867-881.

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003a). A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell* **12**, 525-532.

de la Mata, M., Alonso, C.R., Kadener, S., Fededa, J.P., Blaustein, M., Pelisch, F., Cramer, P., Bentley, D., and Kornblihtt, A.R. (2003b). A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell* **12**, 525–532.

de la Mata, M., and Kornblihtt, A.R. (2006). RNA polymerase II C-terminal domain mediates regulation of alternative splicing by SRp20. *Nat Struct Mol Biol* **13**, 973–980.

Derr, L.K., and Strathern, J.N. (1993). A role for reverse transcripts in gene conversion. *Nature* **361**, 170–173.

Du, K., Peng, Y., Greenbaum, L.E., Haber, B.A., and Taub, R. (1997). HRS/SRp40-mediated inclusion of the fibronectin EIIIB exon, a possible cause of increased EIIIB expression in proliferating liver. *Mol Cell Biol* **17**, 4096–4104.

Eldridge, A.G., Li, Y., Sharp, P.A., and Blencowe, B.J. (1999). The SRm160/300 splicing coactivator is required for exon-enhancer function. *Proc Natl Acad Sci U S A* **96**, 6125–6130.

Emerick, M.C., Stein, R., Kunze, R., McNulty, M.M., Regan, M.R., Hanck, D.A., and Agnew, W.S. (2006). Profiling the array of Ca(v)3.1 variants from the human T-type calcium channel gene CACNA1G: alternative structures, developmental expression, and biophysical variations. *Proteins* **64**, 320–342.

Fairbrother, W.G., Yeh, R.F., Sharp, P.A., and Burge, C.B. (2002). Predictive identification of exonic splicing enhancers in human genes. *Science* **297**, 1007–1013.

Fickett, J.W., and Wasserman, W.W. (2000). Discovery and modeling of transcriptional regulatory regions. *Curr Opin Biotechnol* **11**, 19–24.

Fish, R.N., and Kane, C.M. (2002). Promoting elongation with transcript cleavage stimulatory factors. *Biochim Biophys Acta* **1577**, 287–307.

Fong, N., and Bentley, D.L. (2001). Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev.* **15**, 1783–1795.

Fong, Y.W., and Zhou, Q. (2001). Stimulatory effect of splicing factors on transcriptional elongation. *Nature* **414**, 929–933.

Gall, J.G., Bellini, M., Wu, Z., and Murphy, C. (1999). Assembly of the nuclear transcription and processing machinery: Cajal bodies (coiled bodies) and transcriptosomes. *Mol Biol Cell* **10**, 4385-4402.

Ge, H., and Manley, J.L. (1990). A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro. *Cell* **62**, 25-34.

Ge, H., Si, Y., and Wolffe, A.P. (1998). A novel transcriptional coactivator, p52, functionally interacts with the essential splicing factor ASF/SF2. *Mol Cell* **2**, 751-759.

George, E.L., Georges-Labouesse, E.N., Patel-King, R.S., Rayburn, H., and Hynes, R.O. (1993). Defects in mesoderm, neural tube and vascular development in mouse embryos lacking fibronectin. *Development* **119**, 1079-1091.

Gil, G., Smith, J.R., Goldstein, J.L., and Brown, M.S. (1987). Optional exon in the 5'-untranslated region of 3-hydroxy-3-methylglutaryl coenzyme A synthase gene: conserved sequence and splicing pattern in humans and hamsters. *Proc Natl Acad Sci U S A* **84**, 1863-1866.

Gilbert, W. (1978). Why genes in pieces? *Nature* **271**, 501.

Gornemann, J., Kotovic, K.M., Hujer, K., and Neugebauer, K.M. (2005). Cotranscriptional spliceosome assembly occurs in a stepwise fashion and requires the cap binding complex. *Mol Cell* **19**, 53-63.

Grabowski, P.J., and Black, D.L. (2001). Alternative RNA splicing in the nervous system. *Prog Neurobiol* **65**, 289-308.

