

O velho axioma “um gene, uma proteína” não faz mais sentido. Quanto mais complexo o organismo, maior a probabilidade de ter se tornado assim ao sintetizar várias proteínas a partir de um único gene

Genoma Alternativo

Por Gil Ast

O GENOMA do camundongo e o do homem são 88% semelhantes. Muitas das diferenças vêm do modo como nós, humanos, editamos a informação existente em nossos genes



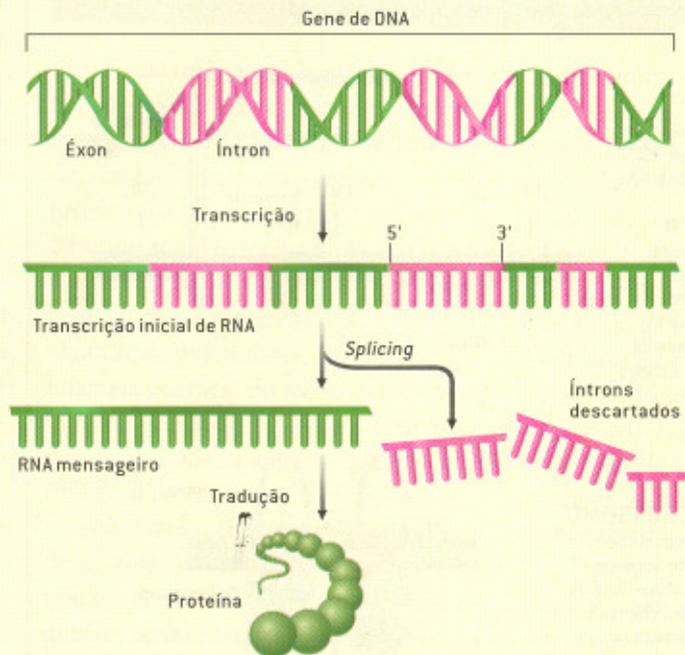
UM GENE, MUITAS PROTEÍNAS

A visão clássica da expressão dos genes era simples: um gene de DNA é inicialmente transcrito em forma de RNA, depois o mecanismo de *splicing* celular remove os trechos de "lixo" chamados íntrons e liga as partes significativas, os éxons, em uma versão final de RNA mensageiro, que então é traduzida para uma proteína. Mas agora

está claro que essas regras nem sempre se aplicam. Em organismos complexos, a transcrição inicial de RNA pode sofrer *splicing* alternativo – os éxons podem ser descartados, e os íntrons, ou partes deles, mantidos – para produzir múltiplos RNA mensageiros, e portanto proteínas diferentes, a partir de um mesmo gene.

EXPRESSIONE DE GENES CLÁSSICA

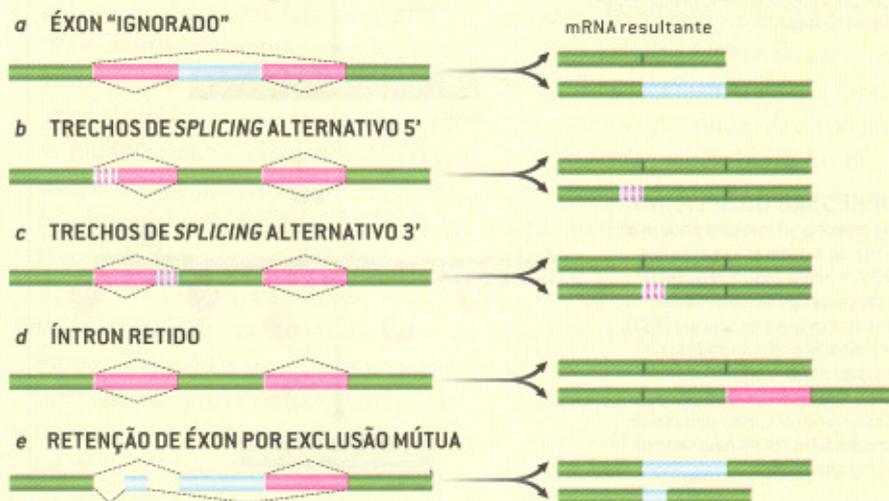
Uma sequência de DNA é transcrita numa cópia de uma só fita feita de RNA. O mecanismo celular então faz o *splicing* nessa transcrição inicial: os íntrons, cada qual definido por sequências particulares de nucleotídeos em seu princípio e seu fim, conhecidas respectivamente como pontos de *splicing* 5' (cinco-linha) e 3' (três-linha), são removidos e descartados, enquanto os éxons são ligados numa versão de RNA mensageiro (mRNA) do gene, que será traduzido numa proteína pela célula



