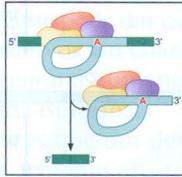


CAPÍTULO 14



Processamento do RNA

A SEQUÊNCIA CODIFICADORA DE UM GENE CODIFICADOR DE PROTEÍNAS é uma série de códons, compostos por três nucleotídeos (trincas), que especifica a sequência linear dos aminoácidos no produto polipeptídico. Até aqui, assumiu-se que a sequência codificadora é contínua: o códon correspondente a um aminoácido está imediatamente adjacente ao códon do aminoácido seguinte na cadeia polipeptídica. Isso é verdadeiro para a grande maioria das bactérias e seus fagos. No entanto, raramente isto é verdade para os genes eucarióticos. Nestes casos, em geral, a sequência codificadora é periodicamente interrompida por segmentos com sequências não codificadoras.

Muitos genes eucarióticos são, portanto, mosaicos compostos por blocos com sequências codificadoras separadas entre si por blocos com sequências não codificadoras. As sequências codificadoras são chamadas **éxons** e as sequências intercaladas, não codificadoras, são os **íntrons**. Ao serem transcritos em uma molécula de RNA, os íntrons devem ser removidos e os éxons, unidos para criar um mRNA para o gene. Na verdade, tecnicamente, o termo *éxon* aplica-se a qualquer região mantida em um RNA maduro, sendo ou não codificadora. Éxons não codificadores incluem as regiões 5' e 3' não traduzidas de um mRNA; todas as porções de RNAs não codificadores estáveis e removidos, como o regulador da inativação do cromossomo X, *Xist* (Cap. 20); e regiões que dão origem a RNAs funcionais, como os micro-RNAs que serão encontrados no Capítulo 20.

A Figura 14-1 apresenta um gene eucariótico típico no qual a região codificadora é interrompida por três íntrons que fragmentam o gene em quatro éxons. O número de íntrons encontrado em um gene varia muito – desde um, no caso da maioria dos genes contendo íntrons em levedura (e alguns poucos genes humanos), a 50, no caso do gene do colágeno de galinha, *pro α 2*, e chegando a 363 no caso do gene humano *Titin*. A Figura 14-2 mostra o número médio de íntrons por gene em uma variedade de organismos. Fica claro que o número médio aumenta à medida que se passa dos eucariotos unicelulares simples, como as leveduras, para organismos mais complexos, como vermes e moscas, até chegar aos seres humanos.

O tamanho dos éxons e íntrons também é variável. Muito frequentemente, os íntrons são bem mais longos do que os éxons que eles separam. Assim, por exemplo, em geral, os éxons possuem cerca de 150 nucleotídeos, enquanto os íntrons – embora também possam ser curtos – podem ter até 800.000 nucleotídeos (800 kb). Como outro exemplo, o gene da enzima di-hidrofolatorredutase de mamíferos tem mais de 31 kb de comprimento, contendo seis éxons dispersos que, juntos, correspondem a 2 kb do mRNA. Assim sendo, neste caso, a região codificadora do gene corresponde a < 10% de seu comprimento total.

SUMÁRIO

Química do Processamento de RNA, 469

Maquinaria do Spliceossomo, 473

Vias de Processamento, 474

Variantes do Processamento, 482

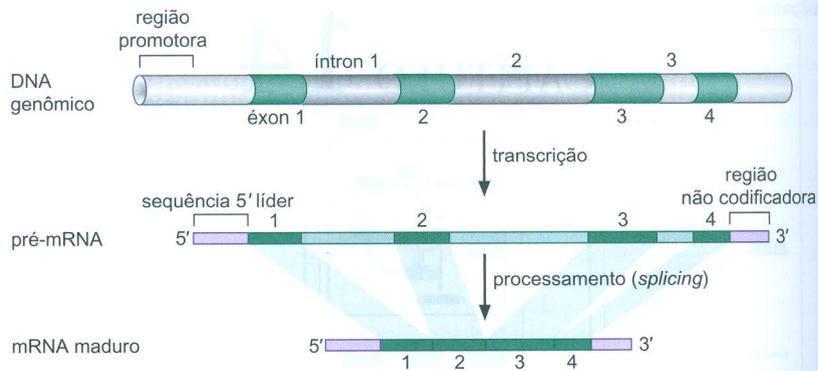
Processamento Alternativo, 483

Embaralhamento de Éxons, 496

Edição de RNA, 500

Transporte de mRNA, 503

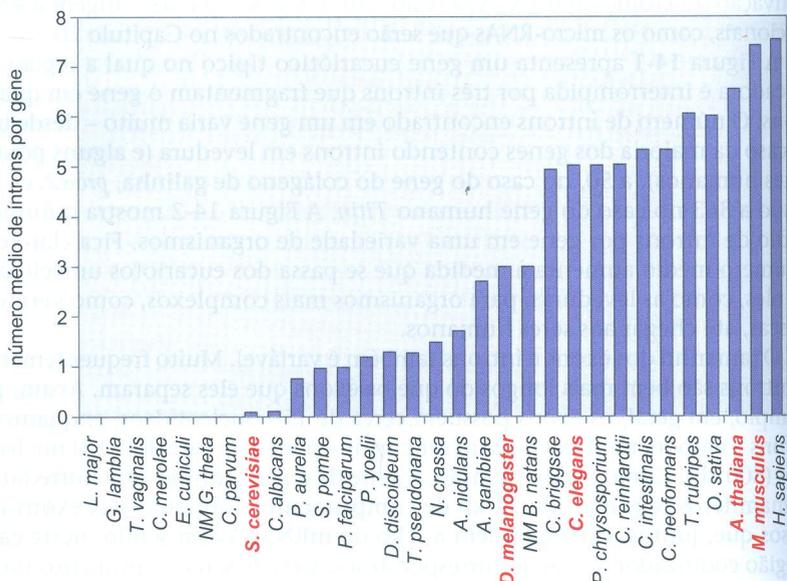
FIGURA 14-1 Um típico gene eucariótico. O gene representado contém quatro éxons separados por três íntrons. A transcrição a partir do promotor gera o pré-mRNA, apresentado na linha intermediária, que contém todos os éxons e íntrons. O processamento remove os íntrons e une os éxons, formando o mRNA maduro, que continua a ser modificado (ver poliadenilação, Cap. 13) e é exportado do núcleo, para ser traduzido em um produto proteico. Tecnicamente, as regiões 5' líder e 3' não codificadoras também são éxons porque elas são mantidas no mRNA maduro. Elas estão aqui representadas em lilás para indicar sua condição de éxons não codificadores.



Da mesma maneira que os genes contínuos dos procariotos, os genes descontínuos dos eucariotos são transcritos em uma cópia de RNA do gene inteiro – o transcrito primário para um gene eucariótico típico contém íntrons e éxons. Isso é mostrado na parte central da Figura 14-1. Devido ao comprimento e ao número de íntrons, o transcrito primário (ou **pré-mRNA**) pode ser, de fato, bastante longo. No caso extremo do gene da *distrofina* humana, a RNA-polimerase precisa percorrer 2.400 kb de DNA para copiar o gene inteiro em RNA. (Considerando que a transcrição avança a uma velocidade de 40 nucleotídeos por segundo, são necessárias 17 horas para sintetizar um único transcrito desse gene.) Isso levanta a possibilidade de que a abundância e o comprimento dos éxons podem ter efeito significativo sobre a taxa de expressão dos genes, assunto que será retomado quando for considerada a regulação gênica durante o desenvolvimento, no Capítulo 21.

Como se afirmou, os transcritos primários dos genes que contêm íntrons devem ser processados para a remoção dos íntrons antes de sua tradução em proteínas. O processo de remoção dos íntrons, chamado **processamento do RNA**, ou simplesmente *splicing*, converte o pré-mRNA em mRNA maduro e precisa ocorrer com grande precisão, evitando a perda ou adição, ainda que de apenas um nucleotídeo, nos sítios de junção de éxons. Como será visto nos Capítulos 15 e 16, os códons de trincas de nucleotídeos do mRNA são traduzidos em uma sequência (ou fase) de leitura constante, determinada pelo primeiro códon da sequência codificadora da proteína. A falta de precisão no processamento – se, por exemplo, uma base for perdida ou adicionada na fronteira entre dois éxons – deixaria as fases de leitura dos éxons fora de or-

FIGURA 14-2 Número de íntrons por gene em várias espécies eucarióticas. O número médio de íntrons por gene é apresentado para uma seleção de espécies eucarióticas. Os nomes em vermelho são os dos organismos experimentais comuns (Apêndice 1): levedura (*Saccharomyces cerevisiae*), mosca-da-fruta (*Drosophila melanogaster*), nematódeo (*Caenorhabditis elegans*), planta (*Arabidopsis thaliana*) e camundongo (*Mus musculus*). As outras espécies apresentadas são: *Anopheles gambiae*; *Aspergillus nidulans*; nucleomorfo de *Bigeloviella natans*; *Caenorhabditis briggsae*; *Candida albicans*; *Chlamydomonas reinhardtii*; *Ciona intestinalis*; *Cryptococcus neoformans*; *Cryptosporidium parvum*; *Cyanidioschyzon merolae*; *Dictyostelium discoideum*; *Encephalitozoon cuniculi*; *Giardia lamblia*; nucleomorfo de *Guillardia theta*; *Homo sapiens*; *Leishmania major*; *Neurospora crassa*; *Oryza sativa*; *Paramecium aurelia*; *Phanerochaete chrysosporium*; *Plasmodium falciparum*; *Plasmodium yoelii*; *Schizosaccharomyces pombe*; *Takifugu rubripes*; *Thalassiosira pseudonana*; e *Trichomonas vaginalis*. (Redesenhada, com permissão, de Roy S.W. e Gilbert W. 2006. *Nat. Rev. Genet.* 7: 212, Fig. 1. © Macmillan.)



dem: os códons a jusante seriam selecionados incorretamente e aminoácidos errados seriam incorporados nas proteínas.

Alguns pré-mRNAs podem ser processados de mais de uma maneira. Assim, mRNAs contendo diferentes grupos de éxons podem ser gerados a partir de um mesmo pré-mRNA. Esse processo é denominado **processamento alternativo** e, por meio dessa estratégia, um gene pode dar origem a mais de um produto polipeptídico. Esses produtos alternativos são chamados de **isoformas**. Estima-se que 90% ou mais dos genes do genoma humano possam sofrer processamentos alternativos, gerando mais de uma isoforma.

O número de formas variantes que um determinado gene pode codificar por esse processo varia de duas a centenas ou até milhares. Por exemplo, o gene *Slo* de ratos, que codifica um canal de potássio expresso nos neurônios, tem potencial para codificar 500 formas alternativas desse produto. Como será visto, existe um gene específico de *Drosophila* que pode codificar 38 mil diferentes produtos em função do processamento alternativo. O processamento alternativo é geralmente um processo regulado, com a produção de diferentes isoformas ocorrendo em resposta a diferentes sinais ou em diferentes tipos celulares.

Neste capítulo, são discutidos não apenas os mecanismos e a regulação do processamento do RNA, mas também hipóteses sobre o porquê dos genes eucarióticos terem regiões codificadoras interrompidas. Também é descrita a edição do RNA, outro modo pelo qual os transcritos primários podem ser alterados para modificar o produto codificado.

O processamento foi descoberto em estudos de expressão gênica de adenovírus de mamíferos, como descrito no Quadro 14-1, Os adenovírus e a descoberta do processamento.

QUÍMICA DO PROCESSAMENTO DE RNA

Sequências no RNA determinam onde ocorre o processamento

Agora, serão considerados os mecanismos moleculares da reação de processamento. Como os íntrons e os éxons são distinguidos entre si? Como os íntrons são removidos? Como os éxons são emendados com alta precisão? As fronteiras entre íntrons e éxons estão marcadas por sequências nucleotídicas específicas presentes nos pré-mRNAs. Essas sequências determinam onde o processamento ocorrerá. Assim, como demonstra a Figura 14-3, a fronteira éxon-íntron – isto é, a fronteira na extremidade 5' de um íntron – é marcada por uma sequência chamada **sítio de processamento 5'** (ou sítio 5' de *splicing*). A fronteira íntron-éxon na extremidade 3' do íntron é marcada como **sítio de processamento 3'** (ou sítio 3' de *splicing*). (Os sítios de processamento 5' e 3' eram algumas vezes referidos, respectivamente, como sítios **doador** e **receptor**, mas atualmente essa nomenclatura é pouco utilizada.)

A Figura 14-3 mostra uma terceira sequência necessária para o processamento. Ela é chamada de **sítio do ponto de ramificação** (ou sequência do

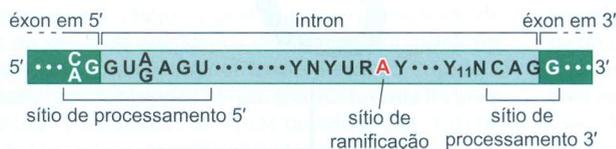


FIGURA 14-3 Sequências das fronteiras entre íntron e éxon. A figura apresenta as sequências consenso dos sítios de processamento 5' e 3' e também o resíduo de A conservado no sítio de ramificação. Como em outros casos de sequências consenso, quando duas bases alternativas são igualmente favorecidas, ambas são indicadas na posição. Nesta figura, as sequências consenso apresentadas são para seres humanos. O mesmo é válido para todas as outras figuras deste capítulo, a menos que conste o contrário.

ponto de ramificação) e localiza-se inteiramente dentro do íntron, geralmente próxima à sua extremidade 3', sendo seguida por uma série de pirimidinas (o trato de polipirimidinas ou trato Py).

A sequência consenso para cada um desses elementos é mostrada na Figura 14-3. As sequências mais bem conservadas são as GU no sítio de processamento 5', as AG no sítio de processamento 3' e o resíduo A no sítio de ramificação. Todos esses nucleotídeos altamente conservados são encontrados no próprio íntron – o que não surpreende, uma vez que as sequências da maioria dos éxons, ao contrário dos íntrons, são limitadas pela necessidade de codificar os aminoácidos específicos de um produto proteico.

À medida que éxons adjacentes são unidos, o íntron é removido na forma de laço

Inicia-se considerando a química do processamento. Um íntron é removido por meio de duas reações sucessivas de **transesterificação**, nas quais as ligações fosfodiéster do pré-mRNA são clivadas, e novas ligações são formadas (Fig. 14-4). A primeira reação é desencadeada pela 2'-OH do A conservado no sítio de ramificação. Este grupo atua como nucleófilo para atacar o grupo fosforil da G conservada no sítio de processamento 5'. (Esta é uma reação S_N2 , que ocorre por meio de um intermediário de fósforo pentavalente.)

Como consequência desta primeira reação, a ligação fosfodiéster entre o açúcar e o fosfato na junção 5' entre o íntron e o éxon é clivada. A extremidade 5' liberada do íntron é unida ao A do sítio de ramificação. Assim, além das ligações 5' e 3' do esqueleto, uma terceira ligação fosfodiéster estende-se da 2'-OH do A, originando uma junção tríplice (por isso, denominada sítio de ramificação). A estrutura da junção tríplice é ilustrada na Figura 14-5.

Observa-se que o grupo que sai na primeira reação de transesterificação é o éxon de 5'. Na segunda reação, o éxon em 5' (mais precisamente, a extremidade 3'-OH recém-liberada do éxon de 5') reverte seu papel e torna-se um nucleófilo que ataca o grupo fosfato no sítio de processamento 3' (Fig. 14-4). Esta segunda reação tem duas consequências. A primeira, e mais importante, é que ela une os éxons de 5' e 3'; portanto, esta é a etapa em que as duas sequências codificadoras são efetivamente "unidas". A segunda é que esta mesma reação libera o íntron, como grupo de saída. Como a extremidade 5' do íntron foi ligada ao A do sítio de ramificação na primeira reação de transesterificação, o íntron recém-liberado tem o formato de um **laço**.