Grabowski, P.J., and Sharp, P.A. (1986). Affinity chromatography of splicing complexes: U2, U5, and U4 + U6 small nuclear ribonucleoprotein particles in the spliceosome. *Science* **233**, 1294-1299.

Graveley, B.R. (2000). Sorting out the complexity of SR protein functions. *RNA* **6**, 1197-1211.

Graveley, B.R., Hertel, K.J., and Maniatis, T. (2001). The role of U2AF35 and U2AF65 in enhancer-dependent splicing. *RNA* **7**, 806-818.

Green, R.E., Lewis, B.P., Hillman, R.T., Blanchette, M., Lareau, L.F., Garnett, A.T., Rio, D.C., and Brenner, S.E. (2003). Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. *Bioinformatics* **19 Suppl 1**, i118-121.

Gutman, A., and Kornblihtt, A.R. (1987). Identification of a third region of cell-specific alternative splicing in human fibronectin mRNA. *Proc Natl Acad Sci U S A* **84**, 7179-7182.

Hall, S.L., and Padgett, R.A. (1994). Conserved sequences in a class of rare eukaryotic nuclear introns with non-consensus splice sites. *J Mol Biol* **239**, 357-365.

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**, 557-580.

Hanamura, A., Caceres, J.F., Mayeda, A., Franza, B.R., Jr., and Krainer, A.R. (1998). Regulated tissue-specific expression of antagonistic pre-mRNA splicing factors. *RNA* **4**, 430-444.

Harrison, P.M., Kumar, A., Lang, N., Snyder, M., and Gerstein, M. (2002). A question of size: the eukaryotic proteome and the problems in defining it. *Nucleic Acids Res* **30**, 1083-1090.

Hertel, K.J., Lynch, K.W., and Maniatis, T. (1997). Common themes in the function of transcription and splicing enhancers. *Curr Opin Cell Biol* **9**, 350-357.

Hirose, Y., and Manley, J.L. (1997). Creatine phosphate, not ATP, is required for 3' end cleavage of mammalian pre-mRNA in vitro. *J Biol Chem* **272**, 29636-29642.

Hirose, Y., Tacke, R., and Manley, J.L. (1999). Phosphorylated RNA polymerase II stimulates pre-mRNA splicing. *Genes Dev.* **13**, 1234-1239.

Hoffman, B.E., and Grabowski, P.J. (1992). U1 snRNP targets an essential splicing factor, U2AF65, to the 3' splice site by a network of interactions spanning the exon. *Genes Dev* **6**, 2554-2568.

Izban, M.G., and Luse, D.S. (1991). Transcription on nucleosomal templates by RNA polymerase II in vitro: inhibition of elongation with enhancement of sequence-specific pausing. *Genes Dev* **5**, 683-696.

Jamison, S.F., and Garcia-Blanco, M.A. (1992). An ATP-independent U2 small nuclear ribonucleoprotein particle/precursor mRNA complex requires both splice sites and the polypyrimidine tract. *Proc Natl Acad Sci U S A* **89**, 5482-5486.

Jeffares, D.C., Mourier, T., and Penny, D. (2006). The biology of intron gain and loss. *Trends Genet* **22**, 16-22.

Jiang, Y., Liu, M., Spencer, C.A., and Price, D.H. (2004). Involvement of transcription termination factor 2 in mitotic repression of transcription elongation. *Mol Cell* **14**, 375-385.

Kadener, S., Cramer, P., Nogues, G., Cazalla, D., de la Mata, M., Fededa, J.P., Werbajh, S.E., Srebrow, A., and Kornblihtt, A.R. (2001). Antagonistic effects of T-Ag and VP16 reveal a role for RNA pol II elongation on alternative splicing. *EMBO J* **20**, 5759–5768.

Kadener, S., Fededa, J.P., Rosbash, M., and Kornblihtt, A.R. (2002). Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation. *Proc Natl Acad Sci U S A* **99**, 8185–8190.