SPLICING ALTERNATIVO

A transcrição inicial do gene pode ser editada de várias formas (à direita), em que a atividade de *splicing* está representada em linhas tracejadas. Um éxon pode ser deixado de fora (a). O mecanismo de *splicing* pode reconhecer os pontos de *splicing* alternativos 5' para um íntron (b) ou os pontos de *splicing* 3' (c). Um íntron pode ser mantido na transcrição final de mRNA (d). E éxons podem ser mantidos num sistema de exclusão mútua (e)

- Éxon que sempre sofre *splicing*
- Éxon de *splicing* alternativo
- Íntron



proteína a partir de um mesmo gene foi feita há 25 anos, mas o fenômeno era considerado raro. Comparações recentes mostram que ele não só é comum como fundamental, em mais uma drástica reviravolta da visão clássica de como a informação armazenada no gene se traduz em proteína. A maioria dos conhecimentos básicos ainda se sustenta: os genomas contêm todas as

instruções necessárias para a criação e a manutenção do organismo, codificadas em linguagem formada por nucleotídeos de DNA de quatro letras (A, G, C e T). Nos cromossomos humanos, 3 bilhões de nucleotídeos estão ligados entre si em cada uma das duas fitas complementares, para formar a dupla hélice. Quando chega o momento de as instruções do gene ser

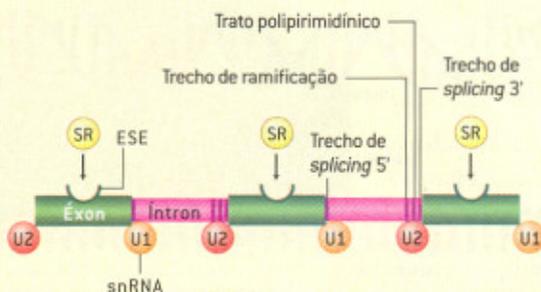
"expressas", o zíper de duas fitas do DNA se abre apenas pelo tempo suficiente para que a cópia de uma fita da sequência do gene possa ser produzida por um primo químico, o RNA. Cada sequência de nucleotídeos de DNA que se transcreve na versão de RNA dessa maneira é chamada de gene. Algumas das moléculas de RNA resultantes acabam nunca se traduzindo

MÁQUINA DE SPLICING

Quando a transcrição inicial em RNA de um gene é criada, uma estrutura chamada spliceossomo executa a edição do RNA. Em organismos complexos, esse processo é controlado por proteínas reguladoras do *splicing* (SR), que definem os éxons e encaminham os spliceossomos para pontos de *splicing* específicos. Essas moléculas reguladoras têm, portanto, o poder de determinar quando e como o *splicing* alternativo de um gene ocorrerá. As proteínas SR são produzidas, elas próprias, de formas diferentes em tecidos e tipos de célula diversos, em estágios variados do desenvolvimento do mesmo tecido.

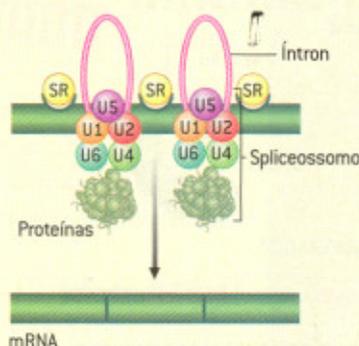
DEFINIÇÃO DO ÉXON

Uma proteína SR se liga a cada éxon na transcrição em uma sequência particular de nucleotídeos chamada ativador exônico de *splicing* (ESE). A ligação da proteína SR define o éxon para o mecanismo de processamento, ao recrutar moléculas de RNA nucleares de pequeno peso molecular (snRNA) chamadas U1 e U2 para o ponto de *splicing* em ambas as extremidades a íntrons adjacentes