Nessas duas etapas de reação, não há ganho líquido no número de ligações químicas – duas ligações fosfodiéster são rompidas e duas novas são

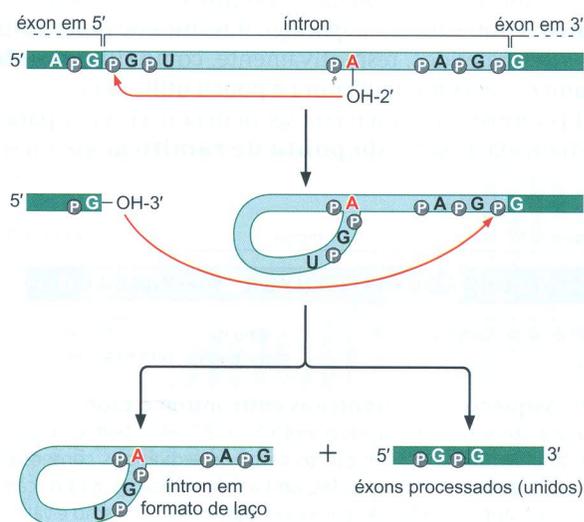


FIGURA 14-4 Reação de processamento. As duas etapas da reação de processamento descritas no texto são mostradas. Na primeira etapa, o RNA forma uma estrutura em alça, apresentada em detalhe na figura seguinte.

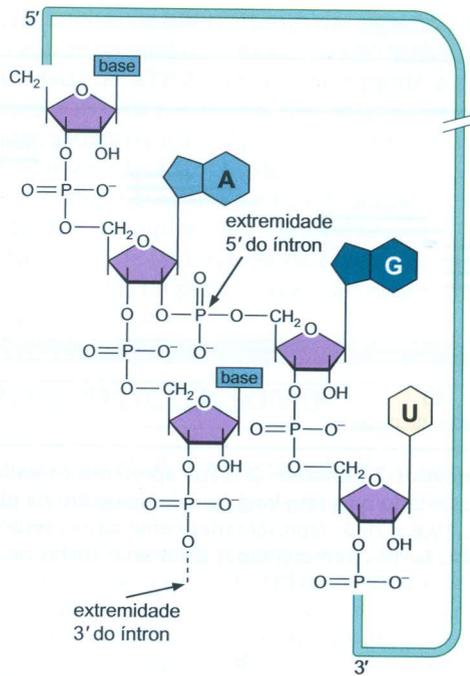


FIGURA 14-5 Estrutura da junção tripla formada durante a reação de processamento.

▶ EXPERIMENTOS - CHAVE

Quadro 14-1 Os adenovírus e a descoberta do processamento

Estudos em bactérias e seus fagos levaram à descoberta de que, com relação à sequência de nucleotídeos, o mRNA é uma réplica exata do gene do qual foi transcrito (ver Cap. 16). Por isso, em 1977, a descoberta de que alguns (e como se sabe hoje, a maioria) mRNAs de eucariotos podem ser clivados e religados, em um trabalho de "corte e colagem", a partir de transcritos primários muito mais longos, causou espanto geral. Como essa surpreendente descoberta foi feita?

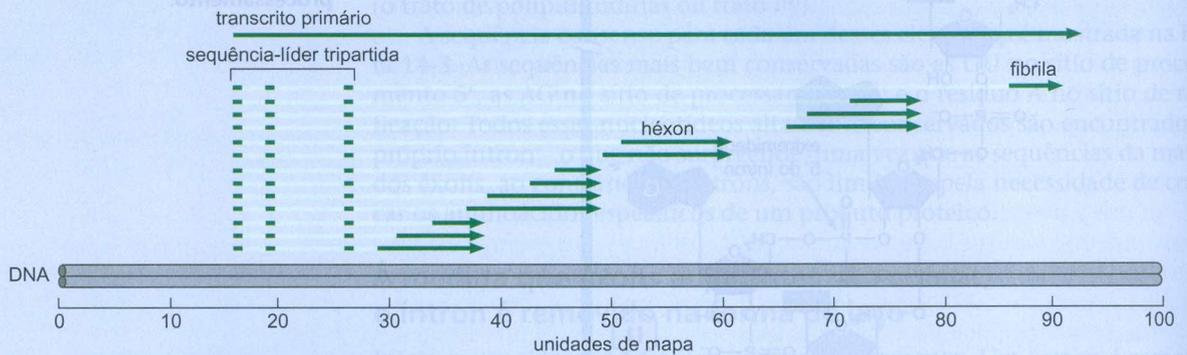
Na tentativa de entender a transcrição gênica em eucariotos, os cientistas detiveram-se em um vírus de DNA humano chamado **adenovírus**. Acreditava-se que o adenovírus pudesse servir como modelo para entender a biologia molecular dos genes eucarióticos, assim como os fagos T4 e λ haviam servido para os genes procarióticos (ver Apêndice 1). O virion do adenovírus é formado por várias proteínas diferentes, codificadas por ele, e os mRNAs para essas proteínas foram purificados, na expectativa de que suas extremidades 5' indicassem precisamente os sítios de início da transcrição de cada gene do genoma viral. Em vez disso, todos os mRNAs, embora codificassem proteínas diferentes, apresentavam a mesma sequência 5'. Hoje se sabe que todos os mRNAs de proteínas do virion do adenovírus derivam de um promotor único, conhecido como promotor tardio principal. O início a partir desse promotor produz longos transcritos que incluem as sequências codificadoras de várias proteínas (Fig. 1 deste quadro). Então, este transcrito sofre processamento alternativo para gerar mRNAs separados para componentes de virions individuais, como as proteínas héxon e fibrila. Todos os mRNAs possuem a mesma sequência 5', que resulta da união de três sequências curtas que não codificam proteínas, conhecida como sequência-líder tripartida. A líder sofre proces-

samento alternativo às sequências codificadoras do héxon, das fibrilas e de outras proteínas do virion, gerando cada um dos futuros mRNAs virais.

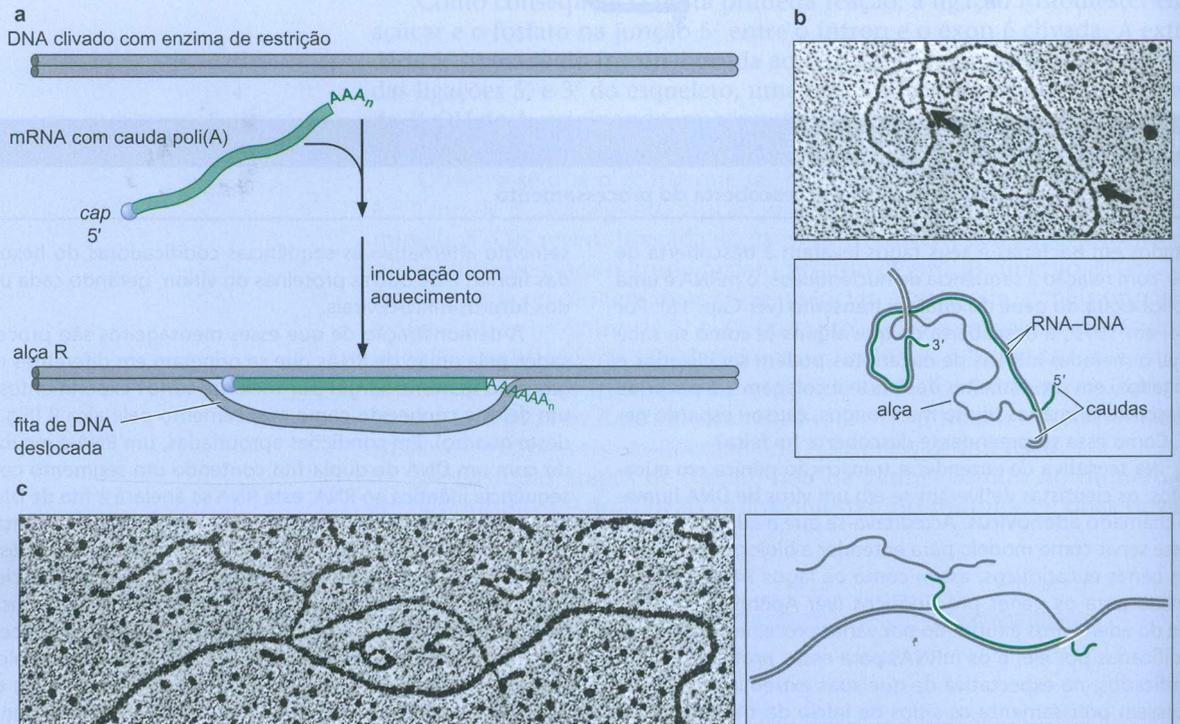
A demonstração de que esses mensageiros são processados pela união de RNAs que se originam em diferentes regiões do genoma surgiu por meio de vários experimentos – um deles é conhecido como mapeamento pela alça R (Fig. 2 deste quadro). Em condições apropriadas, um RNA é incubado com um DNA de dupla-fita contendo um segmento com sequência idêntica ao RNA; este RNA se anelará à fita de DNA complementar, deslocando a fita de DNA não complementar, deixando-a em forma de alça (Fig. 2a deste quadro). Após a técnica de coloração utilizada para visualizar ácidos nucleicos, a alça R pode ser observada ao microscópio eletrônico, uma vez que os duplícies de RNA-DNA e DNA-DNA aparecem mais espessos do que os ácidos nucleicos de fita simples. Quando esse experimento foi realizado em mensageiros de adenovírus, observou-se que as alças R não eram totalmente contíguas a uma única região do DNA. Em vez disso, e dependendo do fragmento de DNA viral utilizado, uma ou ambas as extremidades do RNA projetavam-se das alças de RNA como caudas de fita simples (Fig. 2b deste quadro). Em outros casos, uma das caudas anela a um fragmento de DNA de uma região diferente do genoma viral (Fig. 2c deste quadro). Claramente, esses mRNAs eram moléculas compostas, oriundas de sequências complementares a regiões do genoma que não eram contíguas e que haviam sido ligadas. Esses e outros tipos de experimentos de anelamento DNA-RNA foram usados para deduzir o padrão de processamento alternativo mostrado na Figura 1 deste quadro.

(continua)

Quadro 14-1 (Continuação)



QUADRO 14-1 FIGURA 1 Mapa do genoma do adenovírus-2 humano. O mapa apresenta os padrões de transcrição dos mRNAs tardios, incluindo o transcrito primário (representado pela seta longa em verde-escuro na parte superior); as sequências-líderes tripartidas encontradas nas posições 16,6, 19,6 e 26,6 (representadas como barras verdes); e as posições no mapa das sequências de DNA que codificam os vários mRNAs tardios (representados pelas setas curtas em verde-escuro).



QUADRO 14-1 FIGURA 2 Mapeamento de alça R dos mRNAs tardios do adenovírus-2. (a) A representação esquemática mostra a formação de uma estrutura de alça R. Um fragmento de DNA de dupla-fita gerado por digestão com uma endonuclease de restrição é incubado com mRNA e aquecido um pouco acima da temperatura de desnaturação do DNA, em 80% de formamida. O híbrido formado entre o mensageiro e sua sequência complementar de DNA causa o deslocamento da segunda fita de DNA. Observa-se que a cauda poli(A) do mRNA (não codificada pelo DNA) (ver Cap. 12) projeta-se a partir do fim do dúplex híbrido. (b) Micrografia eletrônica e diagrama esquemático de uma alça R observada após a incubação do mRNA de héxon com uma sequência de DNA complementar da região tardia do genoma do adenovírus-2. Observam-se as extensões das extremidades 5' e 3' no mensageiro. (Linhas cinzas) DNA; (linhas verdes) RNA. (Reproduzida, com permissão, de Berget S.M. et al. 1977. *Proc. Natl. Acad. Sci.* 74: 3171-3175. © National Academy of Sciences.) (c) Micrografia eletrônica e diagrama esquemático de uma alça R observada após a incubação do mRNA de fibra (fibrila) com dois DNAs, o genoma completo do adenovírus e um fragmento de endonuclease de restrição derivado de uma região inicial do genoma. (Reproduzida, com permissão, de Chow L.T. et al. 1977. *Cell* 12: 1-8, p. 2. © Elsevier.)

feitas. Como se trata apenas de um rearranjo das ligações, não há necessidade do fornecimento de energia para o processo. Como será visto a seguir, porém, uma grande quantidade de ATP é consumida durante a reação de processamento. Essa energia não é necessária para a reação química, mas sim para formar e operar adequadamente a maquinaria do processamento.

Outro ponto relacionado à reação de processamento é a direção: o que assegura que o processamento ocorra apenas para a frente – isto é, na direção dos produtos ilustrados na Figura 14-4? Em princípio, as reações poderiam ocorrer na outra direção e, de fato, isso pode ocorrer de maneira forçada, sob circunstâncias especiais. Mas na prática, isso não ocorre na célula, e será descrito aqui como isso é garantido.

MAQUINARIA DO SPLICEOSSOMO

O processamento do RNA é executado por um grande complexo chamado spliceossomo

As reações de transesterificação recém-descritas são promovidas por uma grande “máquina” molecular, chamada **spliceossomo**. Este complexo compreende cerca de 150 proteínas e cinco RNAs e seu tamanho equivale ao de um ribossomo, a “máquina” que traduz o mRNA em proteína (Cap. 15). Para realizar o processamento, mesmo que apenas uma reação, o spliceossomo hidrolisa várias moléculas de ATP. Surpreendentemente, acredita-se que muitas funções do spliceossomo são realizadas pelos RNAs que o compõe e não pelas proteínas – novamente, outra semelhança com o ribossomo. São esses RNAs que localizam os elementos nas fronteiras entre íntrons e éxons e, provavelmente, participam da catálise da reação de processamento.

Os cinco RNAs (U1, U2, U4, U5 e U6) são chamados, conjuntamente, de **pequenos RNAs nucleares (snRNAs, small nuclear RNAs)**. Cada um deles tem entre 100 e 300 nucleotídeos na maioria dos eucariotos e está complexado a várias proteínas. Estes complexos RNA-proteína são chamados de **pequenas proteínas ribonucleares (snRNPs, small nuclear ribonuclear proteins)**. No Capítulo 6, viu-se a estrutura cristalográfica de uma seção do snRNA U1 ligada a uma das proteínas da snRNP U1 (Fig. 6-18).

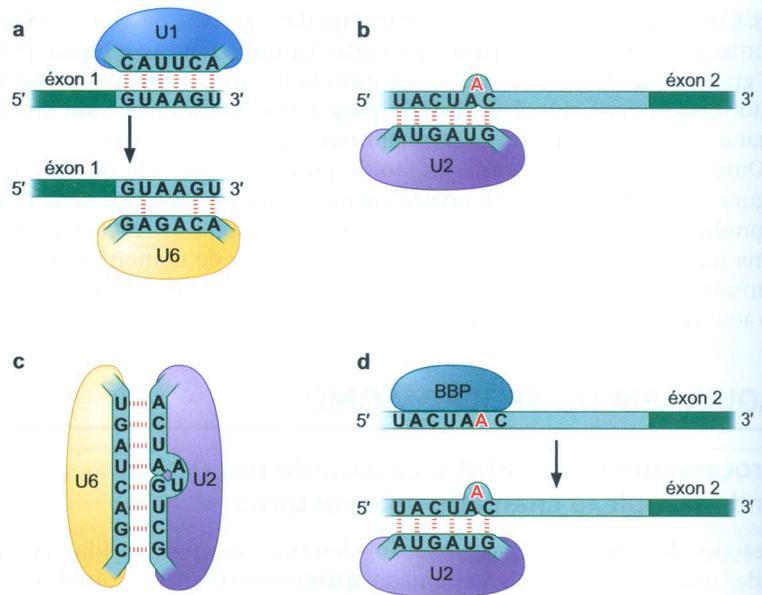
O spliceossomo é o grande complexo formado por estas snRNPs, mas sua constituição exata difere em estágios distintos da reação de processamento: diferentes snRNPs vêm e vão em períodos diferentes, cada uma delas realizando funções específicas na reação. No spliceossomo, também existem várias proteínas que não fazem parte das snRNPs, e outras ligadas fracamente ao complexo.