Kan, J.L., and Green, M.R. (1999). Pre-mRNA splicing of IgM exons M1 and M2 is directed by a juxtaposed splicing enhancer and inhibitor. *Genes Dev* **13**, 462–471.

Kim, E., Magen, A., and Ast, G. (2007). Different levels of alternative splicing among eukaryotes. *Nucleic Acids Res* **35**, 125–131.

Kim, T.H., Barrera, L.O., Zheng, M., Qu, C., Singer, M.A., Richmond, T.A., Wu, Y., Green, R.D., and Ren, B. (2005). A high-resolution map of active promoters in the human genome. *Nature* **436**, 876–880.

Kim, Y.K., and Kim, V.N. (2007). Processing of intronic microRNAs. *EMBO J* **26**, 775–783.

Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Luhrmann, R., Garcia-Blanco, M.A., and Manley, J.L. (1994). Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature* **368**, 119–124.

Konarska, M.M., and Sharp, P.A. (1986). Electrophoretic separation of complexes involved in the splicing of precursors to mRNAs. *Cell* **46**, 845–855.

Kornberg, R. (2007). *The Molecular Basis of Eukaryotic Transcription (Nobel Lecture)*. *Angew Chem Int Ed Engl*.

Kornberg, R.D. (2005). Mediator and the mechanism of transcriptional activation. *Trends Biochem Sci* **30**, 235–239.

Kornblihtt, A.R. (2004). Shortcuts to the end. *Nat Struct Mol Biol* **11**, 1156–1157.

Kornblihtt, A.R., Pesce, C.G., Alonso, C.R., Cramer, P., Srebrow, A., Werbajh, S., and Muro, A.F. (1996). The fibronectin gene as a model for splicing and transcription studies. *FASEB J* **10**, 248–257.

Kornblihtt, A.R., Vibe-Pedersen, K., and Baralle, F.E. (1984). Human fibronectin: cell specific alternative mRNA splicing generates polypeptide chains differing in the number of internal repeats. *Nucleic Acids Res* **12**, 5853-5868.

Kotovic, K.M., Lockshon, D., Boric, L., and Neugebauer, K.M. (2003). Cotranscriptional recruitment of the U1 snRNP to intron-containing genes in yeast. *Mol Cell Biol* **23**, 5768-5779.

Kozak, M. (2000). Do the 5'untranslated domains of human cDNAs challenge the rules for initiation of translation (or is it vice versa)? *Genomics* **70**, 396-406.

Krainer, A.R., Conway, G.C., and Kozak, D. (1990). The essential pre-mRNA splicing factor SF2 influences 5' splice site selection by activating proximal sites. *Cell* **62**, 35-42.

Krainer, A.R., and Maniatis, T. (1985). Multiple factors including the small nuclear ribonucleoproteins U1 and U2 are necessary for pre-mRNA splicing in vitro. *Cell* **42**, 725-736.

Krecic, A.M., and Swanson, M.S. (1999). hnRNP complexes: composition, structure, and function. *Curr Opin Cell Biol* **11**, 363-371.

Ladd, A.N., and Cooper, T.A. (2002). Finding signals that regulate alternative splicing in the post-genomic era. *Genome Biol* **3**, reviews0008.

Lamond, A.I., and Spector, D.L. (2003). Nuclear speckles: a model for nuclear organelles. *Nat Rev Mol Cell Biol* **4**, 605-612.

Lavigueur, A., La Branche, H., Kornblihtt, A.R., and Chabot, B. (1993). A splicing enhancer in the human fibronectin alternate ED1 exon interacts with SR proteins and stimulates U2 snRNP binding. *Genes Dev* **7**, 2405-2417.

LeMaire, M.F., and Thummel, C.S. (1990). Splicing precedes polyadenylation during *Drosophila* E74A transcription. *Mol Cell Biol* **10**, 6059-6063.