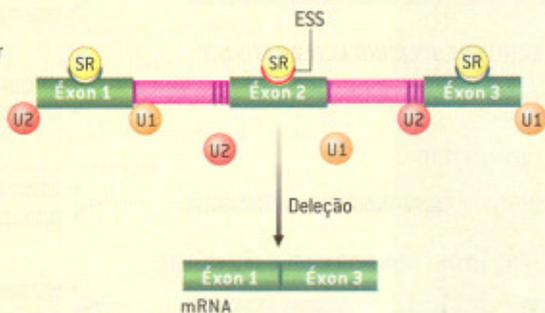
FORMAÇÃO DO SPLICEOSSOMO

Quando os snRNA originais reconhecem os pontos de *splicing* do íntron, formam um complexo com outros snRNA e mais de cem proteínas. Esse complexo, chamado spliceossomo, elimina os íntrons e une os éxons para produzir um RNA mensageiro maduro (mRNA)



SUPRESSÃO DO SPLICING

Uma proteína SR também pode inibir em vez de aumentar a ligação dos snRNA, e nesse caso a sequência à qual ela se liga é chamada de supressor exônico de *splicing* (ESS). A proteína SR portanto pode fazer com que um éxon seja deixado de fora do mRNA final. Em humanos e outros mamíferos, essa deleção de éxons é a forma mais freqüente de *splicing* alternativo



em proteínas, mas realizam funções regulatórias e de manutenção dentro do núcleo. As transcrições de RNA dos genes que realmente codificam proteínas são lidas por mecanismos celulares e traduzidas na sequência correspondente de aminoácidos. Mas, primeiro, a transcrição preliminar tem de passar por um processo de edição.

Em 1977, Phillip Sharp, do Instituto de Tecnologia de Massachusetts (MIT), e

colegas descobriram que essas transcrições de RNA iniciais ou primárias são como livros que contêm muitos capítulos sem nenhum significado, inseridos em intervalos do texto. Os capítulos sem sentido, chamados íntrons, precisam ser extraídos para que os capítulos com sentido se liguem, permitindo ao RNA "contar uma história" coerente. No *splicing*, processo de corte-e-ligação ou processamento, os íntrons são removidos da

transcrição inicial. Os segmentos da transcrição que contêm sequências que codificam proteínas, chamados éxons, são ligados para formar a versão final da transcrição, conhecida como RNA mensageiro (mRNA) (ver quadro na pág. 53).

Mas em 1980, Randolph Wall, da Universidade da Califórnia em Los Angeles (UCLA), já havia mostrado que essa visão básica do processamento pré-mRNA, em que todos os íntrons são sempre descartados e todos os éxons são sempre incluídos, nem sempre é verdadeira. Na realidade, os mecanismos celulares podem "decidir" remover um éxon ou manter um íntron, ou partes dele, na transcrição final do mRNA. Essa capacidade de editar de forma alternativa as transcrições de mRNA aumenta significativamente a versatilidade dos genes e dá ao mecanismo de processamento o enorme poder de determinar o quanto a célula produzirá de certo tipo de proteína em detrimento de outros, codificados pelo mesmo gene.

Em 1984, Tom Maniatis, da Universidade Harvard, desenvolveu um procedimento para revelar o mecanismo molecular que executa a retirada dos íntrons e a ligação dos éxons entre si. Os detalhes sobre seu funcionamento e sobre o sistema regulatório que o controla ainda estão sendo desvendados, mas a pesquisa já revelou que se trata de um sistema primorosamente intrincado.