As snRNPs desempenham três funções no processamento: elas reconhecem o sítio de processamento 5' e o sítio de ramificação; aproximam esses sítios, quando necessário; e catalisam (ou auxiliam a catálise) a clivagem do RNA e as reações de religação. As interações RNA-RNA, RNA-proteína e proteína-proteína são importantes para a realização dessas funções. Iniciaremos considerando algumas das interações RNA-RNA. Essas interações estão presentes em cada snRNP, entre snRNPs diferentes e entre snRNPs e o pré-mRNA.

A Figura 14-6a, por exemplo, mostra a interação, por meio de pareamento de bases complementares, entre o snRNA U1 e o sítio de processamento 5' do pré-mRNA. Na continuação da reação, esse sítio de processamento é reconhecido pelo snRNA U6. Outro exemplo, apresentado na Figura 14-6b, mostra o sítio de ramificação sendo reconhecido pelo snRNA U2. Um terceiro exemplo, na Figura 14-6c, mostra uma interação entre os snRNAs U2 e U6. Esta aproxima o sítio de processamento 5' e o sítio de ramificação. São essas interações e outras semelhantes, bem como os rearranjos que delas resultam, que promovem a reação de processamento e contribuem para sua precisão, com será visto adiante.

Como mencionado anteriormente, algumas não snRNPs estão envolvidas no processamento. Um exemplo disso, a U2AF (fator auxiliar U2), reco-

FIGURA 14-6 Alguns híbridos RNA-RNA formados durante a reação de processamento. Em alguns casos, (a) diferentes snRNPs reconhecem as mesmas sequências (ou sequências sobrepostas) no pré-mRNA, como mostrado aqui para o reconhecimento do sítio de processamento 5' pela U1 e U6. (b) snRNP U2 faz o reconhecimento do sítio de ramificação. (c) O pareamento RNA:RNA das snRNPs U2 e U6 é mostrado. (d) Por fim, a mesma sequência em um pré-mRNA é reconhecida por uma proteína (que não faz parte de snRNP) em uma etapa e é substituída por uma snRNP em outra. Cada uma dessas alterações é acompanhada pela chegada ou pela saída dos componentes do spliceossomo e o rearranjo estrutural necessário para a continuação da reação de processamento. As sequências nesta figura são de levedura.



nhece o trato de polipirimidina (Py)/sítio de processamento 3' e, na etapa inicial da reação de processamento, ajuda outra proteína, a BBP (proteína de ligação ao ponto de ramificação), a ligar-se ao sítio de ramificação. A BBP (também chamada de SF1 em sistemas mamíferos) é, então, deslocada pela snRNP U2, como mostrado na Figura 14-6d. Outras proteínas envolvidas na reação de processamento incluem fatores de anelamento ao RNA, que ajudam a carregar as snRNPs no mRNA, e as proteínas helicases de *DEAD-box*. Essas proteínas usam sua atividade de ATPase para dissociar certas interações RNA-RNA, permitindo a formação de pares alternativos e, assim, promovendo os rearranjos que ocorrem durante a reação de processamento. Elas também são necessárias para remover o mRNA processado do spliceossomo e desencadear a sua desmontagem.

VIAS DE PROCESSAMENTO

Formação, rearranjo e catálise no spliceossomo: a via de processamento

As etapas do processamento são mostradas na Figura 14-7. O que está representado é uma via canônica e, em um caso qualquer, determinadas etapas podem ser levemente diferentes em sua ordem ou podem até mesmo ter ordem inversa, situações essas que serão retomadas mais tarde. Mas a via, como primeiramente apresentada, revela a série extraordinária de eventos realizados pelo dinâmico spliceossomo para dirigir a reação de processamento na célula.

Inicialmente, o sítio de processamento 5' é reconhecido pela snRNP U1 (pelo pareamento de bases entre seu snRNA e o pré-mRNA, como mostrado na Fig. 14-6). O U2AF é formado por duas subunidades, a maior delas (65) liga-se ao trato Py e a menor (35), ao sítio de processamento 3'. A primeira subunidade interage com a BBP (SF1), auxiliando em sua ligação ao sítio de ramificação. Este arranjo de proteínas e RNA é chamado de complexo inicial (E, *early*).

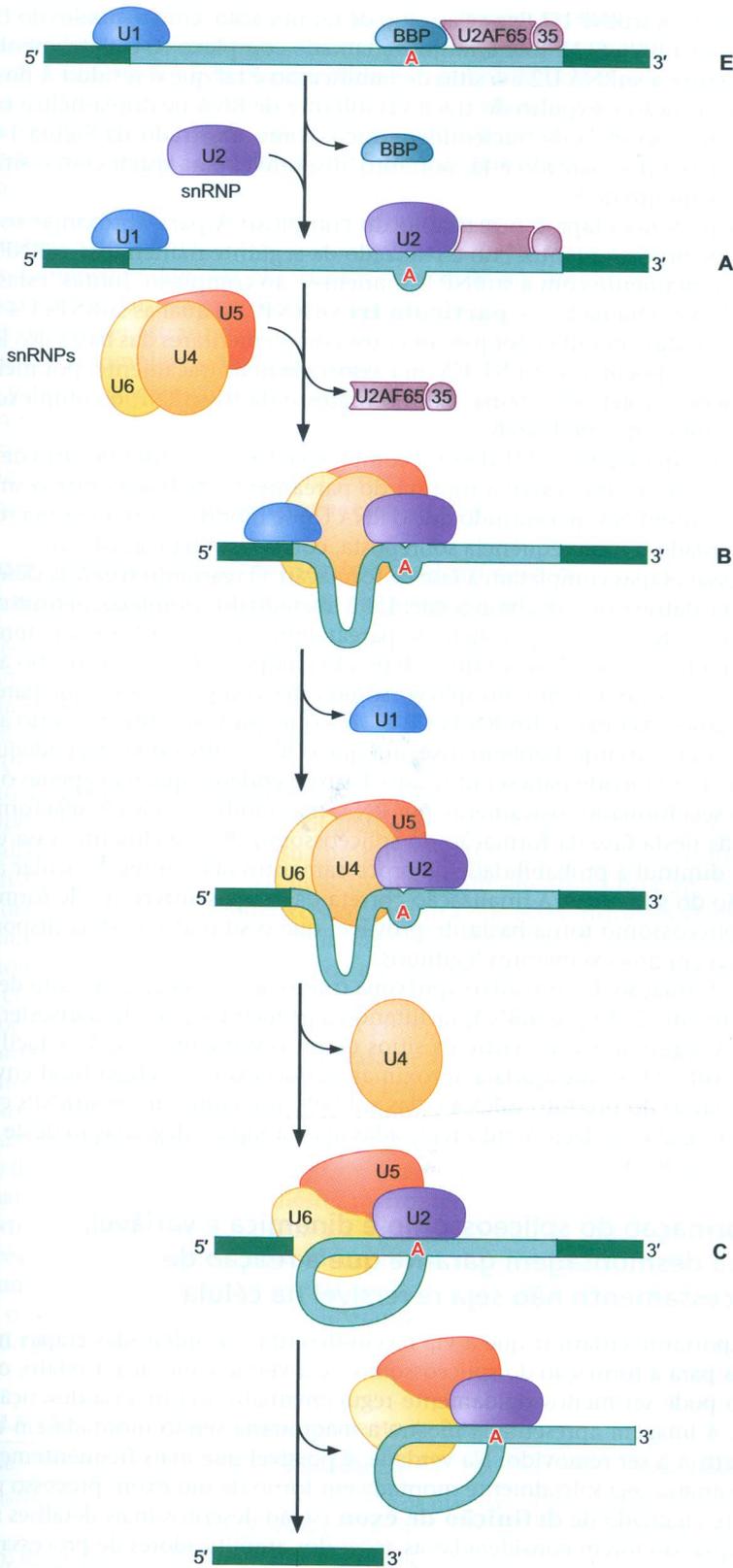


FIGURA 14-7 Etapas da reação de processamento mediada pelo spliceossomo. A formação e a atuação do spliceossomo estão ilustradas, e os detalhes de cada etapa estão descritos no texto. Determinados componentes da maquinaria de processamento entram ou saem do complexo a cada etapa, e essas alterações estão associadas aos rearranjos estruturais necessários para a sequência do processamento. Observa-se que o nome de cada complexo é mostrado à direita. As evidências sugerem que alguns dos componentes mostrados não chegam ou saem exatamente nos momentos indicados nesta figura; eles podem, por exemplo, continuar presentes, mas enfraquecer sua associação com o complexo, em vez de se dissociarem completamente. Também não é possível ter certeza da ordem de algumas alterações mostradas, sobretudo as duas etapas que envolvem alterações no pareamento de U6: quando ela substitui U1 no sítio de processamento 5', comparado a quando ela substitui U4 na ligação a U2. Apesar dessas incertezas, o envolvimento fundamental dos diversos componentes da maquinaria, nas diferentes etapas da reação de processamento, e a natureza dinâmica geral do spliceossomo seguem o esquema ilustrado na figura.

Então, a snRNP U2 liga-se ao sítio de ramificação, com o auxílio do U2AF e desloca a BBP (SF1). Este arranjo é chamado complexo A. O pareamento de bases entre a snRNA U2 e o sítio de ramificação é tal que o resíduo A do sítio de ramificação é expulso do trecho resultante de RNA de dupla-hélice como uma protuberância de nucleotídeo único, como mostrado na Figura 14-6b. O resíduo A não pareado está, portanto, disponível para reagir com o sítio de processamento de 5'.

A próxima etapa é o rearranjo do complexo A para aproximar os três sítios de processamento. Isso é realizado da seguinte maneira: as snRNPs U4 e U6, juntamente com a snRNP U5, unem-se ao complexo. Juntas, estas três snRNPs são chamadas de **partícula tri-snRNP**, na qual as snRNPs U4 e U6 são mantidas em união por pareamentos complementares das bases dos RNAs que as compõem, e a snRNP U5 está associada mais fracamente, por meio de interações proteína-proteína. Com o ingresso da tri-snRNP, o complexo A é convertido em complexo B.

Na etapa seguinte, U1 deixa o complexo e U6 o substitui no sítio de processamento 5'. Isso exige a ruptura do pareamento de bases entre o snRNA U1 e o pré-mRNA, permitindo que o RNA U6 se hibridize a esta mesma região (na verdade, a uma sequência sobreposta, como visto na Fig. 14-6a).

Essas etapas completam a fase de formação. O rearranjo seguinte desencadeia a catálise e ocorre como segue: U4 é liberado do complexo, permitindo a interação de U6 e U2 (por meio do pareamento de bases RNA:RNA, apresentado na Fig. 14-6c). Esse arranjo, chamado complexo C, produz o sítio ativo. Isto é, o rearranjo reúne no spliceossomo estes componentes – que parecem ser regiões exclusivas dos RNAs U2 e U6 – que, juntos, formam o sítio ativo. O mesmo rearranjo também assegura que o RNA substrato esteja adequadamente posicionado para ser utilizado. É surpreendente que não apenas o sítio ativo seja formado basicamente por RNA, mas também, que ele seja formado apenas nesta fase da formação do spliceossomo. Provavelmente, essa estratégia diminui a probabilidade de processamentos aberrantes. Vincular a formação do sítio ativo à finalização correta das etapas anteriores de formação do spliceossomo torna bastante provável que o sítio ativo esteja disponível apenas em processamentos legítimos.

A formação do sítio ativo aproxima o sítio de ramificação do sítio de processamento 5' do pré-mRNA, facilitando a primeira reação de transesterificação. A segunda reação, entre os sítios de processamento 5' e 3', é facilitada pela snRNP U5, que ajuda a aproximar os dois éxons. A etapa final envolve a liberação do produto mRNA e das snRNPs. Inicialmente, as snRNPs continuam ligadas ao laço, sendo recicladas após a rápida degradação deste fragmento de RNA.

A formação do spliceossomo é dinâmica e variável, e sua desmontagem garante que a reação de processamento não seja reversível na célula

É importante enfatizar que a via recém-descrita – a ordem das etapas necessárias para a formação do spliceossomo – é a versão canônica. De fato, o processo pode ser menos rigidamente regulamentado do que essa descrição sugere. A imagem apresentada mostra a maquinaria sendo montada em torno do íntron a ser removido. Na verdade, é possível que mais frequentemente a maquinaria seja inicialmente montada em torno de um éxon, processo geralmente chamado de **definição de éxon** (serão descritos mais detalhes sobre isso quando forem consideradas as ações dos amplificadores de processamento). Além disso, a ordem precisa dos eventos provavelmente varia em alguma extensão – por exemplo, o pareamento do sítio de processamento pode ocorrer antes ou após o recrutamento de tri-snRNP: os detalhes dependerão das sequências de RNA e de fatores de limitação de taxas em um caso qualquer. Além disso, muitas etapas durante a formação do spliceossomo podem ser revertidas.

Anteriormente, mencionou-se que as duas reações principais do processamento poderiam, em princípio, ocorrer bidirecionalmente, mas que na célula isso não é observado. Esta direcionalidade é garantida pela rápida desmontagem do spliceossomo, imediatamente após a realização da segunda reação. A desmontagem é dirigida por uma das helicases DEAD-box que foi mencionada, esta chamada de Prp22. Esta proteína é necessária para a segunda etapa catalítica do processamento e também para remover o mRNA processado do spliceossomo. Mutações que eliminam esta última função também bloqueiam a desmontagem do spliceossomo e, em tal situação, as reações de processamento podem ser observadas ocorrendo em ambas as direções em spliceossomos purificados.

Pode parecer estranho que a maquinaria e o mecanismo de processamento sejam tão complexos. Por que teriam evoluído desse modo? Não teria sido mais simples ligar os éxons em uma única reação, em vez das duas reações, como recém-descrito? Para analisar essa questão, será discutido a seguir um grupo de íntrons que – ao contrário dos considerados até agora – são capazes de *autoexcisão* do pré-mRNA, sem a necessidade de spliceossomo. Eles são chamados de **íntrons de autoprocessamento**.

Íntrons de autoprocessamento revelam que o RNA pode catalisar o processamento de RNA

As três classes de processamento encontradas nas células (excluindo o processamento de tRNAs, que será discutido no Cap. 15) são apresentadas na Tabela 14-1. Até o momento, nós só discutimos o processamento de pré-mRNAs nucleares, mediado pelo spliceossomo encontrado em todos os eucariotos. A Tabela 14-1 também apresenta os chamados íntrons de autoprocessamento do **grupo I** e do **grupo II**. Por “autoprocessamento”, entende-se que o próprio íntron, no RNA precursor, dobra-se em uma conformação específica e catalisa a reação química de sua própria liberação (deve-se lembrar das características gerais das enzimas de RNA que foram discutidas no Cap. 5). Em uma definição prática, a “autoexcisão” significa que esses íntrons podem realizar autorremoção dos seus RNAs, *in vitro*, na ausência de quaisquer proteínas ou moléculas de RNA. Os íntrons de autoprocessamento são agrupados em duas classes, de acordo com a sua estrutura e mecanismo de processamento. A rigor, os íntrons de autoexcisão não são enzimas (“catalisadores”), porque promovem apenas um ciclo de processamento de RNA (como será visto no Quadro 14-2, Conversão dos íntrons do grupo I em ribozimas).

No caso dos íntrons do grupo II, a química de processamento e os intermediários de RNA produzidos são iguais aos dos pré-mRNAs nucleares. Isto é, como ilustrado na Figura 14-8, o íntron usa um resíduo A do sítio de ramificação para atacar a ligação fosfodiéster na fronteira entre a sua extremidade 5' e a extremidade do éxon que está a 5' – isto é, no sítio de processamento 5'. Essa reação produz o laço ramificado, como se viu anteriormente, e é seguida por uma nova reação, em que a extremidade 3'-OH recém-liberada do éxon ataca o sítio de processamento 3', liberando o íntron na forma de um laço e ligando os éxons a 3' e a 5'.