Lenasi, T., Peterlin, B.M., and Dovc, P. (2006). Distal regulation of alternative splicing by splicing enhancer in equine beta-casein intron 1. *RNA* **12**, 498-507.

Lerner, M.R., and Steitz, J.A. (1979). Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc Natl Acad Sci U S A* **76**, 5495-5499.

Lewis, B.P., Green, R.E., and Brenner, S.E. (2003). Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. *Proc Natl Acad Sci U S A* **100**, 189-192.

Li, Y., and Blencowe, B.J. (1999). Distinct factor requirements for exonic splicing enhancer function and binding of U2AF to the polypyrimidine tract. *J Biol Chem* **274**, 35074-35079.

Liao, Y.F., Gotwals, P.J., Kotliansky, V.E., Sheppard, D., and Van De Water, L. (2002). The EIIIA segment of fibronectin is a ligand for integrins alpha 9beta 1 and alpha 4beta 1 providing a novel mechanism for regulating cell adhesion by alternative splicing. *J Biol Chem* **277**, 14467-14474.

Listerman, I., Sapra, A.K., and Neugebauer, K.M. (2006). Cotranscriptional coupling of splicing factor recruitment and precursor messenger RNA splicing in mammalian cells. *Nat Struct Mol Biol* **13**, 815-822.

Luo, W., and Bentley, D. (2004). A ribonucleolytic rat torpedo RNA polymerase II. *Cell* **119**, 911-914.

Lyon, M., Rushton, G., Askari, J.A., Humphries, M.J., and Gallagher, J.T. (2000). Elucidation of the structural features of heparan sulfate important for interaction with the Hep-2 domain of fibronectin. *J Biol Chem* **275**, 4599-4606.

Malik, S., and Roeder, R.G. (2005). Dynamic regulation of pol II transcription by the mammalian Mediator complex. *Trends Biochem Sci* **30**, 256-263.

Manabe, R., Oh-e, N., and Sekiguchi, K. (1999). Alternatively spliced EDA segment regulates fibronectin-dependent cell cycle progression and mitogenic signal transduction. *J Biol Chem* **274**, 5919-5924.

Maniatis, T., and Reed, R. (2002). An extensive network of coupling among gene expression machines. *Nature* **416**, 499-506.

Mardon, H.J., Sebastio, G., and Baralle, F.E. (1987). A role for exon sequences in alternative splicing of the human fibronectin gene. *Nucleic Acids Res* **15**, 7725-7733.

Martin, W., and Koonin, E.V. (2006). Introns and the origin of nucleus-cytosol compartmentalization. *Nature* **440**, 41-45.

Mattick, J.S., and Gagen, M.J. (2001). The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Mol Biol Evol* **18**, 1611-1630.

Mayeda, A., Helfman, D.M., and Krainer, A.R. (1993). Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. *Mol Cell Biol* **13**, 2993-3001.

Mayeda, A., Munroe, S.H., Caceres, J.F., and Krainer, A.R. (1994). Function of conserved domains of hnRNP A1 and other hnRNP A/B proteins. *EMBO J* **13**, 5483-5495.

McCracken, S., Fong, N., Rosonina, E., Yankulov, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Shuman, S., and Bentley, D.L. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev* **11**, 3306-3318.

McCullough, A.J., and Berget, S.M. (1997). G triplets located throughout a class of small vertebrate introns enforce intron borders and regulate splice site selection. *Mol Cell Biol* **17**, 4562-4571.

Misteli, T., Caceres, J.F., and Spector, D.L. (1997). The dynamics of a pre-mRNA splicing factor in living cells. *Nature* **387**, 523-527.

Misteli, T., and Spector, D.L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol. Cell* **3**, 697-705.

Moen, P.T., Jr., Johnson, C.V., Byron, M., Shopland, L.S., de la Serna, I.L., Imbalzano, A.N., and Lawrence, J.B. (2004). Repositioning of muscle-specific genes relative to the periphery of SC-35 domains during skeletal myogenesis. *Mol Biol Cell* **15**, 197-206.