Máquina de Emendar

EM ORGANISMOS COMPLEXOS, dois níveis diferentes de equipamento molecular estão envolvidos no *splicing* das transcrições pré-mRNA. O chamado mecanismo basal, encontrado em todos os organismos cujos genomas possuem íntrons, foi bastante conservado ao longo da evolução, dos fungos até os seres humanos. Ele consiste em cinco moléculas de RNA nuclear de pequeno peso molecular (snRNA) identificadas como U1, U2, U4, U5 e U6. Essas moléculas se aliam a até 150 proteínas para formar um complexo chamado spliceossomo, responsável por reconhecer os locais onde os íntrons começam e terminam,

removendo-os da transcrição pré-mRNA e ligando os éxons para formar o mRNA.

Quatro curtas seqüências de nucleotídeos dentro dos íntrons servem como sinal para indicar ao spliceossomo onde cortar (ver quadro na pág. 54). Um desses sinais fica no começo do íntron e é chamado de ponto de *splicing* 5' (cinco-linha); os outros, localizados na extremidade final do íntron, são conhecidos como ponto de ramificação, trato polipirimidínico e, finalmente, ponto de *splicing* 3' (três-linha).

Um outro sistema regulatório controla o processamento, direcionando o mecanismo basal para esses pontos de *splicing*. Já foram identificadas mais de dez proteínas reguladoras de *splicing* (SR) diferentes, e cada uma delas pode ser expressa em estágios distintos do desenvolvimento dentro do mesmo tecido. As proteínas SR podem se ligar a seqüências curtas de nucleotídeos localizadas dentro dos éxons da transcrição pré-mRNA. Esses pontos de ligação são chamados de ativadores exônicos de *splicing* (ESE, na sigla em inglês) porque, quando a proteína SR adequada se liga a um ESE, essa ação recruta os snRNA do mecanismo basal para os pontos de *splicing* adjacentes a cada extremidade do éxon. E a proteína SR também pode se ligar a uma seqüência supressora exônica de *splicing* (ESS) dentro do éxon, anulando a capacidade do mecanismo basal de se ligar às extremidades daquele éxon, o que resulta em sua remoção do mRNA final.

As conseqüências de deixar de fora um único éxon podem ser dramáticas para um organismo. Na mosca-das-frutas, por exemplo, o *splicing* alternativo controla o sistema que determina o sexo. A deleção (remoção) de éxons é o tipo mais comum de *splicing* alternativo entre os mamíferos, respondendo por cerca de 38% desse tipo de processamento genético em humanos. Mas



As conseqüências de excluir um único éxon podem ser dramáticas para um organismo

várias outras formas também já foram identificadas, como uma que faz os íntrons ser mantidos no mRNA maduro, mais comum em plantas e em formas inferiores de vida multicelular. É provável que a retenção de íntrons seja em termos evolutivos a versão mais antiga do *splicing* alternativo. Mesmo hoje em dia, o mecanismo de *splicing* de organismos unicelulares, como os fungos, funciona por meio do reconhecimento de íntrons, sistema diferente das proteínas SR dos organismos superiores, que define os éxons para o mecanismo basal.

No sistema unicelular, o mecanismo de *splicing* só reconhece seqüências intrônicas de menos de 500 nucleotídeos, o que funciona bem para os fungos, que têm poucos íntrons, em média com apenas 270 nucleotídeos de comprimento. Mas, conforme os genomas foram se expandindo ao longo da evolução, seus trechos intrônicos se multiplicaram e cresceram, e o mecanismo celular de *splicing* provavelmente foi obrigado a mudar de um sistema que reconhecia seqüências intrônicas curtas dentro dos éxons para um que reconhecesse éxons curtos no meio de um mar de íntrons. Em média, um gene humano dos que codificam proteínas tem 28 mil nucleotídeos, com 8,8 éxons separados por 7,8 íntrons. Os éxons são relativamente curtos, em geral têm cerca de 120 nucleotídeos, enquanto os íntrons variam de 100 a 100 mil nucleotídeos de comprimento.

O tamanho e a quantidade dos íntrons humanos – temos o maior número de íntrons por gene de todos os organismos

– suscitam uma questão interessante. Manter os íntrons é um hábito caro. Uma grande proporção da energia que gastamos todo dia é dedicada à manutenção e ao reparo de íntrons na forma de DNA, à transcrição do pré-mRNA e à remoção dos íntrons, e até a sua ruptura ao final da reação de *splicing*. Além disso, o sistema pode causar equívocos com conseqüências graves. Cada erro no corte e na ligação do pré-mRNA leva a uma alteração na seqüência de codificação da proteína da transcrição do gene, e possivelmente à síntese de uma proteína defeituosa.