TABELA 14-1 Três classes de processamento de RNA

Classe	Abundância	Mecanismo	Maquinaria catalítica
Pré-mRNA nuclear	Muito comum; utilizado na maioria dos genes eucarióticos	Duas reações de transesterificação; sítio de ramificação A	Spliceossomos principal e minoritário
Íntrons do grupo II	Raros; em alguns genes de organelas eucarióticas e em procariotos	Igual ao do pré-mRNA	Enzima de RNA codificada pelo íntron (ribozima)
Íntrons do grupo I	Raros; no rRNA nuclear de alguns eucariotos, em genes de organelas e em alguns genes procarióticos	Duas reações de transesterificação; sítio de ramificação G	Igual ao grupo II

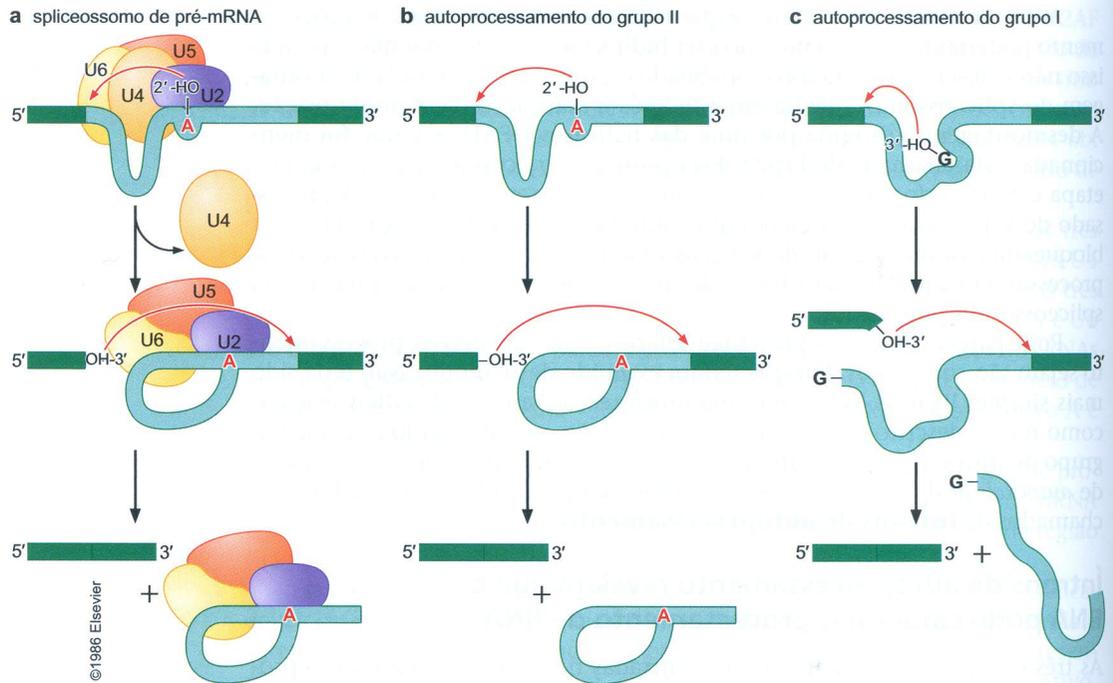


FIGURA 14-8 Íntrons do grupo I e do grupo II. A figura compara as reações de autoexcisão de íntrons dos grupos I e II e da reação mediada pelo spliceossomo, já descrita. No caso dos íntrons do grupo II, a reação química é essencialmente idêntica à dos spliceossomos, com uma adenina altamente reativa do íntron iniciando o processamento e levando à formação de um produto em laço. No caso do íntron do grupo I, o RNA dobra-se para formar um bolso de ligação à guanina, permitindo que a molécula se ligue a um nucleotídeo de guanina livre para iniciar o processamento. Embora, *in vitro*, esses íntrons possam realizar autorremoção do RNA sem interferência de proteínas, *in vivo*, em geral, há a necessidade de diversos componentes proteicos para estimular a reação. (Adaptada, com permissão, de Cech T.R. 1986. *Cell* **44**: 207-210, Fig. 1. © Elsevier.)

Íntrons do grupo I liberam um íntron linear em vez de um laço

Os íntrons do grupo I são processados por uma via diferente (Fig. 14-8c). Em vez de usar um resíduo A do sítio de ramificação, eles usam um nucleotídeo ou nucleosídeo G livre. Esse resíduo G liga-se ao RNA, e seu grupo 3'-OH é apresentado ao sítio de processamento 5'. O mesmo tipo de reação de transesterificação, que nos casos anteriores resultava na formação do laço, aqui liga G à extremidade 5' do íntron. Então, a segunda reação prossegue como nos exemplos anteriores: a extremidade 3' liberada do éxon ataca o sítio de processamento 3'. Isso liga os dois éxons e libera o íntron que, neste caso, é linear e não uma estrutura em laço.

Os íntrons do grupo I, que são menores que os do grupo II, possuem uma estrutura secundária conservada comum (os dobramentos do RNA foram discutidos no Cap. 5). A estrutura dos íntrons do grupo I inclui um bolso (ou fenda) de ligação que acomodará qualquer nucleotídeo ou nucleosídeo de guanina, desde que ele esteja na forma de ribose. Além desse bolso, os íntrons do grupo I contêm uma “sequência interna guia”, cujas bases pareiam com as bases da sequência do sítio de processamento 5', determinando o local exato para o ataque nucleofílico pelo nucleotídeo G (ver Quadro 14-2).

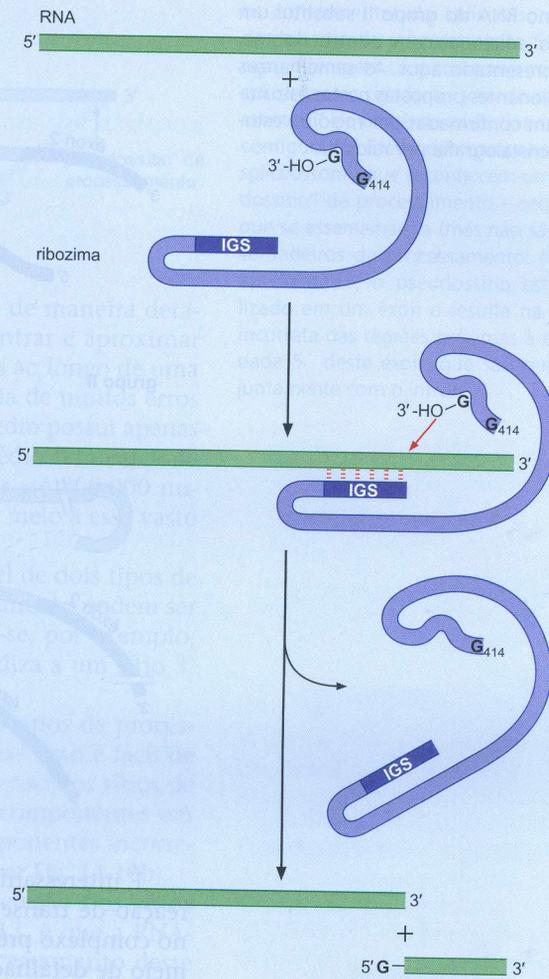
Um íntron de autoprocessamento típico tem entre 400 e 1.000 nucleotídeos de comprimento e, ao contrário dos íntrons removidos por spliceossomos, grande parte da sequência de um íntron de autoprocessamento é essencial para a reação de processamento. Esta restrição na sequência existe pois o íntron deve sofrer um dobramento exato para que possa realizar a reação

▶ EXPERIMENTOS - CHAVE

Quadro 14-2 Conversão dos íntrons do grupo I em ribozimas

Uma vez que um íntron de autoprocessamento do grupo I tenha sido removido, o sítio ativo que ele contém permanece intacto. Então, o que impede a autorreversão da reação de processamento? Um fator é a alta concentração celular de nucleotídeos G, o que favorece muito um dos sentidos da reação. Além disso, o íntron sofre uma reação subsequente que o impede de participar efetivamente da reação reversa. Convenientemente, no final da extremidade 3' do íntron há um G, que pode ligar-se ao bolso de ligação a G. Enquanto isso, a extremidade 5' do íntron pode ligar-se ao longo da sequência-guia interna. Assim, uma terceira reação de transesterificação pode ocorrer para circularizar o íntron. A nova ligação formada com o G terminal é fraca e hidrolisa-se espontaneamente. Em consequência, o íntron é relinearizado, mas é truncado e, assim, impedido de realizar a reação de processamento reversa.

Conforme explicado anteriormente, os íntrons do grupo I (e do II) não são enzimas porque possuem um número de *turnover* de apenas 1. Mas eles podem ser prontamente convertidos em enzimas (ribozimas) da seguinte maneira (Fig. 1 deste quadro): o íntron relinearizado descrito retém seu sítio ativo. Se fornecermos um G livre e um substrato que contenha uma sequência complementar à sua sequência interna guia, ele catalisará, repetidamente, a clivagem de moléculas de substrato. Assim, pode-se converter um íntron de grupo I em uma ribozima, de modo semelhante ao da enzima de autoclivagem Hammerhead (cabeça-de-martelo), que pode ser convertida em ribozima ao separar-se o sítio ativo do substrato (Cap. 5). Pode-se ir um passo adiante neste processo alterando-se a sequência interna guia e, assim, produzindo ribonucleases sob medida que clivam uma molécula de RNA à escolha.

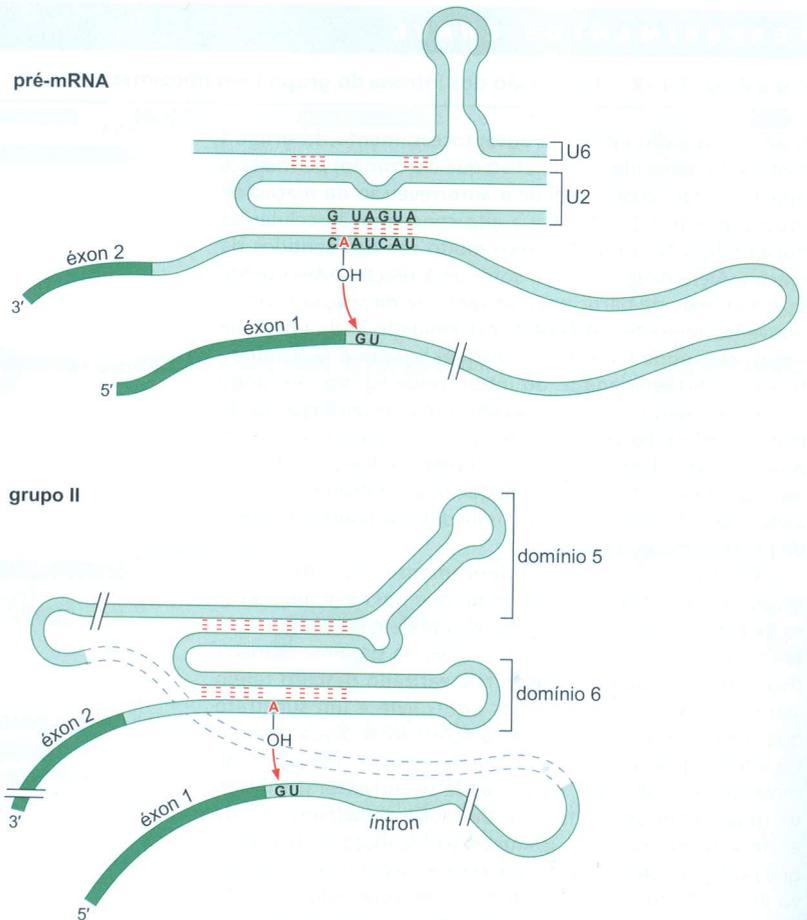


QUADRO 14-2 FIGURA 1 Os íntrons do grupo I podem ser convertidos em ribozimas verdadeiras.

química necessária. Além disso, *in vivo*, o íntron é complexado a várias proteínas que ajudam a estabilizar a estrutura correta – em parte, pelo isolamento das regiões do esqueleto umas em relação às outras. Assim, esse dobramento requer que certas seções do esqueleto do RNA estejam próximas entre si e, se não forem protegidas, as cargas negativas dos fosfatos dessas regiões seriam repelidas entre si. *In vitro*, altas concentrações salinas (e, portanto, íons positivos) compensam a ausência dessas proteínas. É por isso que se sabe que as proteínas não são necessárias para a reação de processamento propriamente dita.

Acredita-se que a semelhança entre a reação química do autoprocessamento e a do processamento mediado pelo spliceossomo seja reflexo de uma relação evolutiva. Talvez íntrons ancestrais semelhantes aos íntrons de auto-excisão do grupo II tenham sido o ponto de partida para a evolução do atual processamento do pré-mRNA. As funções catalíticas fornecidas pelo RNA foram mantidas, mas a necessidade de extensiva especificidade de sequência no íntron propriamente dito foi aliviada pela função em *trans* de snRNAs e suas proteínas associadas. Desta maneira, os íntrons precisaram apenas manter o mínimo de elementos de sequência necessários para desencadear o processamento nos locais corretos. Assim, permitiu-se a existência de muitas e variadas sequências e tamanhos de íntrons.

FIGURA 14-9 Dobramento das regiões catalíticas do RNA, proposto para processamentos de íntrons do grupo II e de pré-mRNAs. A região traçada no RNA do grupo II substitui um domínio adicional com quatro dobras, não representado aqui. As semelhanças impressionantes propostas nestas estruturas foram confirmadas por meio de estudos de cristalografia por raios X.



É interessante que a estrutura da região catalítica que realiza a primeira reação de transesterificação seja muito parecida nos íntrons do grupo II e no complexo pré-mRNA/snRNP (Fig. 14-9). Essa observação, confirmada por meio de detalhados estudos de cristalografia por raios X, recentemente, fomenta a grande especulação (discutida no Cap. 17) de que, nos primórdios da evolução dos organismos modernos, muitas funções catalíticas eram realizadas por RNAs, funções essas que, em sua maioria, foram substituídas por proteínas desde então. No caso dos spliceossomos e ribossomos, entretanto, as funções não foram completamente substituídas por proteínas. Em vez disso, resquícios de mecanismos catalisados por RNA continuam a ser fundamentais na complexa maquinaria atual.

Como o spliceossomo encontra os sítios de processamento com precisão?

Um mecanismo de proteção contra processamentos impróprios já foi visto: o sítio ativo do spliceossomo é formado apenas sobre sequências de RNA que tenham sido eficientemente reconhecidas por vários elementos durante a formação do spliceossomo. Desse modo, o sítio de processamento 5', por exemplo, precisa primeiro ser reconhecido pela snRNP U1 e, depois, pela snRNP U6. É improvável que ambas reconheçam uma sequência incorreta e, por isso, a seleção é estridente. Ainda assim, o reconhecimento do sítio de processamento apropriado no pré-mRNA continua sendo importante.

Considere-se o seguinte: um gene humano médio possui sete ou oito éxons e pode ser processado de três modos alternativos. Existe, porém, um gene humano com 363 éxons e um gene de *Drosophila* que pode ser proces-

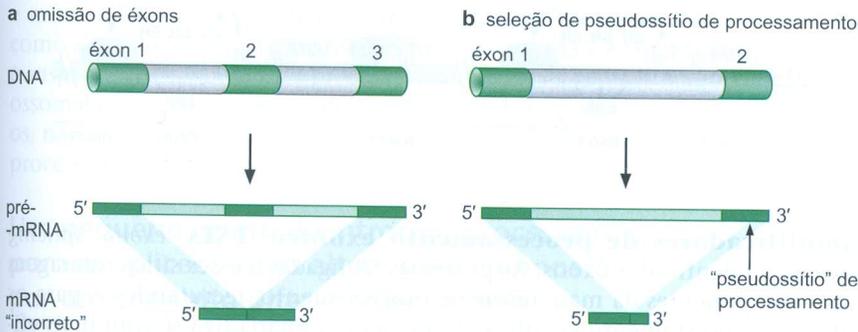


FIGURA 14-10 Erros produzidos por falhas na seleção de sítios de processamento. (a) Consequência da omissão de um éxon. Isso ocorre quando os componentes do spliceossomo ligados ao sítio de processamento 5' de um éxon interagem com componentes de um spliceossomo ligado ao sítio de processamento 3' de um éxon posterior, mas não contíguo. (b) Efeito dos componentes de spliceossomo que reconhecem um "pseudossítio" de processamento – sequências que se assemelham a (mas não são) sítios verdadeiros de processamento. No caso apresentado, o pseudossítio está localizado em um éxon e resulta na excisão incorreta das regiões próximas à extremidade 5' deste éxon, que são removidas juntamente com o íntron.

sado de 38 mil modos alternativos, caso que será descrito de maneira detalhada na próxima seção. Se as snRNPs tivessem que encontrar e aproximar corretamente os sítios de processamento 5' e 3' apropriados ao longo de uma molécula inteira de RNA, sem nenhuma ajuda, a ocorrência de muitos erros seria inevitável. Deve-se lembrar também de que o éxon médio possui apenas 150 nucleotídeos de comprimento, enquanto o íntron médio tem cerca de 3.000 nucleotídeos (como visto, alguns íntrons podem ter até 800.000 nucleotídeos). Portanto, os éxons têm de ser identificados em meio a esse vasto oceano de íntrons.