Mourier, T., and Jeffares, D.C. (2003). Eukaryotic intron loss. *Science* **300**, 1393.

Muro, A.F., Caputi, M., Pariyarath, R., Pagani, F., Buratti, E., and Baralle, F.E. (1999). Regulation of fibronectin EDA exon alternative splicing: possible role of RNA secondary structure for enhancer display. *Mol Cell Biol* **19**, 2657-2671.

Muro, A.F., Chauhan, A.K., Gajovic, S., Iaconcig, A., Porro, F., Stanta, G., and Baralle, F.E. (2003). Regulated splicing of the fibronectin EDA exon is essential for proper skin wound healing and normal lifespan. *J Cell Biol* **162**, 149-160.

Muro, A.F., Iaconcig, A., and Baralle, F.E. (1998). Regulation of the fibronectin EDA exon alternative splicing. Cooperative role of the exonic enhancer element and the 5' splicing site. *FEBS Lett* **437**, 137-141.

Murzin, A.G., Brenner, S.E., Hubbard, T., and Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**, 536-540.

Neugebauer, K.M., and Roth, M.B. (1997). Transcription units as RNA processing units. *Genes Dev* **11**, 3279-3285.

Newby, M.I., and Greenbaum, N.L. (2002). Sculpting of the spliceosomal branch site recognition motif by a conserved pseudouridine. *Nat Struct Biol* **9**, 958-965.

Nikolov, D.B., Chen, H., Halay, E.D., Usheva, A.A., Hisatake, K., Lee, D.K., Roeder, R.G., and Burley, S.K. (1995). Crystal structure of a TFIIB-TBP-TATA-element ternary complex. *Nature* **377**, 119-128.

Nilsen, T.W. (1996). A parallel spliceosome. *Science* **273**, 1813.

Nogues, G., Munoz, M.J., and Kornblihtt, A.R. (2003). Influence of polymerase II processivity on alternative splicing depends on splice site strength. *J Biol Chem* **278**, 52166-52171.

Norton, P.A., and Hynes, R.O. (1987). Alternative splicing of chicken fibronectin in embryos and in normal and transformed cells. *Mol Cell Biol* **7**, 4297-4307.

Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S., and Sharp, P.A. (1986). Splicing of messenger RNA precursors. *Annu Rev Biochem* **55**, 1119-1150.

Pal, M., and Luse, D.S. (2002). Strong natural pausing by RNA polymerase II within 10 bases of transcription start may result in repeated slippage and reextension of the nascent RNA. *Mol Cell Biol* **22**, 30-40.

Pal, M., and Luse, D.S. (2003). The initiation-elongation transition: lateral mobility of RNA in RNA polymerase II complexes is greatly reduced at +8/+9 and absent by +23. *Proc Natl Acad Sci U S A* **100**, 5700-5705.

Payne, J.M., Laybourn, P.J., and Dahmus, M.E. (1989). The transition of RNA polymerase II from initiation to elongation is associated with phosphorylation of the carboxyl-terminal domain of subunit IIa. *J Biol Chem* **264**, 19621-19629.

Peters, J.H., and Hynes, R.O. (1996). Fibronectin isoform distribution in the mouse. I. The alternatively spliced EIIIB, EIIIA, and V segments show widespread codistribution in the developing mouse embryo. *Cell Adhes Commun* **4**, 103-125.

Phair, R.D., and Misteli, T. (2000). High mobility of proteins in the mammalian cell nucleus. *Nature* **404**, 604–609.

Prasad, J., Colwill, K., Pawson, T., and Manley, J.L. (1999). The protein kinase Clk/Sty directly modulates SR protein activity: both hyper- and hypophosphorylation inhibit splicing. *Mol. Cell. Biol.* **19**, 6991–7000.

Price, D.H. (2000). P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol Cell Biol* **20**, 2629–2634.

Proudfoot, N.J. (1989). How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem Sci* **14**, 105–110.