Por exemplo, uma doença hereditária que investigo, a disautonomia familiar, ou Síndrome de Riley Day, resulta de uma mutação em um único nucleotídeo do gene chamado IKBKAP, que faz com que ele sofra *splicing* alternativo em tecidos do sistema nervoso. A redução da disponibilidade da proteína padrão IKBKAP decorrente disso causa desenvolvimento anômalo do sistema nervoso. Pelo menos 15% das mutações que provocam doenças genéticas e provavelmente também determinados tipos de câncer fazem isso ao afetar o *splicing* do pré-mRNA. Então, por que a evolução preservou um sistema tão complicado e que ainda pode causar doenças? Talvez porque as vantagens superem os riscos.

Vantagens nas Alternativas

COMO GERA MAIS DE UM tipo de molécula de mRNA e, portanto, mais de uma proteína por gene, o *splicing* alternativo permite aos humanos produzir mais de 90 mil proteínas sem precisar manter 90 mil genes. Em média, cada um de nossos genes dá origem a três mRNA por meio do *splicing* alternativo. Mesmo assim, esse número não explica nossa necessidade de ter tantos íntrons nem por que eles ocupam tanto território dentro dos

OAUTOR

GIL AST é professor do Departamento de Genética Humana e Medicina Molecular da Faculdade de Medicina da Universidade de Tel Aviv, em Israel. Sua pesquisa se concentra na mecânica molecular do *splicing* do pré-mRNA, na evolução e na regulação do *splicing* alternativo e nos defeitos de *splicing* associados a cânceres e doenças hereditárias. Recentemente, colaborou com cientistas da Compugen no desenvolvimento de um sistema de bioinformática para prever o *splicing* alternativo, de forma a detectar novas variantes de proteínas.

genes, de modo que as seqüências exônicas correspondam a apenas entre 1% e 2% do genoma humano.

Depois de as equipes de seqüenciamento terem revelado esse panorama aparentemente desolador em 2001, outro enigma surgiu quando o genoma do camundongo foi publicado, em 2002. Revelou-se que esse roedor possui quase o mesmo número de genes do homem. Embora tenham se passado 100 milhões de anos desde que ambos tiveram um ancestral comum, 99% dos genes tanto do camundongo quanto dos humanos derivam daquele ancestral. A maioria possui a mesma organização em íntrons e éxons, e as seqüências de nucleotídeos dentro de seus éxons também são altamente conservadas. Portanto, a pergunta passa a ser: se há diferenças tão pequenas entre o genoma dos humanos e o dos camundongos, o que é que nos torna tão diferentes dos roedores?

Christopher J. Lee e Barmak Modrek, da UCLA, recentemente revelaram que um quarto dos éxons que sofrem *splicing* alternativo em ambos os genomas são específicos ou dos seres humanos ou dos camundongos. Esses éxons têm, portanto, o potencial de criar proteínas especiais para determinadas espécies, que podem ser responsáveis pela diferenciação entre elas. Uma das categorias dos éxons de processamento alternativo é realmente exclusiva dos primatas e pode ter contribuído para sua divergência em relação a outros mamíferos. Estudando o processo do nascimento de um éxon, podemos começar a enxergar as vantagens dos íntrons em geral, e a energia que gastamos para mantê-los parece se justificar.