O reconhecimento do sítio de processamento é passível de dois tipos de erros (Fig. 14-10). O primeiro é quando os sítios de processamento podem ser "omitidos" e deixados para trás, e os componentes ligam-se, por exemplo, a um determinado sítio de processamento 5' que se hibridiza a um sítio 3' adiante do sítio correto.

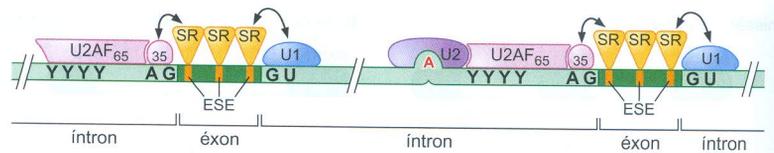
Segundo, outros sítios, com sequência semelhante aos sítios de processamento verdadeiros, podem ser reconhecidos erroneamente. Isso é fácil de compreender quando se recorda que as sequências de consenso dos sítios de processamento são pouco definidas. Assim, por exemplo, componentes em um dado sítio de processamento 5' podem parear com componentes incorretamente ligados neste "pseudossítio" de processamento 3' (ver Fig. 14-10b).

Existem duas maneiras de aumentar a precisão da seleção dos sítios de processamento. A primeira, como se viu no Capítulo 13, é que a RNA-polimerase II carrega várias proteínas que atuam no processamento deste RNA, durante a transcrição de um gene (ver Cap. 13, Fig. 13-19). Isso inclui as proteínas envolvidas na remoção dos íntrons. Quando um sítio de processamento 5' é encontrado em um RNA recém-sintetizado, certos componentes são transferidos da "cauda" carboxiterminal da polimerase (a parte da enzima à qual se ligam) para o RNA. Neste local, os componentes do sítio de processamento 5' estão prontos para interagir com fatores que se liguem ao próximo sítio de processamento 3' a ser sintetizado. Portanto, o sítio de processamento 3' correto pode ser reconhecido antes que outros sítios competidores a jusante tenham sido transcritos. Esse processo de carregamento cotranscricional de componentes do processamento reduz bastante a probabilidade de omissão ou salto de éxons.

(É importante ressaltar que, apesar de grande parte da maquinaria de processamento ser montada enquanto o gene está sendo transcrito, isso não significa, necessariamente, que os íntrons sejam removidos nessa ordem. Assim, contrastando com diversos outros processos – transcrição, replicação e outros –, parece não haver um mecanismo de "percurso" envolvido, o qual faça com que a maquinaria se forme em uma extremidade do gene ou mensagem e atue à medida que se desloca para a outra extremidade.)

Um segundo mecanismo que evita o uso incorreto dos sítios assegura que os sítios de processamento próximos de éxons (e, portanto, provavelmente verdadeiros) sejam reconhecidos preferencialmente. As chamadas proteínas SR (ricas em serina e arginina) ligam-se a sequências chamadas

FIGURA 14-11 As proteínas SR recrutam componentes do spliceossomo para os sítios de processamento 5' e 3'. Os sítios de processamento autênticos são reconhecidos pela maquinaria de processamento por estarem próximos dos éxons. Assim, as proteínas SR ligam-se a sequências ESEs (do inglês, *exonic splicing enhancers*) do éxon e dali recrutam U2AF e snRNP U1, respectivamente, para os sítios de processamento 5' a jusante e 3' a montante. Isso dá início à formação da maquinaria de processamento nos sítios corretos, e o processamento pode prosseguir como foi descrito no texto. (Adaptada, com permissão, de Maniatis T. e Tasic B. 2002. *Nature* 418: 236-243. © Macmillan.)



amplificadores de processamento exônico (ESEs, *exonic splicing enhancers*) dentro dos éxons. As proteínas SR ligadas a esses sítios interagem com componentes da maquinaria de processamento, recrutando-os para os sítios de processamento vizinhos. Assim, a maquinaria liga-se com mais eficiência a esses sítios de processamento do que a sítios incorretos, distantes de éxons. Especificamente, as proteínas SR recrutam as proteínas U2AF para o sítio de processamento 3' e a snRNP U1 para o sítio 5' (Fig. 14-11). Como se viu anteriormente, esses fatores demarcam os sítios de processamento, para que o restante da maquinaria se posicione corretamente (Fig. 14-7). Este recrutamento ocorre por meio da interação direta entre as proteínas SR e as proteínas do spliceossomo ou por meio da interação com, e estabilização de, híbridos de RNA:RNA formados durante a formação e a atuação do spliceossomo.

Ao recrutar fatores de processamento para cada lado de um dado éxon, este processo encoraja a chamada “definição de éxon”, que foi mencionada quando se discutiu a ordem dos eventos durante a formação do spliceossomo. Ou seja, os componentes do spliceossomo são recrutados inicialmente em torno de éxons, em vez de ao redor do íntron a ser removido. Subsequentemente, os componentes próximos a um éxon irão parear com os próximos a um éxon adjacente para eliminar o íntron interveniente.

As proteínas SR são essenciais para o processamento. Elas não apenas asseguram a precisão e a eficiência do processamento constitutivo (como visto anteriormente), mas também regulam o processamento alternativo (como será visto agora). As proteínas SR apresentam uma grande diversidade: algumas são controladas por sinais fisiológicos, outras são constitutivamente ativas. Algumas são expressas preferencialmente em determinados tipos celulares, controlando o processamento de acordo com padrões específicos do tipo celular. Na seção sobre processamento alternativo, serão discutidos alguns exemplos específicos das funções das proteínas SR.

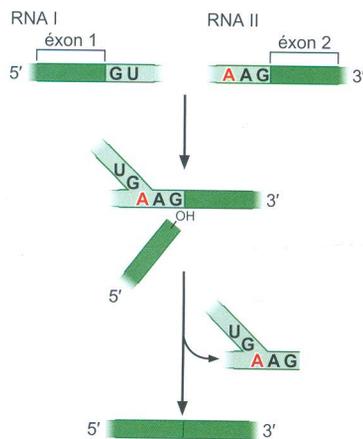


FIGURA 14-12 *Transprocessamento*. No *transprocessamento*, dois éxons localizados inicialmente em duas moléculas de DNA diferentes são processados em um mesmo mRNA. A química desta reação é a mesma do processamento-padrão descrito anteriormente, e o produto final é indistinguível. No *transprocessamento*, a única diferença é que o produto normal – o laço da reação-padrão – é substituído por uma estrutura ramificada em formato de Y. Isso ocorre porque a reação inicial une duas moléculas de RNA, em vez de formar uma alça em uma única molécula.

VARIANTES DO PROCESSAMENTO

Antes de abordar o processamento alternativo, serão descritas brevemente duas variantes da maquinaria de processamento e de reações de processamento discutidas até o momento. No primeiro caso, são considerados exemplos nos quais os dois éxons que estão sendo unidos residem em moléculas de RNA diferentes e, no segundo, uma versão especializada da maquinaria de processamento utilizada para processar um subconjunto de íntrons.

Éxons de diferentes moléculas de RNA podem ser ligados pelo *transprocessamento*

Na descrição anterior do processamento, considerou-se que o sítio de processamento 5' de um éxon é ligado ao sítio de processamento 3' do éxon que o sucede imediatamente. Porém, nem sempre é assim. No **processamento alternativo**, alguns éxons podem estar ausentes, ou um determinado éxon pode ser ligado a outro situado bem mais adiante (como será visto posteriormente). Em alguns casos, dois éxons diferentes localizados em diferentes moléculas de RNA podem ser ligados em um processo denominado **transprocessamento**. Embora seja um evento geralmente raro, o *transprocessamento* ocorre em quase todos os mRNAs de tripanossomos. No verme nematódeo (*Caenorhabditis elegans*), todos os mRNAs sofrem *transprocessamento* (para ligar uma sequência-líder 5'),

e muitos deles sofrem também *cis*processamento. A Figura 14-12 demonstra como a reação básica de processamento supradescrita é adaptada para realizar o *trans*processamento. O *trans*processamento utiliza a mesma maquinaria spliceossomal que o *cis*processamento, exceto por U1 que, pelo menos nos nematódeos, não é necessária para o *trans*processamento. Agora serão abordados casos de processamentos nos quais a maquinaria é bastante distinta.

Um pequeno grupo de íntrons é processado por um spliceossomo alternativo composto por um conjunto diferente de snRNPs

Os eucariotos superiores (incluindo mamíferos, plantas, etc.) empregam a maquinaria principal de processamento, discutida até aqui, para promover o processamento da maioria de seus pré-mRNAs. Mas nesses organismos (diferentemente de leveduras), alguns pré-mRNAs são processados por uma forma mais rara de spliceossomos. Essa forma possui alguns componentes comuns aos spliceossomos principais, mas também apresentam alguns componentes exclusivos. Assim, nessa reação de processamento, os componentes U11 e U12 do spliceossomo alternativo possuem as mesmas funções que U1 e U2 têm na forma principal, mas reconhecem sequências diferentes. U4 e U6 apresentam formas equivalentes em ambas as formas de spliceossomos – embora sejam snRNPs diferentes, elas compartilham os mesmos nomes. Finalmente, o componente U5 é idêntico no spliceossomo principal e no alternativo (chamado minoritário).

O spliceossomo minoritário reconhece íntrons de ocorrência rara com sequências consenso diferentes da maioria dos íntrons de pré-mRNAs. Deve-se enfatizar que, embora esses íntrons sejam raros, eles estão amplamente distribuídos – aproximadamente 800 genes humanos contêm pelo menos um intron minoritário. Além disso, mutações em snRNAs minoritários foram associadas a algumas doenças genéticas humanas raras.

Essa forma minoritária também é conhecida como spliceossomo AT-AC, porque as extremidades dos íntrons raros originalmente identificados contêm AU no sítio de processamento 5' e AC no sítio 3' (no RNA, ou AT e AC no DNA). Mais tarde, verificou-se que muitos íntrons processados por essa via têm extremidades GT-AG (como os íntrons comuns), mas, no restante, suas sequências de consenso diferem das dos íntrons da via principal.

Apesar de os sítios de processamento e de ramificação reconhecidos pelos dois sistemas serem diferentes, as formas principal e minoritária dos spliceossomos removem íntrons usando a mesma reação química (Fig. 14-13). Em conformidade com a conservação desse mecanismo, as diferenças de sequências nos sítios de processamento reconhecidas por essas snRNPs são correspondidas por diferenças complementares nas sequências de seus snRNAs. Portanto, conserva-se a capacidade de pareamento de bases entre as sequências dos snRNAs e as sequências dos sítios de processamento, e não a presença de determinadas sequências nas suas estruturas.

Também é importante registrar que os íntrons AT-AC poderiam enquadrar-se no esquema evolutivo discutido anteriormente. Como já foi mencionado, tem sido proposto que os íntrons do grupo II representam a forma mais antiga de íntrons. Além disso, sugere-se que os íntrons AT-AC evoluíram dos íntrons do grupo II e, estes, por fim, originaram os íntrons do pré-mRNA principal.

PROCESSAMENTO ALTERNATIVO

Um mesmo gene pode originar diferentes produtos pelo processamento alternativo

Como descrito na introdução deste capítulo, muitos genes de eucariotos superiores codificam RNAs que podem ser processados de modos alternativos, gerando dois ou mais RNAs diferentes e, portanto, diferentes produtos

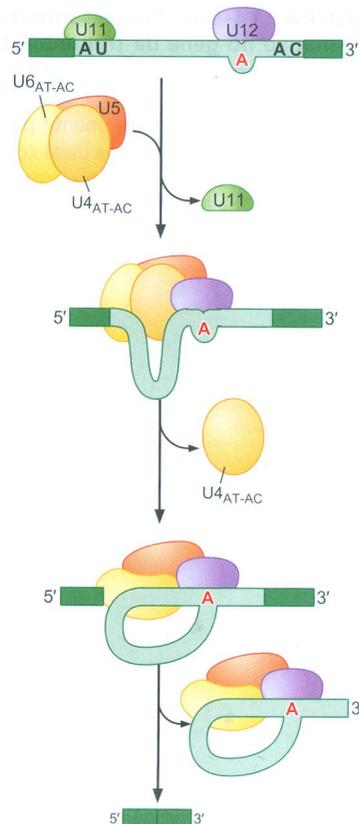
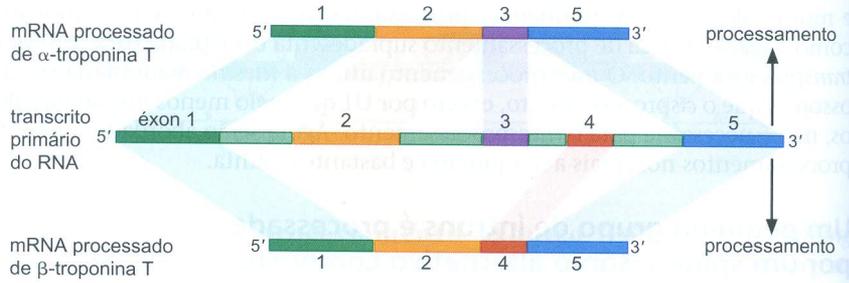


FIGURA 14-13 Processamento catalisado pelo spliceossomo AT-AC (minoritário). Esse spliceossomo funciona em uma minoria de éxons (p. ex., talvez 1:1.000 éxons de seres humanos), que têm sequências diferentes nos sítios de processamento. Apesar disso, a reação química e alguns componentes do spliceossomo são os mesmos, ao passo que outros são muito semelhantes.

FIGURA 14-14 Processamento alternativo no gene da troponina T.

A figura mostra uma região do gene da troponina T que codifica cinco éxons que gera duas formas alternativamente processadas, como indicado. Uma delas contém os éxons 1, 2, 4 e 5; a outra contém os éxons 1, 2, 3 e 5.



proteicos (ou isoformas). Hoje acredita-se que pelo menos 40% dos genes de *Drosophila* e até 90% dos genes humanos sofram processamento alternativo. Muitos genes alternativamente processados geram apenas dois produtos alternativos, mas em alguns casos, o número de potenciais alternativas que podem ser geradas a partir de um único gene é impressionante – centenas (p. ex., no caso do gene humano *Slo*) ou até mesmo milhares (para o gene de *Drosophila* *Dscam*). Às vezes, o processamento alternativo é usado como maneira de gerar diversidade, com formas alternativas sendo geradas estocasticamente. Mas em muitos casos, o processo é regulado para assegurar que diferentes produtos proteicos sejam feitos em diferentes tipos celulares ou em resposta a diferentes condições.

Para um caso simples de processamento alternativo, considere-se o gene de uma proteína muscular de mamíferos, a troponina T. A Figura 14-14 ilustra uma região de pré-mRNA produzido por esse gene, contendo cinco éxons. Esse pré-RNA é processado, formando dois mRNAs maduros alternativos, cada um contendo quatro éxons. Em cada um dos dois mRNAs, um éxon diferente é eliminado, de modo que as duas mensagens têm três éxons em comum, mas cada uma tem um éxon exclusivo.