Proudfoot, N.J., Furger, A., and Dye, M.J. (2002). Integrating mRNA processing with transcription. *Cell* **108**, 501–512.

Quelle, D.E., Zindy, F., Ashmun, R.A., and Sherr, C.J. (1995). Alternative reading frames of the INK4a tumor suppressor gene encode two unrelated proteins capable of inducing cell cycle arrest. *Cell* **83**, 993–1000.

Query, C.C., Moore, M.J., and Sharp, P.A. (1994). Branch nucleophile selection in pre-mRNA splicing: evidence for the bulged duplex model. *Genes Dev* **8**, 587–597.

Robberson, B.L., Cote, G.J., and Berget, S.M. (1990). Exon definition may facilitate splice site selection in RNAs with multiple exons. *Mol. Cell. Biol.* **10**, 84–94.

Roberts, G.C., Gooding, C., Mak, H.Y., Proudfoot, N.J., and Smith, C.W. (1998). Co-transcriptional commitment to alternative splice site selection. *Nucleic Acids Res.* **26**, 5568–5572.

Romano, M., Marcucci, R., and Baralle, F.E. (2001). Splicing of constitutive upstream introns is essential for the recognition of intra-exonic suboptimal splice sites in the thrombopoietin gene. *Nucleic Acids Res* **29**, 886–894.

Rosonina, E., Kaneko, S., and Manley, J.L. (2006). Terminating the transcript: breaking up is hard to do. *Genes Dev* **20**, 1050–1056.

Roy, S.W., and Gilbert, W. (2006). The evolution of spliceosomal introns: patterns, puzzles and progress. *Nat Rev Genet* **7**, 211–221.

Saitoh, N., Spahr, C.S., Patterson, S.D., Bubulya, P., Neuwald, A.F., and Spector, D.L. (2004). Proteomic analysis of interchromatin granule clusters. *Mol Biol Cell* **15**, 3876–3890.

Sakharkar, M.K., Chow, V.T., and Kanguene, P. (2004). Distributions of exons and introns in the human genome. *In Silico Biol* **4**, 387–393.

Saunders, A., Core, L.J., and Lis, J.T. (2006). Breaking barriers to transcription elongation. *Nat Rev Mol Cell Biol* **7**, 557–567.

Sengbusch, J.K., He, W., Pinco, K.A., and Yang, J.T. (2002). Dual functions of $[\alpha]_4[\beta]_1$ integrin in epicardial development: initial migration and long-term attachment. *J Cell Biol* **157**, 873–882.

Seraphin, B., and Rosbash, M. (1989). Identification of functional U1 snRNA-pre-mRNA complexes committed to spliceosome assembly and splicing. *Cell* **59**, 349–358.

Shatkin, A.J., and Manley, J.L. (2000). The ends of the affair: capping and polyadenylation. *Nat Struct Biol* **7**, 838–842.

Shilatifard, A., Conaway, R.C., and Conaway, J.W. (2003). The RNA polymerase II elongation complex. *Annu Rev Biochem* **72**, 693–715.

Shopland, L.S., Johnson, C.V., Byron, M., McNeil, J., and Lawrence, J.B. (2003). Clustering of multiple specific genes and gene-rich R-bands around SC-35 domains: evidence for local euchromatic neighborhoods. *J Cell Biol* **162**, 981–990.

Smith, C.W., and Valcarcel, J. (2000). Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends Biochem. Sci.* **25**, 381–388.

Staley, J.P., and Guthrie, C. (1998). Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**, 315–326.

Stamm, S., Riethoven, J.J., Le Texier, V., Gopalakrishnan, C., Kumanduri, V., Tang, Y., Barbosa-Morais, N.L., and Thanaraj, T.A. (2006). ASD: a bioinformatics resource on alternative splicing. *Nucleic Acids Res* **34**, D46–55.