Esses éxons específicos dos primatas derivam de elementos genéticos móveis chamados Alu, que pertencem a uma classe maior de elementos conhecida como retrotranspósons – seqüências curtas de DNA cuja função parece ser gerar cópias de si mesmas e depois reinseri-las no genoma em posições aleatórias, como pequenos parasitas genômicos. Os retrotranspósons estão presentes em quase todos os genomas, e foram profundamente importantes



CHIMPANZÉS E HUMANOS têm 99% dos genes em comum, e alguns deles contêm minúsculos elementos genéticos móveis chamados Alu, encontrados apenas nos primatas. Esses elementos se inserem dentro dos genes e podem ter dado origem, através do *splicing* alternativo, a novas proteínas que levaram à diferenciação dos primatas em relação a outros mamíferos. A diferenciação dos humanos em relação aos outros primatas também pode se dever, em parte, ao *splicing* alternativo: estudos recentes mostram que humanos e outros primatas produzem basicamente as mesmas proteínas na maioria dos tecidos, com a exceção do cérebro, onde os humanos fabricam um conjunto significativamente diferente de proteínas, e isso se deve em grande parte ao *splicing* alternativo dos genes

por terem contribuído para a expansão genômica que acompanhou a evolução dos organismos multicelulares. Quase metade do genoma humano é composta de elementos transponíveis, e os Alus são os mais abundantes.

Os elementos Alu possuem apenas 300 nucleotídeos de comprimento, com uma seqüência particular que termina em uma “cauda poli-A”. Mas nosso genoma contém cerca de 1,4 milhão de cópias Alu, e muitos desses elementos continuam se expandindo e se reinserindo em novos locais dentro do genoma, a um ritmo de cerca de uma nova inserção a cada 100 ou 200 nascimentos humanos.

Por muito tempo, os Alu foram considerados mero lixo genômico, mas ganharam um pouco de respeito quando os geneticistas perceberam que eles podem expandir a capacidade de geração de proteína do gene. Cerca de 5% dos éxons que sofrem *splicing* alternativo no genoma humano contêm uma seqüência Alu. É bastante provável que esses éxons tenham nascido quando um elemento Alu “pulou” para dentro do íntron de um gene, em local

onde a inserção normalmente não teria nenhuma consequência negativa para o primata, já que a maioria dos íntrons é removida e descartada. Através de mutações subsequentes, porém, o Alu conseguiu transformar o íntron no qual vivia em uma seqüência de informação genética com significado – um éxon. Isso pode acontecer se as alterações na seqüência Alu criarem um ponto de *splicing* 5’ ou 3’ dentro do íntron, fazendo com que parte dele seja reconhecida como “éxon” pelo spliceossomo. (Mutações como essa normalmente aparecem durante a divisão celular.)

Se o novo éxon Alu só for incluído no *splicing* alternativo, o organismo desfruta do melhor de dois mundos. Ao inserir o éxon Alu, suas células são capazes de produzir uma nova proteína. Mas a nova habilidade não interfere na função original do gene, porque os tipos antigos de mRNA ainda continuam a ser sintetizados quando o éxon Alu é deixado de fora no *splicing*. Apenas quando um Alu que sofreu mutação passa a ser constitutivo, ou seja, o éxon de Alu começa a ser incluído em todos os mRNA produzidos pelo gene, é que

a questão se torna problemática, porque pode desencadear doenças genéticas causadas pela ausência da proteína anterior. Até hoje foram identificadas três doenças genéticas provocadas por seqüências Alu inseridas em algum ponto do genoma: as síndromes de Alport e de Sly e a atrofia girata de coróide e retina.

Meus colegas e eu já demonstramos que basta a alteração de uma única letra na seqüência de DNA dos Alus para convertê-los de elementos silenciosos intrônicos a éxons reais. Hoje, o genoma humano contém aproximadamente 500 mil elementos Alu localizados dentro de íntrons, e 25 mil deles podem se tornar novos éxons se passarem por essa mutação em um único ponto. As seqüências Alu possuem, portanto, o potencial de continuar enriquecendo bastante o estoque disponível de informação genética significativa para a produção de novas proteínas humanas.

Terapia de RNA

CERCA DE 3 MIL CIENTISTAS no mundo todo tentam entender as complexas reações envolvidas no *splicing* alternativo. Embora essa pesquisa ainda esteja em estágio inicial, os investigadores concordam que as descobertas recentes apontam para futuras aplicações terapêuticas, como novas estratégias de terapia gênica que explorem o mecanismo de *splicing* para tratar tanto doenças hereditárias como adquiridas, a exemplo do câncer.