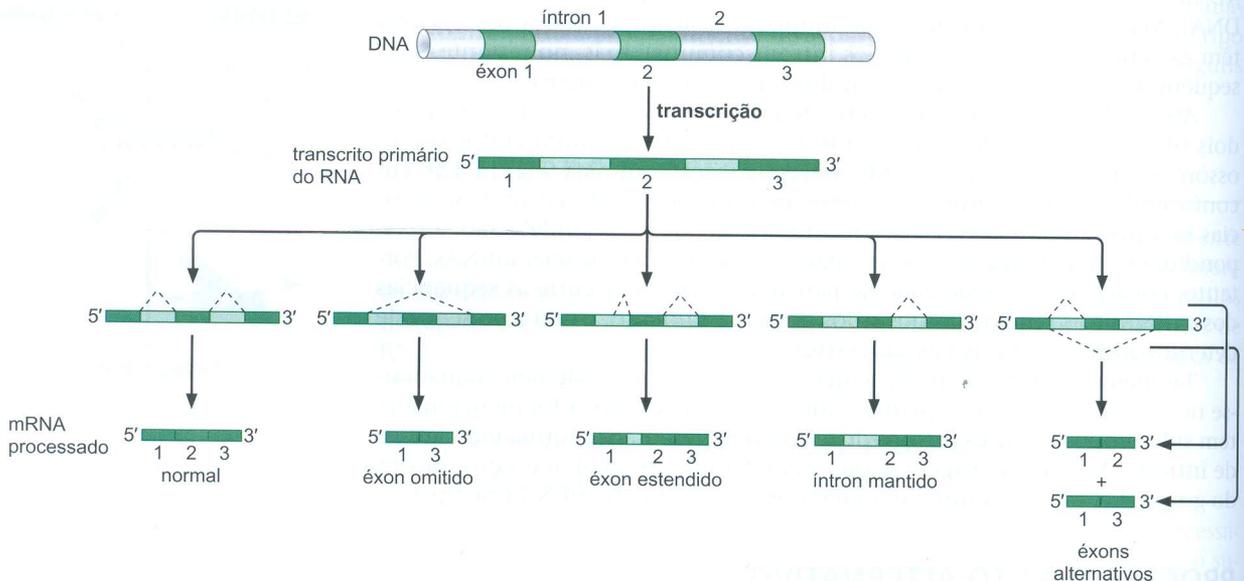


FIGURA 14-15 Cinco maneiras de realizar o processamento de um RNA. No topo, está representado um gene que codifica três éxons. Ele é transcrito em um pré-mRNA, apresentado na parte intermediária e, então, processado de cinco modos alternativos. Se todos os éxons forem incluídos, é gerado um mRNA com os três éxons. A omissão de éxon gera um mRNA que contém apenas os éxons 1 e 3. Na extensão de éxons, parte do íntron 1 é mantida junto com os três éxons. No outro caso, um íntron inteiro é mantido no mRNA maduro. Finalmente, os éxons 2 e 3 podem ser utilizados de maneira alternativa, gerando uma mistura de mRNAs, e cada um deles contém os éxons 1 e 2 ou os éxons 1 e 3.

No entanto, como mostrado na Figura 14-15, o processamento alternativo pode acontecer de várias maneiras. Assim, além de éxons alternativos, os éxons podem ser estendidos (pela seleção de um sítio de processamento alternativo 5' a jusante, ou 3' a montante). Em outros casos, os éxons podem ser "omitidos ou saltados" (deliberadamente) ou os íntrons podem ser mantidos no mensageiro maduro. Alguns processamentos alternativos resultam da transcrição de um gene a partir de promotores alternativos, permitindo que um transcrito inclua um éxon 5' ausente no outro. De maneira semelhante, sítios poli(A) alternativos permitem que éxons 3' terminais sejam estendidos ou éxons 3' terminais alternativos sejam usados em alguns transcritos de um determinado gene. Existem, ainda, casos de *trans*processamento alternativo (ver Fig. 14-12).

Em um exemplo de éxon estendido, a Figura 14-16 mostra o caso do antígeno T do vírus SV40 de símios. O gene do antígeno T codifica dois produtos proteicos: o antígeno T grande (T-ag) e o antígeno t pequeno (t-ag). As duas proteínas resultam do processamento alternativo de pré-mRNAs do mesmo gene. Assim, como ilustrado na Figura 14-16, o gene possui dois éxons, e dois mRNAs maduros diferentes resultam do uso alternativo de dois sítios diferentes de processamento 5'. No mRNA que codifica T-ag, o éxon 1 é ligado diretamente ao éxon 2, e o íntron entre eles é removido. O mRNA do t-ag, por outro lado, é formado utilizando um sítio de processamento 5' alternativo no interior do íntron. Neste caso, o mRNA inclui também uma parte do íntron. (Portanto, esse é um exemplo de "extensão de éxon" apresentada na Fig. 14-15.) A existência de um códon de terminação na fase de leitura do mRNA mais longo explica porque este codifica uma proteína mais curta.

Ambas as formas de antígenos T são feitas em uma célula infectada por SV40, mas eles possuem funções diferentes. T-ag induz a transformação e o reinício do ciclo celular, enquanto t-ag bloqueia a resposta apoptótica de células forçadas a esta via. A proporção das duas formas produzidas difere, dependendo do nível do regulador de processamento SF2/ASF. Se presente em níveis elevados, ela favorece o uso do sítio de processamento 5' que gera uma maior quantidade de mRNA de t-ag. A SF2/ASF é uma proteína SR e, quando abundante, provavelmente liga-se a sítios no interior do éxon 2 e auxilia na formação do spliceossomo no local.

Em estudos do genoma inteiro, as formas de processamento alternativo mais comumente observadas são casos nos quais éxons completos são incluídos

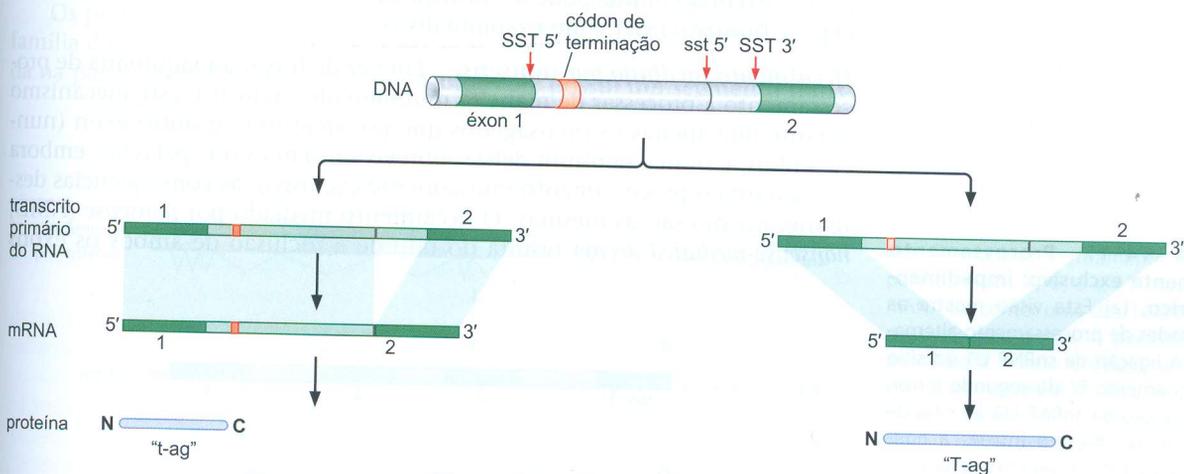


FIGURA 14-16 Processamento alternativo do antígeno T de SV40. Processamento do RNA do antígeno T de SV40. Normalmente, as duas formas de mRNA e ambas as proteínas são sintetizadas, após a infecção. O antígeno t menor é codificado pelo mais longo dos dois mRNAs; este mensageiro contém um códon de terminação em fase, a montante do éxon 2. SST 5' refere-se ao sítio de processamento 5' utilizado para gerar o mRNA de T maior; sst 5' refere-se ao sítio de processamento 5' utilizado para gerar o t menor. SST 3' é o sítio de processamento 3' utilizado para gerar ambos os mRNAs.

ou excluídos do mensageiro maduro. Estes éxons são geralmente chamados de **éxons cassete**. Em cerca de 10% dos casos, os éxons cassete ocorrem aos pares, e apenas um é incluído no mensageiro processado, assim como se viu no caso da α -troponina T (Fig. 14-14). Nestes casos, é preciso haver mecanismos que garantam que os éxons sejam processados de maneira mutuamente exclusiva.

Existem vários mecanismos para garantir o processamento mutuamente exclusivo

Existem vários mecanismos para assegurar que a seleção de éxons alternativos seja mutuamente exclusiva – ou seja, quando um é escolhido, o outro não o é (ou, para citar novamente o exemplo da α -troponina T, quando o éxon 3 é escolhido, o éxon 4 é sempre excluído, e vice-versa). Aborda-se, aqui, cada um destes mecanismos e, então, na próxima seção, discute-se um caso extremo no qual um mecanismo especial é necessário.

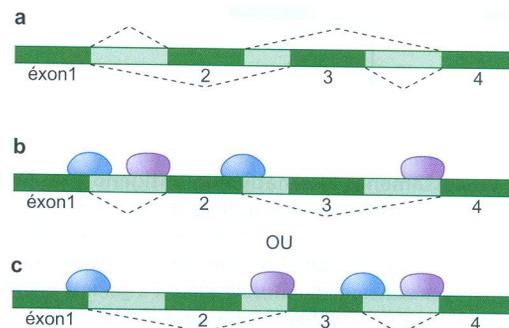
Impedimento estérico Considere dois éxons alternativos separados por um íntron. Se os sítios de processamento do íntron estiverem muito próximos, os fatores de processamento não poderão ligar-se a ambos os sítios ao mesmo tempo. Assim, a Figura 14-17 mostra um caso no qual a ligação da snRNP U1 ao sítio de processamento 5' do íntron entre dois éxons alternativos (éxons 2 e 3) impede a ligação da snRNP U2 ao ponto de ramificação do mesmo íntron (Fig. 14-17b). Alternativamente, a ligação da snRNP U2 ao ponto de ramificação exclui o uso do sítio de processamento 5' (Fig. 14-17c). O processamento dos éxons 3 e 4 da α -troponina é mutuamente exclusivo seguindo esse mecanismo.

Esse arranjo pode surgir por meio das posições relativas dos sítios de processamento de um íntron ou porque o íntron é simplesmente muito pequeno para funcionar; em *Drosophila*, qualquer íntron com menos de 59 nucleotídeos entra nesta categoria.

Combinações de sítios de processamento principais e minoritários Como visto anteriormente, há uma forma de spliceossomo chamada spliceossomo minoritário que reconhece sítios de processamento distintos dos reconhecidos pelo spliceossomo principal. Nenhum dos spliceossomos pode remover um íntron que contém uma combinação de sítios (i.e., um sítio de processamento 5' de um tipo e um 3', de outro tipo). Assim, pelo arranjo criterioso de sítios de processamento 5' e 3' reconhecidos pelos spliceossomos alternativos, a exclusão mútua pode ser alcançada, como mostrado na Figura 14-18a. O gene humano *JNK1* é um exemplo disso.

Decaimento mediado por nonsense Em vez de forçar a maquinaria de processamento a processar de maneira mutuamente exclusiva, este mecanismo garante que apenas os mensageiros que possuem um ou outro éxon (nunca ambos, e nunca nenhum deles) sobrevivam. Em outras palavras, embora não garanta o processamento mutuamente exclusivo, as consequências deste mecanismo são as mesmas. O decaimento mediado por nonsense (NMD, *nonsense-mediated decay*) resulta do fato de a inclusão de ambos os éxons

FIGURA 14-17 Processamento mutuamente exclusivo: impedimento estérico. (a) Esta visão mostra as possibilidades de processamento alternativo. (b) A ligação da snRNP U1 ao sítio de processamento 5' do segundo íntron exclui a ligação da snRNP U2 ao sítio de ramificação do mesmo íntron; a ligação de U2 ao íntron seguinte resulta na exclusão do íntron 3. (c) Aqui, a ligação da snRNP U2 ao sítio de ramificação do segundo éxon exclui a ligação de U1 ao sítio de processamento 5' do mesmo íntron. Neste caso, a ligação de U1 ao sítio de processamento 5' do primeiro íntron resulta na exclusão do éxon 2.



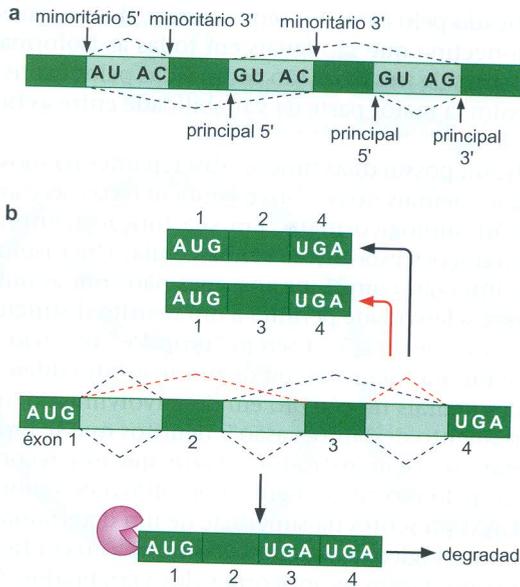


FIGURA 14-18 Processamento mutuamente exclusivo. (a) Sítios de processamento reconhecidos pelos spliceossomos principal e minoritário. (b) Decaimento mediado por nonsense.

produzir um mRNA que contém um códon de terminação prematuro (Fig. 14-18b). Estes mensageiros são destruídos pelo NMD, cujos detalhes serão descritos no Capítulo 15 (ver Fig. 15-51).

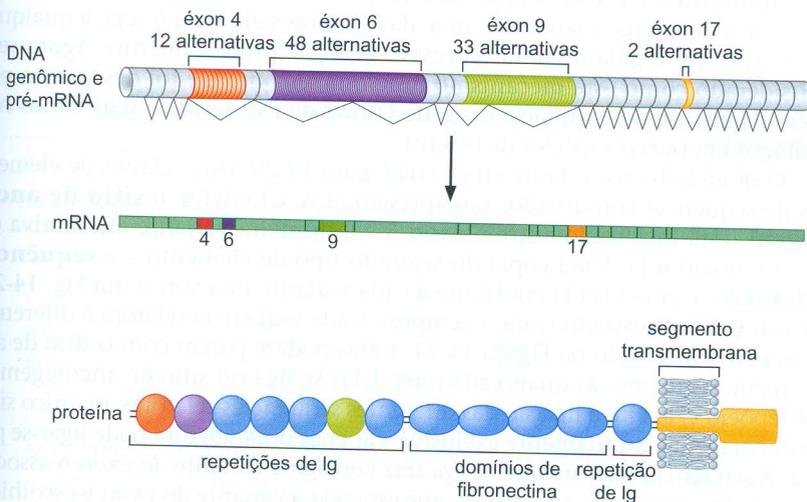
O curioso caso do gene de *Drosophila Dscam*: processamento mutuamente exclusivo em grande escala

O gene de *Drosophila Dscam* (do inglês, *Down syndrome cell-adhesion molecule* [molécula de adesão celular da síndrome de Down]) codifica potencialmente 38.016 isoformas proteicas. Como mostrado na Figura 14-19, cada mRNA possível gerado a partir do gene *Dscam* contém 24 éxons, e 20 deles são sempre os mesmos, mas quatro (éxons 4, 6, 9 e 17) existem em múltiplas formas alternativas no pré-mRNA. Assim, há 12 versões possíveis do éxon 4, 48 do éxon 6, 33 do éxon 9, e duas do éxon 17. As permutações que isso possibilita ($12 \times 48 \times 33 \times 2$) dão origem ao grande número de formas possíveis.

Os produtos do gene *Dscam* são proteínas de superfície celular da super-família das imunoglobulinas (Ig). Uma forma genérica da proteína é mostrada na parte inferior da Figura 14-19. A molécula possui um segmento trans-

FIGURA 14-19 Os múltiplos éxons do gene *Dscam* de *Drosophila*.