Stoltzfus, A., Spencer, D.F., Zuker, M., Logsdon, J.M., Jr., and Doolittle, W.F. (1994). Testing the exon theory of genes: the evidence from protein structure. *Science* **265**, 202–207.

Sun, H., and Chasin, L.A. (2000). Multiple splicing defects in an intronic false exon. *Mol Cell Biol* **20**, 6414–6425.

Tacke, R., and Manley, J.L. (1999). Determinants of SR protein specificity. *Curr Opin Cell Biol* **11**, 358–362.

Teigelkamp, S., Newman, A.J., and Beggs, J.D. (1995). Extensive interactions of PRP8 protein with the 5' and 3' splice sites during splicing suggest a role in stabilization of exon alignment by U5 snRNA. *EMBO J* **14**, 2602–2612.

Valadkhan, S., and Manley, J.L. (2001). Splicing-related catalysis by protein-free snRNAs. *Nature* **413**, 701–707.

Valadkhan, S., and Manley, J.L. (2003). Characterization of the catalytic activity of U2 and U6 snRNAs. *RNA* **9**, 892–904.

Vibe-Pedersen, K., Kornblihtt, A.R., and Baralle, F.E. (1984). Expression of a human alpha-globin/fibronectin gene hybrid generates two mRNAs by alternative splicing. *EMBO J* **3**, 2511–2516.

Wang, Z., Hoffmann, H.M., and Grabowski, P.J. (1995). Intrinsic U2AF binding is modulated by exon enhancer signals in parallel with changes in splicing activity. *RNA* **1**, 21–35.

Wetterberg, I., Bauren, G., and Wieslander, L. (1996). The intranuclear site of excision of each intron in Balbiani ring 3 pre-mRNA is influenced by the time remaining to transcription termination and different excision efficiencies for the various introns. *RNA* **2**, 641–651.

Wuarin, J., and Schibler, U. (1994). Physical isolation of nascent RNA chains transcribed by RNA polymerase II: evidence for cotranscriptional splicing. *Mol Cell Biol* **14**, 7219–7225.

Xing, Y., Yu, T., Wu, Y.N., Roy, M., Kim, J., and Lee, C. (2006). An expectation-maximization algorithm for probabilistic reconstructions of full-length isoforms from splice graphs. *Nucleic Acids Res* **34**, 3150–3160.

Yamada, K.M. (2000). Fibronectin peptides in cell migration and wound repair. *J Clin Invest* **105**, 1507–1509.

Zeng, C., and Berget, S.M. (2000). Participation of the C-terminal domain of RNA polymerase II in exon definition during pre-mRNA splicing. *Mol. Cell. Biol.* **20**, 8290–8301.

- Zhang, Z., and Gilmour, D.S. (2006). Pcf11 is a termination factor in *Drosophila* that dismantles the elongation complex by bridging the CTD of RNA polymerase II to the nascent transcript. *Mol Cell* **21**, 65-74.
- Zhang, Z., Klatt, A., Henderson, A.J., and Gilmour, D.S. (2007). Transcription termination factor Pcf11 limits the processivity of Pol II on an HIV provirus to repress gene expression. *Genes Dev* **21**, 1609-1614.
- Zhou, Z., Licklider, L.J., Gygi, S.P., and Reed, R. (2002). Comprehensive proteomic analysis of the human spliceosome. *Nature* **419**, 182-185.
- Zhu, J., Mayeda, A., and Krainer, A.R. (2001). Exon identity established through differential antagonism between exonic splicing silencer-bound hnRNP A1 and enhancer-bound SR proteins. *Mol Cell* **8**, 1351-1361.
- Zhu, J., Shendure, J., Mitra, R.D., and Church, G.M. (2003). Single molecule profiling of alternative pre-mRNA splicing. *Science* **301**, 836-838.
- Zuo, P., and Maniatis, T. (1996). The splicing factor U2AF35 mediates critical protein-protein interactions in constitutive and enhancer-dependent splicing. *Genes Dev* **10**, 1356-1368.