Uma das possíveis abordagens é fazer com que um trecho curto de nucleotídeos sintéticos de RNA ou DNA, chamados oligonucleotídeos "antisense" (complementares), se liguem a um alvo específico no DNA ou RNA do paciente. Esses oligonucleotídeos poderiam ser colocados dentro de células para mascarar um ponto específico de *splicing* ou alguma outra seqüência regulatória, transferindo, dessa forma, a atividade do *splicing* para outro local. Ryszard Kole, da Universidade da Carolina do Norte, demonstrou essa técnica em células progenitoras do sangue de portadores de beta-talassemia. Essa doença

Por que a evolução preservou um sistema tão complicado e que pode causar doenças?

hereditária ocorre quando um ponto de *splicing* 5' anômalo faz com que as moléculas de hemoglobina fiquem deformadas. Mascarando a mutação, Kole transferiu o *splicing* para o ponto normal, restaurando a produção da hemoglobina funcional.

Mais tarde, Kole demonstrou que a mesma técnica pode ser usada em células cancerosas humanas em cultura. Mascarando um ponto de *splicing* 5' na transcrição do gene Bcl-x, que controla a apoptose, ele conseguiu transferir o *splicing* e produzir a forma do mRNA Bcl-x(S), em vez da forma do Bcl-x(L), reduzindo a síntese da proteína antiapoptótica nas células e aumentando a síntese da proteína pro-apoptótica. Em algumas células cancerosas, essa alteração ativa o programa de morte celular; em outras, aumenta os efeitos apoptóticos da quimioterapia administrada junto com os oligonucleotídeos.

Outra maneira de usar o *splicing* alternativo em terapia foi mostrada em 2003 por Adrian Krainer e Luca Cartegni, do Laboratório Cold Spring Harbor, em Long Island, NY, que descobriram um modo de induzir as células a incluir um éxon que normalmente seria deletado. Eles criaram uma molécula sintética que pode ser programada para se ligar a qualquer trecho de RNA que combine com sua seqüência, e então anexaram a ela a parte da proteína SR que se liga ao RNA. Essa molécula química pode, portanto, ao mesmo tempo se ligar a uma se-

qüência específica do pré-MRNA e recrutar o mecanismo basal para o ponto adequado de *splicing*. Krainer e Cartegni usaram esse método em cultura de células humanas para corrigir defeitos no gene *BRCA 1*, envolvido no câncer de mama, e no *SMN2*, ligado à atrofia muscular espinhal.

Antes da publicação do genoma humano, poucos acreditavam que organismos tão complexos quanto os humanos pudessem sobreviver com meros 25 mil genes. Desde que o seqüenciamento foi concluído, o *splicing* alternativo surgiu como o processo fundamental para permitir que um pequeno número de genes produza uma coleção muito superior de proteínas necessárias para dar origem ao corpo e à mente humana, e para fazer com que essa produção seja orquestrada com precisão em tecidos diferentes e em momentos diferentes. Acima de tudo, o *splicing* explica como a tremenda diversidade entre humanos, camundongos e talvez de todos os mamíferos pôde se originar a partir de genomas tão semelhantes.

A evolução funciona apresentando novas opções aos organismos, e depois selecionando as vantajosas. Novas proteínas criadas pela inclusão de éxons derivados de Alus, portanto, provavelmente ajudaram a fazer do ser humano a espécie que somos hoje. E mais estudos sobre o *splicing* alternativo prometem melhorar nossa qualidade de vida. SA

PARA CONHECER MAIS

Alternative splicing: increasing diversity in the proteomic world. B. R. Graveley, em *Trends in Genetics*, vol. 17, nº 2, págs. 100-107, fevereiro de 2001.

Splicing regulation as a potential genetic modifier. M. Nissim-Rafinia e B. Kerem, em *Trends in Genetics*, vol. 18, nº 3, págs. 123-127, março de 2002.

Páginas ocultas no livro da vida. John S. Mattick, em *SCIENTIFIC AMERICAN*, ed. 30, págs. 54-61, novembro de 2004.

How did alternative splicing evolve? Gil Ast, em *Nature Reviews Genetics*, vol. 5, págs. 773-782, outubro de 2004.