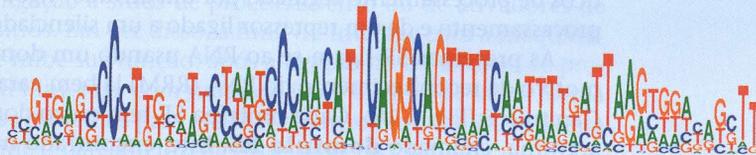
O gene *Dscam* (mostrado no topo) possui 61,2 kb; uma vez transcrito e processado, ele produz uma ou mais versões de um mRNA com 24 éxons e 7,8 kb (a figura mostra a estrutura genérica desses mRNAs). Como mostrado, há várias alternativas mutuamente exclusivas para os éxons 4, 6, 9 e 17. Assim, cada mRNA irá conter uma das 12 alternativas possíveis para o éxon 4 (vermelho), uma de 48 para o éxon 6 (roxo), uma de 33 para o éxon 9 (verde) e uma de duas para o éxon 17 (amarelo). Os éxons 4, 6 e 9 codificam partes dos três domínios de Ig, representadas nas cores correspondentes, e o éxon 17 codifica o domínio transmembrana. Se todas as combinações possíveis desses éxons forem utilizadas, o gene *Dscam* produz 38.016 mRNAs e proteínas diferentes. (Adaptada, com permissão, de Schmucker D. 2000. *Cell* 101: 671, Fig. 8. © Elsevier.)



Quadro 14-3 (Continuação)

```

6.24 GTCATTGTCGAGACTCTT-----TACATCCAATAC - TCAGGCAGT
6.47 GGCTTTCCAGTACCCATTATCAGGTTAGTCCAACA---TCGGGCATATC---CAATTAGACAGAGG
6.22 CAGCTCAATCGTATCC-----AATCCAGCTT---TTAGG---TAAACACTAAGATTTA
6.17 CAGCTGTCAGGACTT-----G-TCCAGGCTG---TTAGGCAGTAAATCG
6.34 TTCAG-CC-----TTAGACCAACA---TCAGG
6.10 GTGGGTTTCCC-----TTTCCCAATCATCAGACAGTTTT
6.4 GGTAACCC-----A-ACCCAACAG- TTAGGC
6.6 GTCAGTCCCT-----TCCCATC---TCAGGCA
6.38 GGCATTCC-----GGTCCAGCTT---TTAGGTTAATAACAATTTTTGGTT
6.33 G---CC-----TTGAACCAACA---TCAGGCTTGGTTTCGATTTCTTTTA
6.15 TGCCC-----TCCCAACAT- TCA
6.8 TCCT-----AGGCCAACAT- TCAAG
6.14 GTCGTTTCTT-ATCCAGCAT--TCAGATAGTAGATTT
6.35 TCCT-----ATCCATACAT--TCCNATGTCGGATAGATT
6.9 TACTTTAAATAAAATCCAAACACATTCAGTCAGTTG-----AATAAGGGA
6.1 TCA-----AGTCCCAATCG--TT-GGCTATCTTCACCTCTTA
6.41 CTCAGGCTTCCCGTTCCCATCAT-AGTAAAGTTTCAGCAA---CAGGCTTCTAGGTT
6.3 CC-----TATCCCAATCG--A-AGG---TGTTTC-CTTTCGA
6.31 TTGGGAATCAGTGTCAATAT-TCAGGCAGTGTATGGAGTTGTTAGAGC
6.37 GGTAA---GCCCA---CATTGAGGAGTT-----AGTAA
6.29 GGTGATTCTGCTCAGA---CATTGAGGAGTTTA-----AGTTATGGCT
6.32 TCFTTATGATTCGCCAT--TCAGACAGTT
6.39 TGTTGATAAACGCATTT--TCAGTCAGTTTTCA
6.23 GAGTCCCTGGTTCCTTTTTCATGTAGTTTT---CAT-----GGTCT
6.7 TATCCGTGATTTTCACT-TC-CCGAGTTTCAAT-----TGAGTTAGG
6.42 GCCTTGATCCCGTGTATCAGGCAGTTTCATAGAGATT
6.12 TGCTCAACTCCACATTTGGCAGATTC
6.26 CCCCATCCATTTCCACTCAGGCAGTTTTC-ACTAGACTTCGGTT
6.2 ACCCAGCCA--AC-TGCGCCAGTTTCCAAT-----TGAGATTGCTCGC
6.16 CCCTGT--CCAACATCAGGTATTTCTAGCAT-----GGC
6.19 CTCTGTCTT---GGTC-AGGCAGTTTCAAAGTCCCTTAGCTGATAGGT
6.25 TGTCGAGTT---CCTT-GGGCAGTTTTCAAT-----CTCAG--CAGGGTT
6.20 CATTGCTGAGTACAGTCAGGCAGGTTTTTCATG-AGATTGGG
6.21 CATTGTTGAGTACATCAGACGGTTTTTCCATTTACATAGAAATGTTAGAAGC
6.5 TTTATGCCCAACT--TCAGGCAG--GCTAGA
6.36 ACCCCGCAAGCAGTTCAGTCAGTTTAT-----TGTTGCGTTTAGCT
6.13 AACACCA--T--TCAGACTGTTTTTTTATG
6.27 ATCCCAAGTTCCTCAGGCAGA--C-----TTCAAATGA--CTT
6.28 ATCCCTAGCGATTTAGGCAGTTTTCCGGTTTACTTAG
6.44 ACCCAAACCTATTCAAGCAGTTA-----ATTAAGCGAC
6.40 ACCGTG---T--TCA-ACTGTTCTGTTAGGTTCAAATAGA
6.43 TTTAGCATCAGGCAGTTTTC
6.30 CCAAGATTCGGGAGTTTTCAAT
6.48 CCAAGATTCGGGAGCAAATAGCAT--TTCAAATAGGATCTTA
6.45 GTCAGGCAG-----CTTAGAAGGC--T
6.18 TTCAGGCAGTTTT-----CTTAAAAGA--TCTTCTTAAGCA
6.11 GGGCGTATTCGAA-----TTCAG--GAC
6.46 AGGTAGCCAATA-----AGTAGA---CTTA
    
```

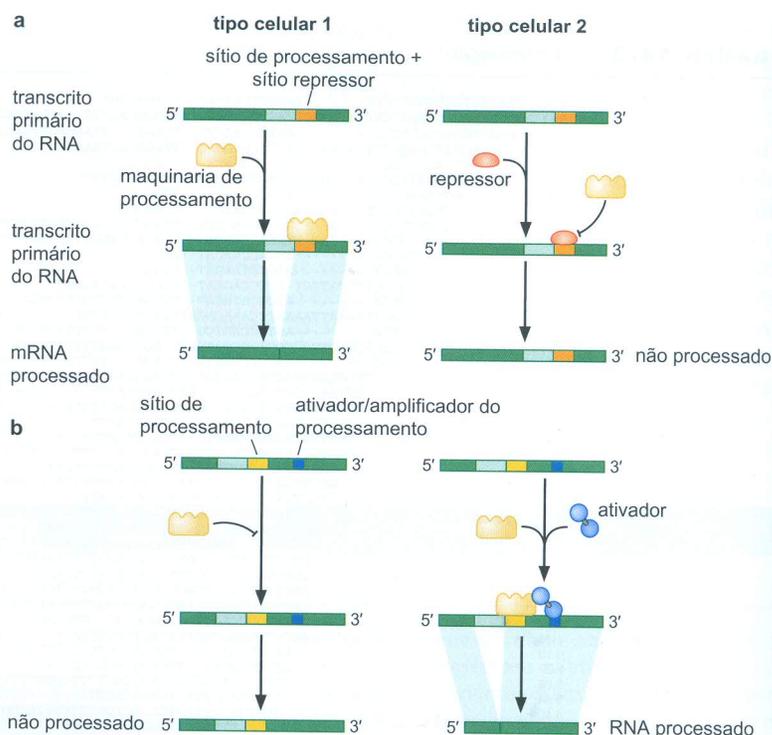


QUADRO 14-3 FIGURA 2 Sequência seletora de consenso de *D. melanogaster*. (Painel superior) As 48 seqüências seletoras e seqüências flanqueadoras foram alinhadas. Os nucleotídeos mais frequentes na porção central do alinhamento estão realçados. (Painel inferior) O alinhamento foi utilizado para gerar uma seqüência seletora de consenso. (Reproduzida, com permissão, de Graveley B.R. 2005. *Cell* 123: 65-73, Fig. 4. © Elsevier.)

O processamento alternativo é regulado por ativadores e repressores

Proteínas que regulam o processamento ligam-se a sítios específicos chamados **amplificadores de processamento exônico** (ou **intrônico**) (**ESE** ou **ISE**) ou **silenciadores** (**ESS** e **ISS**). Os primeiros potencializam, e os segundos reprimem, a remoção de íntrons nas proximidades dos sítios de processamento. Os ativadores e as proteínas SR que se ligam a eles já foram mencionadas (Fig. 14-11). Esses elementos e proteínas são importantes no direcionamento da maquinaria de processamento para muitos éxons, mesmo quando não há envolvimento de processamento alternativo. Além disso, no exemplo do processamento do antígeno T (Fig. 14-16), a proteína SR era a responsável pela ocorrência do processamento alternativo. Entretanto, essa família de proteínas, grande e diversificada, também desempenha funções específicas no processamento alternativo *regulado*, porque direciona a maquinaria de processamento a diferentes sítios de processamento, sob condições diferentes. Desse modo, a presença ou atividade de determinada proteína SR pode determinar a utilização

FIGURA 14-22 Processamento alternativo regulado. (a) Alguns éxons alternativamente processados aparecem em mRNAs, a menos que sejam impedidos de fazê-lo por uma proteína repressora. (b) Outros aparecem apenas se um ativador específico promover a sua inclusão. Ambos os mecanismos podem ser utilizados para regular o processamento de maneira que, em um tipo celular, um determinado éxon seja incluído em um mRNA, enquanto outro, não.



de um dado sítio de processamento em um tipo celular específico ou em uma determinada fase do desenvolvimento. A Figura 14-22 apresenta casos hipotéticos de processamento regulado por um ativador ligado a um amplificador de processamento e de um repressor ligado a um silenciador de processamento.

As proteínas SR ligam-se ao RNA usando um domínio – por exemplo, o motivo de reconhecimento do RNA (RRM) já bem caracterizado e descrito no Capítulo 6 (Fig. 6-18). Cada proteína SR tem outro domínio, rico em serina e arginina, chamado **domínio RS**. Este, encontrado na porção carboxiterminal da proteína, promove as interações entre a proteína SR e as proteínas da maquinaria de processamento, recrutando essa maquinaria para um sítio de processamento próximo.

Um exemplo de ativador que promove um determinado evento de processamento alternativo em um tecido específico é a proteína Half-pint de *Drosophila*. Este ativador regula o processamento alternativo de um conjunto de pré-mRNAs no ovário da mosca. Ele atua pela ligação a sítios próximos ao sítio de processamento 3' de éxons específicos nesses pré-mRNAs e pelo recrutamento do fator de processamento U2AF.

A maioria dos silenciadores é reconhecida por membros da família de ribonucleoproteínas nucleares heterogêneas (hnRNPs). Estas ligam-se ao RNA, mas não possuem domínios RS e, portanto, não podem recrutar a maquinaria de processamento. Em vez disso, elas bloqueiam sítios específicos de processamento, impedindo sua utilização. Viu-se uma função como essa no exemplo anterior de *Dscam*, onde Hrp36 inibe a inclusão de variantes do éxon 6 no mRNA. Outro exemplo é a hnRNPA1, que se liga a um elemento silenciador exônico localizado no interior de um éxon do pré-RNA de *tat* do vírus da imunodeficiência humana (HIV), impedindo a inclusão desse éxon no mRNA final. Esse repressor ligado bloqueia a ligação do ativador SC35 (uma proteína SR) a um elemento ativador próximo. Esse bloqueio não é direto – os dois sítios de ligação não são sobrepostos –, mas o hnRNPA1 promove a ligação cooperativa de outras moléculas de hnRNPA1 a sequências adjacentes, estendendo-se sobre o sítio do ativador (Fig. 14-23). No caso de outra proteína SR (SF2/ASF) estar presente, essa repressão pode ser evitada, porque SF2/ASF

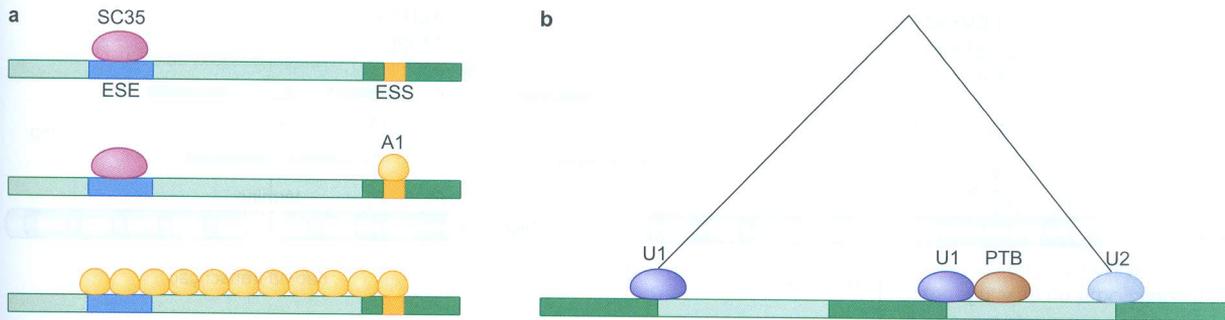


FIGURA 14-23 Dois mecanismos de ação do silenciador. (a) Mecanismo de exclusão do éxon 3 do *tat* de HIV por hnRNP1. O amplificador do processamento SC35 liga-se ao ESE e promove a inclusão do éxon. A1 liga-se ao ESS do éxon e, de lá, através da ligação cooperativa obstrui o ESE e impede a ligação de SC35 por competição. (b) Mecanismo de exclusão de éxon pela proteína hnRNP1 (PTB). A PTB liga-se a um éxon e interage com U1 no sítio de processamento 5', como descrito no texto. Esta interação bloqueia a habilidade de U1 interagir com os componentes do sítio de processamento 3' e, assim, U1 do éxon a montante pareia com U2 do éxon a jusante.

tem afinidade maior com a sequência de amplificação/ativação do que SC35 e, assim, desloca os repressores ligados nas proximidades. Nos Capítulos 18 e 19, serão vistas situações semelhantes envolvendo ligações cooperativas e competitivas em exemplos de regulação da transcrição.

Outro repressor de processamento em mamíferos é a proteína hnRNP1 (ou proteína de ligação ao trato Py, PTB). Essa proteína exclui um determinado éxon do mRNA maduro ao ligar-se a sequências adjacentes ao éxon. Mas o mecanismo pelo qual este opera não é por competição com os componentes do spliceossomo pela ligação a sítios de processamento ou de ativadores de processamento a seus sítios. Em vez disso, a hnRNP1 interage com a maquinaria de processamento e inibe sua função: após a ligação de U1 ao sítio de processamento 5', a hnRNP1 interage com uma região de U1 que de outra forma iria interagir com outras proteínas para facilitar o pareamento de éxon. Desta maneira, hnRNP1 previne o pareamento de éxon (Fig. 14-23b).

A regulação do processamento alternativo determina o gênero das moscas

Agora será considerado um exemplo particularmente elaborado de processamento alternativo regulado – o que envolve o gene *double-sex* de *Drosophila*. O sexo de cada mosca depende de qual entre duas variantes de processamento alternativo do mRNA é produzida.

O sexo da mosca é determinado pela relação/proporção entre cromossomos X e autossomos. A fêmea resulta de uma proporção de 1 (dois X e dois conjuntos de autossomos) e o macho, de uma relação de 0,5. Inicialmente, essa relação é medida no nível da transcrição usando dois ativadores, chamados SisA e SisB (os mecanismos de regulação da transcrição são considerados em detalhe nos Cap. 18 e 19). Como ambos os genes que codificam esses reguladores estão no cromossomo X, no embrião precoce, a futura fêmea prepara duas vezes mais produtos do que o macho (Fig. 14-24).

Esses ativadores ligam-se a sítios da sequência reguladora a montante do gene *Sex-lethal* (*Sxl*). Outro regulador que se liga e controla o gene *Sxl* é um repressor chamado Dpn (Deadpan); este é codificado por um gene encontrado em um dos autossomos (cromossomo 2). Portanto, a relação entre os ativadores e os repressores difere nos dois sexos, e isso determina a diferença entre a ativação (nas fêmeas) e a repressão (nos machos) do gene *Sxl*.

O gene *Sxl* é expresso a partir de dois promotores, P_e e P_m . O primeiro (promotor para estabelecimento) é controlado por SisA e SisB (e, portanto, ex-

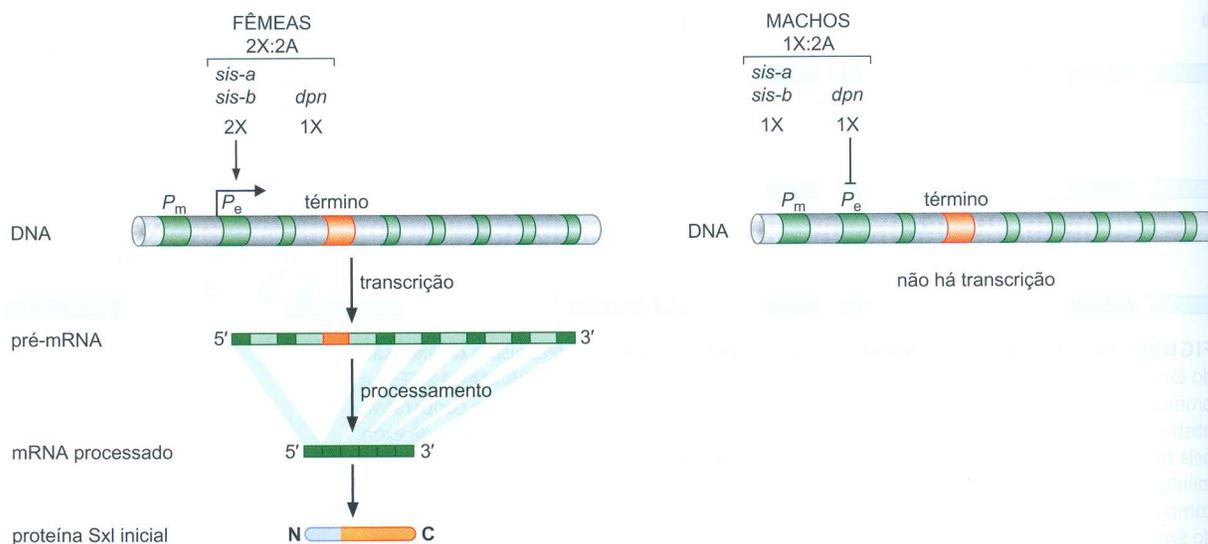


FIGURA 14-24 Regulação transcricional precoce de *Sxl* em moscas machos e fêmeas. Os genes *sisA* e *sisB* são encontrados no cromossomo X e codificam ativadores transcricionais que controlam a expressão do gene *Sxl*. *Dpn*, um repressor de *Sxl*, é codificado por um gene do cromossomo 2. Embora machos e fêmeas expressem a mesma quantidade de *Dpn* codificado de modo autossômico, as fêmeas produzem o dobro de ativadores dos machos (porque as fêmeas possuem dois cromossomos X e os machos, apenas um). A diferença na proporção entre ativadores e repressor assegura a expressão de *Sxl* nas fêmeas, mas não nos machos. Então, a proteína *Sxl* autorregula sua expressão, como descrito no texto e na próxima figura. (Adaptada de Estes P.A. et al. 1995. *Mol. Cell. Biol.* 15: 904-917.)

presso apenas em fêmeas). Mais adiante no desenvolvimento, esse promotor é desligado permanentemente. Nos embriões de fêmeas, a expressão de *Sxl* é mantida pela expressão a partir do P_m (promotor de manutenção).

A transcrição a partir de P_m é constitutiva em fêmeas e machos, mas o RNA produzido a partir desse promotor contém um éxon a mais do que o transcrito produzido a partir de P_e (Fig. 14-23). Se este éxon permanecer no mensageiro maduro, ele não conseguirá produzir uma proteína ativa, que é o que ocorre no macho. Na fêmea, porém, o processamento remove esse éxon e a proteína funcional *Sxl* continua a ser produzida.

Como demonstrado na Figura 14-25, é a própria proteína *Sxl*, presente nas fêmeas, mas não nos machos (graças à expressão precoce a partir de P_e), que promove o processamento do RNA feito a partir de P_m e garante que o éxon inibidor seja excluído da cadeia. *Sxl* realiza essa função atuando como repressor do processamento.

Assim, a proteína *Sxl* funcional continua a ser produzida nas fêmeas. Essa proteína regula o processamento de outros RNAs na fêmea, além de seu próprio RNA. Um deles é o RNA produzido constitutivamente (em machos e fêmeas) pelo gene *tra* (Fig. 14-25). Novamente, na ausência do processamento dirigido por *Sxl*, esse RNA não produz proteínas (em machos), mas na presença de *Sxl*, ele é processado para formar a proteína funcional *Tra* (em fêmeas).

A proteína *Tra* também é um regulador de processamento. Enquanto *Sxl* é repressora do processamento, *Tra* é ativadora (Fig. 14-25). Um de seus alvos é o RNA produzido a partir do gene que codifica *Doublesex* (*Dsx*). Esse RNA é processado de duas formas alternativas, ambas codificadoras de proteínas reguladoras, mas com atividades diferentes. Assim, na presença de *Tra*, o RNA de *dsx* é processado de maneira a originar uma proteína que reprime a expressão de genes específicos de machos. Na ausência da proteína *Tra*, a forma de *Dsx* produzida reprime os genes específicos de fêmeas.

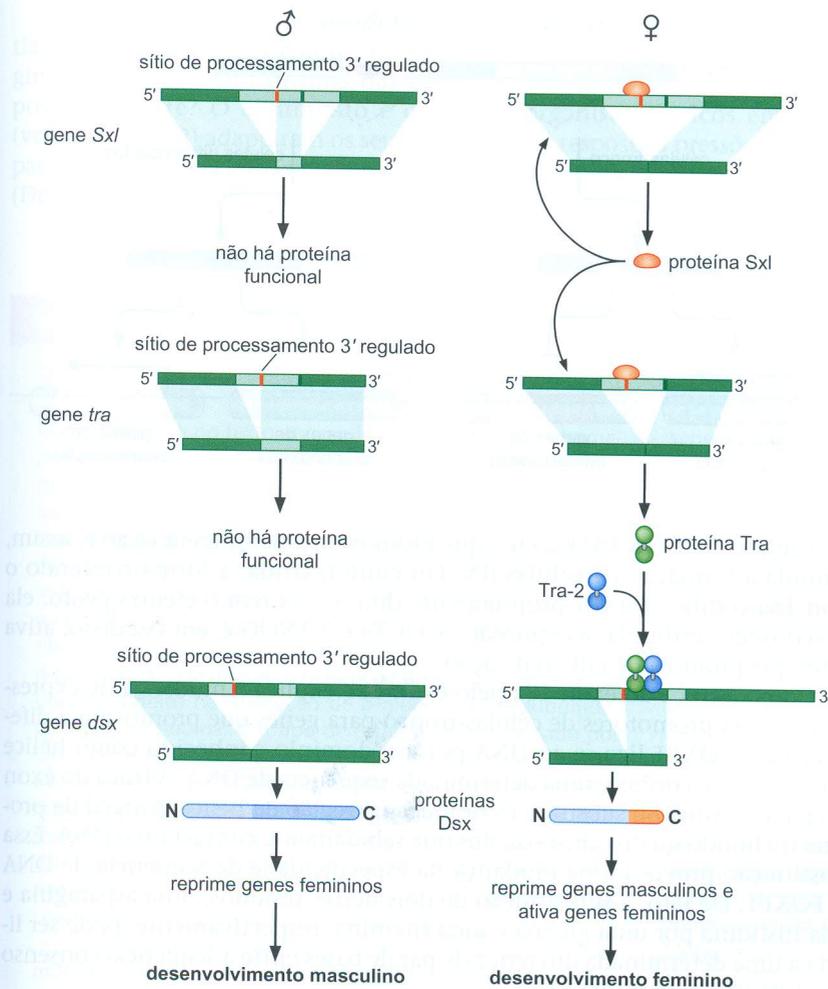


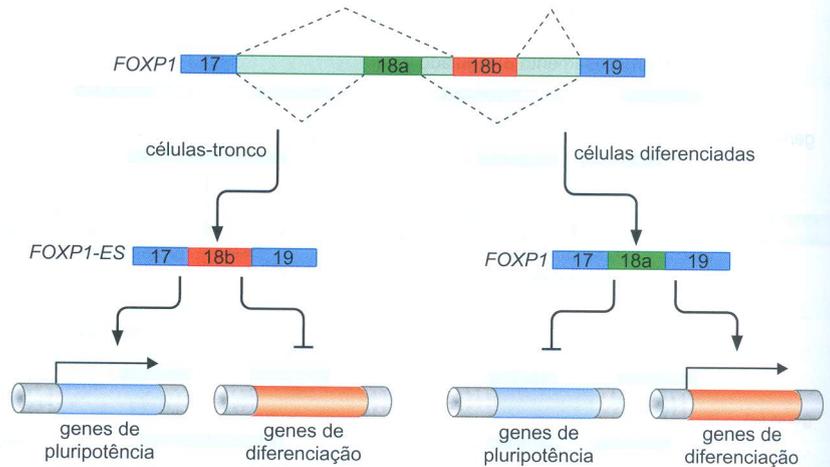
FIGURA 14-25 Uma cascata de eventos de processamento alternativo determina o sexo de uma mosca.

Como descrito de maneira detalhada no texto, a proteína Sex-lethal (Sxl) é produzida em moscas que se desenvolvem em fêmeas (apresentadas à direita da figura), mas não nas que originam machos (apresentadas à esquerda). A presença dessa proteína é mantida pela autorregulação do processamento de sua mensagem. Na ausência da regulação, a proteína funcional não é produzida (nos machos). Sxl também controla o processamento do gene *tra*, produzindo uma proteína funcional Tra nas fêmeas (mas não nos machos). A própria Tra é um regulador de processamento, e atua no pré-mRNA do gene *doublesex* (*dsx*). Quando o mRNA de *dsx* é processado em resposta à proteína Tra, uma versão da proteína Dsx é produzida (em fêmeas) com um trecho de 30 aminoácidos em sua extremidade carboxiterminal que a distingue da forma da proteína produzida na ausência da proteína reguladora Tra (em machos). A forma feminina de Dsx ativa os genes necessários ao desenvolvimento feminino e reprime os genes de desenvolvimento masculino. A forma masculina, que tem um segmento de 150 aminoácidos na extremidade carboxiterminal, reprime os genes que promovem o desenvolvimento feminino. A proteína Sxl atua como repressora do processamento ao ligar-se ao trato de pirimidina no sítio de processamento 3' (ver Fig. 14-3). A proteína Tra, em contrapartida, atua como amplificadora do processamento. Ela liga-se a uma sequência reforçadora (*enhancer*) em um dos éxons do RNA de *dsx* (ver Fig. 14-11).

Uma alteração de processamento alternativo está no cerne da pluripotência

Nenhum assunto gerou mais interesse no campo da biologia do desenvolvimento de mamíferos ou chamou mais a atenção por potenciais avanços em medicina do que o uso de células-tronco embrionárias para gerar tipos celulares especializados no laboratório. As células-tronco embrionárias são células indiferenciadas encontradas no embrião que dão origem a tecidos e tipos celulares do animal adulto e, assim, são chamadas de pluripotentes (ver Apêndice 1). De fato, as células-tronco embrionárias podem ser induzidas em laboratório a se diferenciar em tecidos muscular, nervoso, pancreático e outros tipos celulares especializados. Recentemente, tornou-se possível desencadear a desdiferenciação de células somáticas de volta ao estágio de células pluripotentes (as chamadas células-tronco pluripotentes induzidas ou **células iPS**) ao amplificar artificialmente a produção de um número relativamente pequeno de fatores de transcrição essenciais, sabidamente necessários para a manutenção da pluripotência (Quadro 21-1). Exemplos de tais fatores de transcrição incluem OCT4 e NANOG. A produção de OCT4 e NANOG (e outras proteínas reguladoras essenciais) é estimulada por um fator de transcrição conhecido como FOXP1, membro da família Forkhead de proteínas de ligação ao DNA. Na verdade, existem duas isoformas da proteína FoxP1. Uma delas é o produto de um mRNA processado para incluir um determinado éxon, o 18b, enquanto a outra variante de processamento carrega o éxon 18a (Fig. 14-26). A proteína codificada pelo mRNA que contém o éxon 18b é chamada de FOXP1-ES: ela

FIGURA 14-26 Uma alteração do processamento alternativo está no cerne da pluripotência. O fator de transcrição FOXP1 liga-se a sítios específicos do DNA por meio de um domínio conhecido como hélice Winged. A especificidade do sítio de DNA que ele reconhece pode ser alterada pelo processamento alternativo de seu mRNA. Assim, quando um determinado éxon é incluído (éxon 18a), a isoforma resultante (FOXP1) reconhece sítios de uma sequência, mas quando o processamento alternativo gera uma forma que inclui o éxon 18b em vez do 18a, a isoforma resultante (FOXP1-ES) liga-se a uma sequência diferente de DNA e, assim, controla a expressão de um conjunto diferente de genes.



ativa genes (OCT4, NANOG, etc.) que promovem a desdiferenciação e, assim, estimula a formação de células iPS. Em contrapartida, a forma contendo o éxon 18a codifica FOXP1 propriamente dita, e esta tem o efeito oposto: ela não consegue estimular a expressão de OCT4 e NANOG e, em vez disso, ativa genes que promovem a diferenciação.

Como a troca do éxon 18a pelo éxon 18b explica a mudança de expressão de genes promotores de células-tronco para genes que promovem a diferenciação? FOXP1 liga-se ao DNA por um domínio conhecido como hélice Winged, que reconhece uma determinada sequência de DNA. A troca do éxon 18a para o éxon 18b substitui 35 resíduos na região da hélice Winged da proteína incluindo quatro aminoácidos que sabidamente contactam o DNA. Essa substituição provoca uma mudança na especificidade de sequência de DNA de FOXP1. De fato, a substituição de dois destes resíduos, uma asparagina e uma histidina por uma glicina e uma treonina, respectivamente, pode ser ligada a uma determinada diferença de par de bases entre a sequência consenso para FOXP1 e a para FOXP1-ES.

Assim, no topo da hierarquia de eventos que controlam os circuitos transcripcionais para pluripotência e diferenciação, está uma troca entre formas de processamento alternativo do mRNA para FOXP1. Esta situação nos lembra o processamento alternativo do gene *Sex-lethal*, que, como visto, dispara uma cadeia de eventos que determina o sexo das moscas. No caso do *Sex-lethal*, sabe-se como a proporção entre cromossomos X e autossomos dita qual variante de processamento é produzida em machos e em fêmeas. No caso de FOXP1, entretanto, ainda não foram descobertos os fatores que determinam qual forma é produzida nas células-tronco embrionárias ou nas células em diferenciação. Ainda assim, esses exemplos de moscas e mamíferos ressaltam o papel do processamento alternativo em uma ampla e crescente variedade de eventos reguladores gênicos.

Nós consideramos as várias formas pelas quais o processamento é realizado em sistemas eucarióticos, bem como a diversidade de componentes envolvidos. A perda de função em qualquer um destes componentes pode levar a graves consequências, como é descrito no Quadro 14-4, Defeitos no processamento do pré-mRNA causam doenças humanas.

EMBARALHAMENTO DE ÉXONS

Os éxons são embaralhados por recombinação, produzindo genes que codificam novas proteínas

Como já mencionado, todos os eucariotos possuem íntrons; no entanto, estes são raros – quase inexistentes – em bactérias. Existem duas explicações plausíveis para essa situação.